# VALUES, PRACTICES, AND PHILOSOPHY OF SCIENCE

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

RAŞİT HASAN KELER

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
THE DEPARTMENT OF PHILOSOPHY

MARCH 2016

Approval of the Graduate School of Social Sciences

Prof. Dr. Meliha Altunışık
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Doctor of Philosophy.

Prof. Dr. Ş. Halil Turan
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Doctor of Philosophy.

Prof. Dr. David Grünberg
Supervisor

**Examining Committee Members**

| | | |
|---|---|---|
| Prof. Dr. David Grünberg | (METU, PHIL) | |
| Prof. Dr. Teo Grünberg | (METU, PHIL) | |
| Prof. Dr. Erdinç Sayan | (METU, PHIL) | |
| Assist. Prof. Hilmi Demir | (BİLKENT, PHIL) | |
| Assist. Prof. Cem Kamözüt | (EGE, PHIL) | |

**I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.**

Name, Last name: Raşit Hasan Keler

Signature:

# ABSTRACT

VALUES, PRACTICES, AND PHILOSOPHY OF SCIENCE

Raşit Hasan Keler
Ph.D., Department of Philosophy
Supervisor: Prof. Dr. David Grünberg

March 2016, 219 pages

The main aim of philosophy of science is to clarify the nature of science. I believe that this can only be achieved by specifying the values and practices of particular scientific theories and research traditions. Therefore this text is about practices of sciences and in particular, about what scientists and scientific communities prefer about and choose the theories, research programs, models, tools, data, hypothesis, theoretical frameworks, and all kinds of things they employ. What needs to be done is not to assume the values and practices of science, but rather to find them out. I give numerous examples from the sciences to exhibit different practices, values, and their connections.

After my discussion of practices and values, I move on to discussion of various philosophy of science issues. I believe that the root of many philosophy of science issues is a lack of consideration of practices and values of science. I show that a lot of these issues dissolve once we look at the issues informed by the practices and values. Discussed issues include the aim of science issue, objectivity and unbiasedness, rationality, underdetermination, confirmation, demarcation, and social constructivism.

**Keywords:** Philosophy of science, Values, Scientific practices.

# ÖZ

BİLİMİN DEĞERLERİ, PRATİKLERİ VE FELSEFESİ

Raşit Hasan Keler
Ph.D., Felsefe Bölümü
Tez Yöneticisi: Prof. Dr. David Grünberg

Mart 2016, 219 sayfa

Bilim felsefesinin temel amacı bilimin doğasını ortaya çıkarmaktır. Bunu başarmanın tek yolunun bilimdeki kuramların ve araştırma geleneklerinin değerlerini ve uygulamalarını belirlemek olduğuna inanıyorum. Tezimin konusu da bu nedenle bilimsel kuramların ve araştırma geleneklerinin değerleri ve uygulamaları. Özellikle bilim insanlarının ve bilimsel toplulukların kuramlar, araştırma programları, modeller, araçlar, veriler, hipotezler, ve kullandıkları her türlü şey hakkında neler tercih ettikleri ile ilgilidir. Yapılması gereken bilimdeki değer ve uygulamaları varsaymak değil, araştırıp bulmaktır. Bu tezde farklı uygulamalar, değerler ve bunların bağlantılarını sergilemek için bir çok örnek vereceğim.

Uygulamalar ve değerler hakkındaki bölümler ardından, çeşitli bilim felsefesi tartışmalarına eğileceğim. Ben bilim felsefesi sorunlarının çoğunun altında bilimsel uygulamalara ve değerlere yeterince dikkat edilmemesinin yattığını düşünüyorum. Bilimin bu yanlarını göz önüne alarak bir çok felsefi probleminin hakkından gelineceğini gösteriyorum. Bu problemler arasında bilimin amacı sorunu, nesnellik ve sapmasızlık, rasyonalite, kuramın veriler tarafından belirlenememesi, pekiştirme, tanım meselesi, ve sosyal oluşturmacılık var.

**Anahtar Kelimeler:** Bilim felsefesi, Değerler, Bilimsel uygulamalar

This thesis is dedicated to all short-haired women in the world.

# ACKNOWLEDGMENTS

I was lucky enough to have very good teachers who nourished my diverse and ever-changing interests:

- Ayhan Sol (Philosophy)
- Andreas Tiefenbach (Mathematics)
- Cem Tezer (Mathematics)
- David Grünberg (Philosophy)
- David Peirce (Mathematics)
- Erdinç Sayan (Philosophy)
- Mahmut Kuzucuoğlu (Mathematics)
- Samet Bağçe (Philosophy)
- Süleyman Önal (Mathematics)
- Teo Grünberg (Philosophy)

David Grünberg, Süleyman Önal, and Teo Grünberg have served as my academic advisors at different times and I cannot overemphasize their contributions to my personal and academic growth. Andreas Tiefenbach and Samet Bağçe are responsible for raising my interests in respectively mathematics and philosophy of science years ago when I was as an undergraduate student. Ayhan Sol's and Erdinç Sayan's captivating lectures that I now miss deeply were instrumental for my move to the philosophy department. The feedback they gave me over the years were invaluable. Without the knowledge, teaching, patience, friendship and guidance of all these people, I would not be who I am today.

Ayhan Sol, Cem Kamözüt, Genco Güralp, and Hilmi Demir read preliminary versions of my thesis and made very helpful and detailed comments. Erdinç Sayan suggested a large number of stylistic and linguistic improvements. To write this thesis I relied on the LaTeX knowledge that I learned from Andreas Tiefenbach years ago. My hat's off to all of them.

Even though I did not have time to get into the problems of social and cultural studies in this thesis, the talks on the subject I had with İnci Dirim broadened my perspective and brought various aspects of science that I had not thought about to my attention.

I have included a large number of images in this thesis. Many of these have some version of *Creative Commons* license which can be obtained from https://creativecommons.org/licenses/.

If available, I have given the licensing information about individual images below each one. I appreciate both the work of the image creators as well as the right holders who have given permission to use their work here.

I want to thank my friends for fun and games, as well as stimulating discussions.

My greatest gratitude is to my mother who supported and believed in me all these years.

# TABLE OF CONTENTS

# LIST OF FIGURES

FIGURES

# CHAPTER 1

# INTRODUCTION

According to Sidney Morgenbesser [1967, xiii], there are "five major goals of the philosophy of science" and these are, in his own words,

(1) to clarify the nature and aims of science;

(2) to specify the structure of particular scientific theories;

(3) to criticize and to comment critically on scientific claims in the light of epistemological and ontological theses;

(4) to assess claims about the possible reach of science;

(5) to buttress or test various epistemological theses on the basis of scientific results.[1]

My aim here belongs mainly to the first of the listed goals but I will also touch the second one quite frequently along the way. This text is about practices of sciences and in particular, about what scientists and scientific communities prefer about and choose the theories, research programs, models, tools, data, hypothesis, theoretical frameworks, and all kinds of things they employ.

I will take considerable time to investigate the values relevant in theory appraisal. These values are the key to answer questions such as: Why do scientists or scientific communities accept/reject the theories they accept/reject? What do scientists (dis)like about theories? What are the virtues of a good theory? How do scientists judge theories? Why do they choose a particular theory over another? Do the principles of theory choice differ from discipline to discipline? Do they change in time? Are there common trends? What are the relation of these values to each other?

These values are not only important in theory appraisal but also in theory construction. Scientists do not haphazardly come up with new theories, rather they try to come up with theories that have particular values.

I will *not* be restricting my attention to one type of thing (e.g. theories) used in science but will consider theories, models, tools, data, etc. Since writing "theories, models, tools, data, etc." time and time again is cumbersome, I will generally refrain from making the extension

[1] Following Peter Godfrey-Smith [2014, 4], we can add another item to the list: (6) to understand the world and our place in it in the light of the scientific results.

of my references explicit and use only one from the longer list. For example, the sentence "physicists prefer simple theories" as used by me implies that "physicists prefer simple models" among others. That is, what I am saying about theories, etc. at the time applies to more than that but I will seldom make this clear.

Of course, there are a number of ways the values are involved in science. For example, the code of conduct for scientists involves different values like honesty and integrity. Another involvement of values is the ethics of the use of science and technology. There are values present in deciding on what to study, aims of research, how to frame a question, how to allocate resources, how to publish and present results, how to experiment or carry on clinical trials, and so on. I will not look at these values unless they relate to appraisal of theories/data/models. From now on, I will use the word "value", in a restricted sense that only applies to appraisal unless explicitly stated otherwise. I use "virtue" as a synonym of "value".

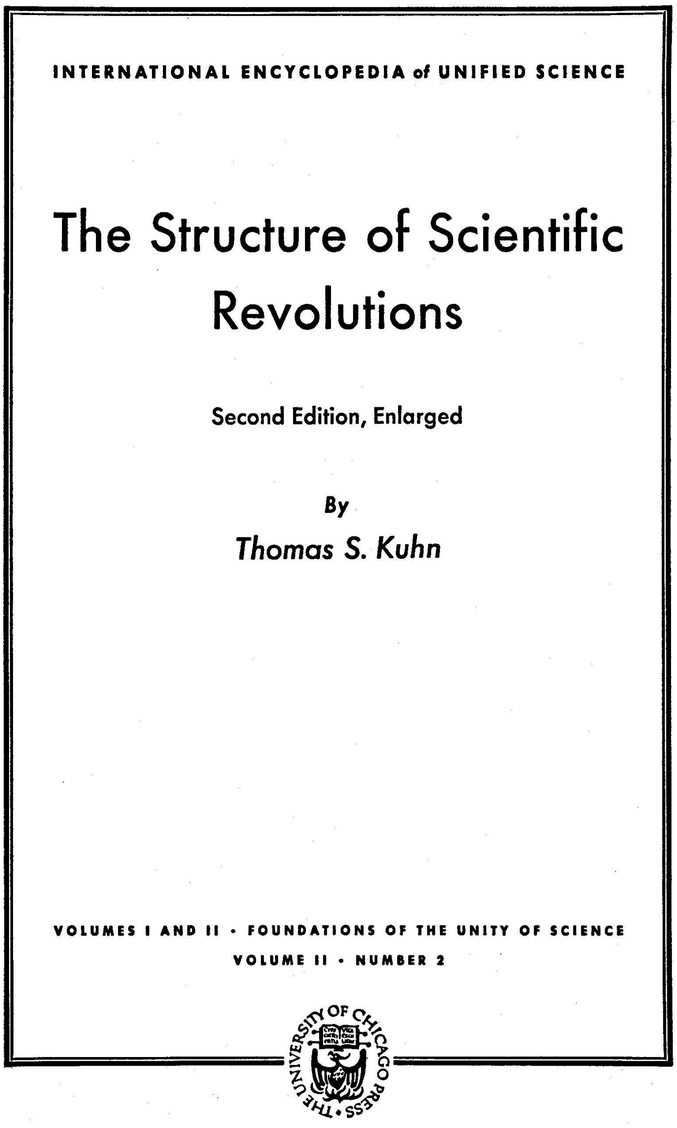
This limited interest in values allows me to restrict the use of the word "value" as well in the following way: In this text it is never an agent that has a value but always a theory/data/model, etc. that has it. I will talk about motivations/aims of scientists (or institutions) and I will talk of values of things (e.g. fruitfulness of research, simplicity of data and so on). This duality of language goes on to show that I am hardly interested in exhausting all value-centric problems in science.

The importance of values in theory choice in science was acknowledged to a limited degree in the pre-Kuhn era. But it was Kuhn's *The Structure of Scientific Revolutions* [1962] and his later article *Objectivity, Value Judgment, and Theory Choice* [1977a] which established the prevalence of values in science. (See figure 1.1.) Following the huge popularity of Kuhn, in the post-Kuhn era values are present in most discussions about science. But most of the writings on values actually focus on non-appraisal related topics. One might think that all that can be said about the role of values in appraisal has already been said; but the issues are far from settled.

> [W]hile it is clear that value-judgements play an important role in deciding how a given area of research should lead us to *act*, it is less clear whether such value judgements should play a role in deciding what to *believe*. In other words, while value-judgements clearly play a legitimate role in the realm of *practice*, do they also play a legitimate role in the realm of *theory*? [Biddle and Winsberg, 2010, 172]

I believe that practices in science along with values are important to understand science and the necessary attention to them is missing in philosophy of science. There is a need for a more thorough and detailed look at practices and values which I will fulfil in this text.

After my discussion of practices and values, I will move on to discussion of various philosophy of science issues. I believe that the root of many philosophy of science issues is a lack of consideration of practices and values of science. I will show that a lot of these issues dissolve once we look at the issues informed by the practices and values.



INTERNATIONAL ENCYCLOPEDIA of UNIFIED SCIENCE

The Structure of Scientific Revolutions

Second Edition, Enlarged

By

*Thomas S. Kuhn*

VOLUMES I AND II • FOUNDATIONS OF THE UNITY OF SCIENCE

VOLUME II • NUMBER 2

THE UNIVERSITY OF CHICAGO PRESS

Figure 1.1: Title page of the second edition of Thomas Kuhn's *The Structure of Scientific Revolutions* [1970]. According to Alexander Bird, "Thomas Samuel Kuhn (1922–1996) is one of the most influential philosophers of science of the twentieth century, perhaps the most influential. His 1962 book The Structure of Scientific Revolutions is one of the most cited academic books of all time. Kuhn's contribution to the philosophy of science marked not only a break with several key positivist doctrines, but also inaugurated a new style of philosophy of science that brought it closer to the history of science. His account of the development of science held that science enjoys periods of stable growth punctuated by revisionary revolutions. To this thesis, Kuhn added the controversial 'incommensurability thesis', that theories from differing periods suffer from certain deep kinds of failure of comparability." [Bird, 2013] Kuhn's views were so controversial at the time that there was a huge backlash against them. Ironically, the first two editions were part of the series *International Encyclopedia of Unified Science* associated with the logical empiricists. The second edition of the *Structure* contains a postscript in which he answers his critics and clarifies his position. Nevertheless, his views continued to be controversial to this day. [Image scanned by the author.]

## 1.1 How to Understand Science

> However, many [philosophers of science] now take interest in a more historical, contextual, and sometimes sociological approach, in which the methods and successes of a science at a particular time are regarded less in terms of universal logical principles and procedures, and more in terms of the then available methods and paradigms, as well as the social context.
>
> Simon Blackburn [2008, 330]

There are different avenues to investigate science and scientific activity. Historiographical studies, sociological investigations, rhetorical analysis, feminist studies, Popper's falsification and all kind of approaches claim to give us a better understanding of science. Imre Lakatos [1978b] sees different philosophical approaches to science to be competing with each other. They might be competing, but different approaches do not necessarily undermine each other but can result in a more complete understanding of science.

Consider for example, the approach of Lakatos [1978a] which is the *Methodology of Scientific Research Programmes* or *MSRP* for short. If I oversimplify MSRP, it involves (1) identifying the essential, indispensable part of a research program (its *core*), and (2) identifying various heuristics that protect this core and those that make possible the core to be applied in various ways. I think that MSRP is effective in untangling a research program. Knowing the core and heuristics leads to a solid understanding of a research program. Identifying the core of a research program would illuminate the nature of some controversies related to that program. For example, there are a number of issues (punctuated equilibrium, sympatric speciation, etc.) debated by evolutionary theorists that do not involve the core of the theory. But those who want to discredit evolutionary biology overemphasise these debates as if they challenge the theory. Only through having some understanding of the core of evolutionary theory it is possible to defend against these unfounded attacks by showing that they are not about the core, but rather side issues that do not threaten the fundamentals of the evolutionary theory.

But all methods, including MSRP, have their limits. MSRP's problem turns out to be that the core and the heuristics are not clear cut and open to interpretation. It might not be possible to divide a theory like that. It is more apt to think of a research program as a loosely related group of theories and practices.

Even if a Lakatosian research program existed in science, a Lakatosian analysis would not exhaust all the philosophically interesting things about science.

Here I suggest another approach, *value analysis*, to investigate science. This involves in identifying the particular values of a theory, contrasting these with general values, comparing with the values of other theories, and see what kind of practices and activities relate to these values. In particular, values of competing theories are to be compared.

Value analysis is not exclusive to analysis of science but it is useful in understanding all parts of social life. But here my interests lie in science and except for a few titbits all my examples come from science.

Like all methods, value analysis has its strong points and blind spots. I will mainly show the strong points in this text but mention some of its blind spots in passing as well. Since I

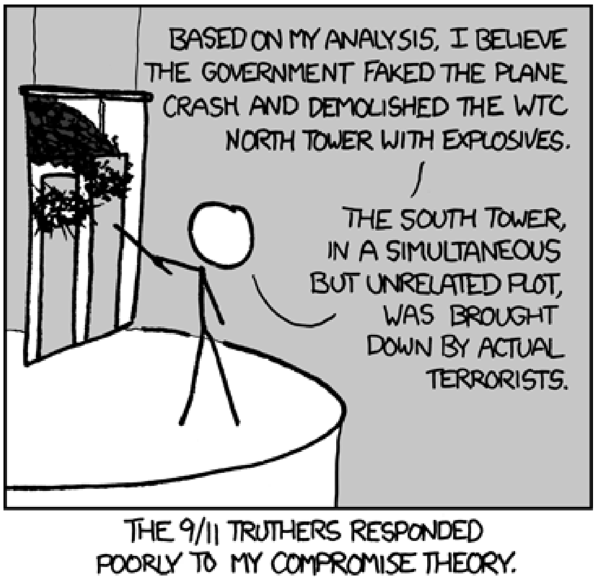



Figure 1.2: A comic by Randall Munroe showing what can go wrong with a middle ground theory. [Image from http://xkcd.com/690/ licensed under Creative Commons Attribution-NonCommercial 2.5 License.]

do not think that value analysis is the only game in town, the blindness can be overcome by other methods. I will even advocate another method in section 5.1.

There is more to using different approaches than combining their results. The output of a method can be an input for another. For example, value analysis can make use of historiographical and sociological studies to determine values (and I make use of them in this text). This is a recursive procedure and they can influence each other *ad infinitum*.

Philosophers should look at all these different analyses and combine them to understand science. I think that there has been a neglect of value analysis by both philosophers and sociologists of science. The former sees the only salient values to be evidential ones. The latter only accepts the so-called societal values. I do not think that this is an informed choice on their part. When they assume that just a single kind of value is operational, their research program turns to a lopsided enterprise. My account of science borrows much from the two camps but it is not a compromise between the two (see figure 1.2).

What needs to be done is not to assume the values, but rather to find them out. In this text I defend the view that we must look at science and its activities to find out about science. This may sound like a truism, but how can it be when there is only a handful of people following it? Truism or not, it needs to be defended and demonstrated.

Steven Shapin's [2010] book *Never Pure* has a delightful subtitle: *Historical Studies of Science as if It Was Produced by People with Bodies, Situated in Time, Space, Culture, and Society, and Struggling for Credibility and Authority*. This is the kind of science I am interested in — real, not hypothetical.

Then the aims of this thesis is to conduct value analysis, subject science to value analysis, stress the *let us look at how science is really done* viewpoint, and to follow the philosophical consequences of these views.

**Basic assumption of this text.** In any kind of research or study, there are always some philosophical underpinnings and beliefs accepted without any discussion.[2] There has to be, or else we would have to build our ground *ad infinitum*. In my case, it is the belief that sciences and their methods, practices, and values are so varied. Historian of science Naomi Oreskes [1999] makes a similar remark for *methods of science*:

> The scientific method, always in the singular, has been taken as monolithic and unproblematic — a textbook cliche, the one sure thing we all know about science. But is there such a thing as the scientific method? The answer is clearly no. From the past two decades of historical scholarship, one insight has emerged unequivocally: the methods of science are complex, variegated, and often local. Throughout history scientists have drawn on a wide variety of epistemic commitments and beliefs, linguistic and conceptual metaphors, and material and cognitive resources, all of which have changed with time and varied in space. At different times or in different social contexts, scientists have preferred either inductive or deductive modes of reasoning and argumentation, experimental or theoretical approaches to problem-solving, laboratory- or field-based methodologies. [Oreskes, 1999, 5]

I believe that research traditions in science differ so much from each other that it is impossible to pigeon-hole them into neat categories. Moreover, I believe that there is no way to know the methods, practices, and values of a tradition without actually examining them.

Since this is an underlying assumption of my text, I will not try to convince you of its truth though there is a chance that the short historical examples I give in the next chapter may sway you in my direction if you are even a little bit sympathetic to the assumption. If you still decline to believe my basic assumption at the end of the next chapter then we can agree to disagree.

## 1.2 A First Look at Values

> **value (n.)** circa 1300, "price equal to the intrinsic worth of a thing;" late 14c., "degree to which something is useful or estimable," from Old French *value* "worth, price, moral worth; standing, reputation" (13c.), noun use of feminine past participle of *valoir* "be worth," from Latin *valere* "be strong, be well; be of value, be worth". The meaning "social principle" is attested from 1918, supposedly borrowed from the language of painting. *Value judgment* (1889) is a loan-translation of German *Werturteil*.
>
> Online Etymology Dictionary

Here is a quick and partial list of candidate values that are considered by different philosophers:

simplicity, parsimony, explanatory simplicity, clarity, plausibility, influence of authority, familiarity, agreement with common-sense, fruitfulness, innovation, usefulness, problem solving effectiveness, beauty, mathematical form, symmetry, agreement with observation,

[2] Section 3.5 contains a discussion of underlying assumptions in science.

empirical adequacy, predictive power, explanatory power, making novel predictions, precision, ethics, politics, culture, religion, incentives, vested interest, future prospect, longevity, application, technology, control of nature, internal coherence, coherence with other theories, compatibility with other theories, dynamism, unification, implying other theories, mechanism, breadth, scope, pedagogical factors, psychological factors, visualisability, fear of implications.

Of course, only a small subset of these are really important in any particular case. What most of these mean should be clear from their name. Since this is a list of values I have written down from all kinds of sources, there are overlapping ones. I will refrain from explicating them or making a more principled list, but I need to clarify some of them.

To begin with, let me start with different versions of simplicity and parsimony. *Theoretical simplicity* (just *simplicity* from now on) is the simplicity of the theory: this favours the simplest theory that is consistent with all of the accepted observations, evidence and background assumptions. On the other hand, *parsimonious* theories consider nature itself to be simple. This is not simplicity of the theory, but of the properties of nature. *Ontological simplicity* is akin to Occam's razor and tells us to choose theories that make as fewest assumptions as possible.

I am using *dynamism* in the sense of Frank [1962, 352]: A theory is dynamic if it is "more fit to expand into unknown territory." This means that the theory can accommodate future developments of science or itself can be included in a future theory which is more general. Frank gives the following example: "Newton's laws originated in generalisations of the Copernican theory, and we can hardly imagine how they could have been formulated if he had started with the Ptolemaic system." Frank concludes that Copernican theory was the more dynamic one.

The values *agreement with observation/experiment, empirical adequacy, predictive power, making novel predictions* form what I call *evidential values*. When I say there is more evidence for a theory compared to an alternative, I mean that evidential values support that theory more.

The *societal values* are *ethics, politics, culture, incentives, agreement with group*, and so on. The societal values are not the only values that have social factors. In section 3.1, I will argue that almost all values are influenced by social factors. But I am calling only a small subset of values "societal values" because it is customary to name them so and also the social factors are much more apparent in these values.

*Repeatability* and *reproducibility* are virtues of experiments but they are different from each other. Repeatability is the consistency of measurements taken by a single person using the same instrument and experimental setting. On the other hand, reproducibility is the ability of recreating the experiment or study by another scientist independently. Sometimes reproducibility is said to be one of the main principles of the scientific method.

One important value that I neglect in this text is the *problem solving effectiveness* although I might be forgiven for this omission because there is already an important book on the subject, namely, Larry Laudan's *Progress and Its Problems* [1977]. But I believe that Laudan goes overboard with the importance he gives to this value and I will have a few words about it in chapter 5.

The above list is by no means exhaustive. There are one of a kind values peculiar to a particular tradition. For example, there is a school of thought which values writing incomprehensible, that is, to be taken seriously you have to be unintelligible. (See figure 1.3.) This goes

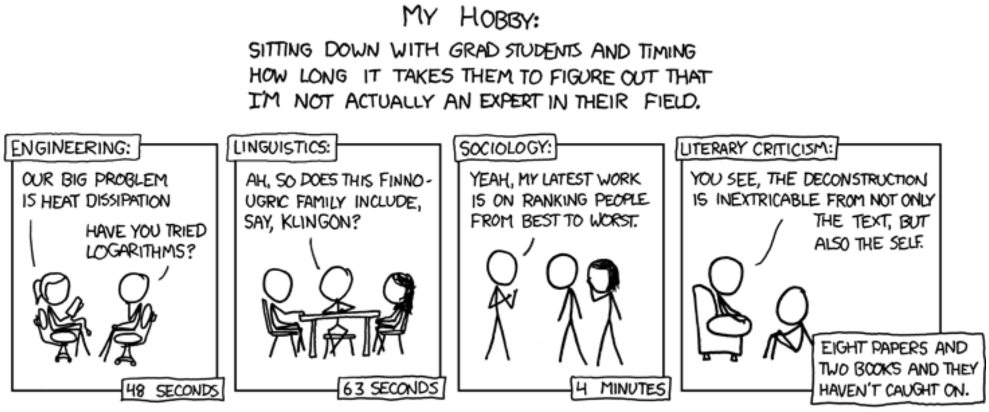



Figure 1.3: A comic by Randall Munroe poking fun at the jargon loaded language of literary criticism. [Image from http://xkcd.com/451/ licensed under Creative Commons Attribution-NonCommercial 2.5 License.]

on to show that a value may be found only in a few theories or disciplines, and you should not expect to have a complete list of values.

There is a caveat about most of the values: they are not for the uninitiated. For example, you need to be aware of physics and mathematics to appreciate the mathematical beauty and simplicity of a theory in theoretical physics. What a physicist finds beautiful in a theory can be quite different than that of a layman. (See figure 1.4)

**The value myth of positivism.** The following paragraph by the philosopher A. C. Grayling written in 2015 shows a typical view of values:

> When two hypotheses are equally adequate to the data, and equal in predictive power, extratheoretical criteria for choosing between them might come into play. They include not just questions about best fit with other hypotheses or theories already predicated to inquiry but also the aesthetic qualities of the competing hypotheses themselves — which is more pleasing, more elegant, more beautiful? — and of course the question of which of them is simpler. [Grayling, 2015, 9]

Grayling is not alone in believing these ideas. This conception of values is so widespread that I will call it the *value myth of positivism*. There are three components of the myth.

The first is the importance given to evidential support which indeed *is* one of the most important values in science. But the myth goes overboard with this idea putting it lonely at the peak. All other values are either non-existent or at best something to be avoided. As we shall see, non-evidential values are not an afterthought.

The second component of the myth is the view that one can decide the role of values by fiat, without actually doing any value analysis of a science. How does one claim that evidence is the prominent value? Through divine intervention? What about actually looking at science to see what is really going on? What the myth claims about science (and in particular about values) is not even based on a cursory look at science. What I call *let us look at how science is really done* viewpoint needs to be stressed and employed.

The third component is considering non-evidential values to be ideally eliminable from theory. On the contrary, non-evidential values are a legitimate part of theory appraisal, and

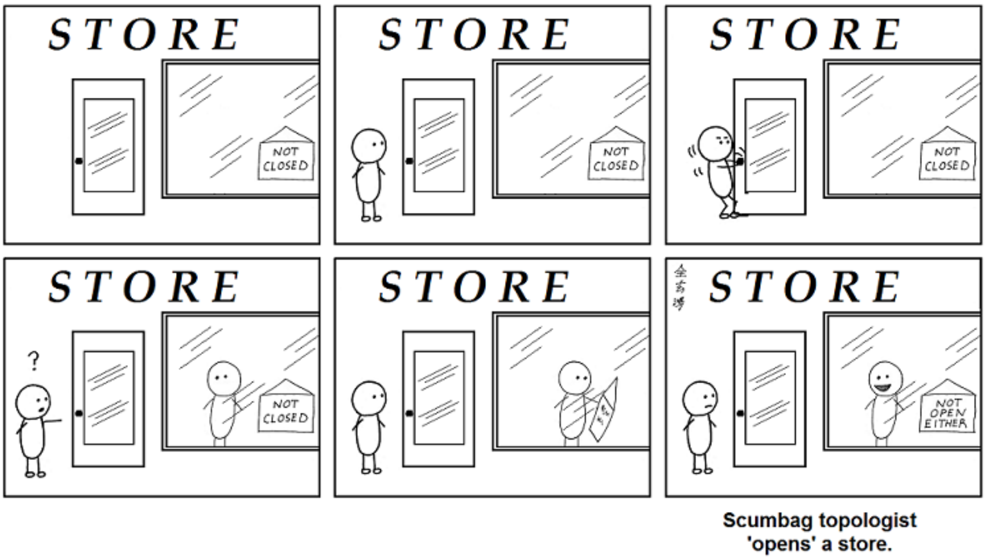



Figure 1.4: A comic from www.abstrusegoose.com/394. What a matematician/physicist finds funny can be quite different than that of a layman. All but the last pane show the window sign "not closed". On the last pane we see that the sign is changed to "not open either". This joke will be lost on you unless you know some topology. [Image licensed under Creative Commons Attribution-Noncommercial 3.0 United States License.]

as we shall see, they are not eliminable even in the ideal case.

To be honest, Grayling is defending a mild version of the myth — at least he acknowledges the presence of some values other than evidential ones. But even this mild version is far from capturing the role of values in science. They are pervasive in science and for good reason.

## 1.3 The Shape of Things to Come

> We write in outline, and necessarily in an elementary history it is only the primary lines that can be given.
>
> H. G. Wells, *The Shape of Things to Come*, 1933

The plan of my text is as follows:

Chapter 2 forms the heart of this text. Although there are works which analyse a single case study in depth, there are certainly not enough of them. My approach here is rather to choose breadth over depth and give a number of short case studies which serve the following purposes: (1) Each is a brief example that shows the related values and/or practices at work. (2) They exhibit different values and/or practices and their relationships. (3) A cumulative and intricate picture of values and practices emerges from them not available from a single example. (4) They prime the reader for different aspects of science that I will consider in later chapters.

Chapter 3 analyses the properties of values pertinent in science and elucidates the role values play in appraisal. In particular, I introduce the important and hitherto neglected distinction of general/concrete values and identify their usage, and, draw conclusions. I also discuss the factors that shape the values. Then I turn my attention to some value related issues discussed in philosophy of science such as societal values, theory-ladenness, etc. This chapter not only sheds light on the role and relationship of values, it also illustrates and discusses how philosophers should investigate values.

The chapter 4 is a look at opposing values which contains novel analyses of this topic. Different values clash regularly in science, and some more frequently than others. As far as I know, this chapter is the first cumulative look at this important aspect of values.

The understanding of values and practices that emerges from previous chapters is then used as a tool in chapter 5 to clarify some philosophical problems. I believe that there are a large number of issues in philosophy of science that would benefit a lot from looking closely at the actual practice of science, including the values at work. The chapter 5 gives a handful of these issues that show how long such an approach can go. I will show that a number of issues are amenable to such a practice/value oriented approach.

The last chapter concludes my discussion.

**A note about quotations.** Let me mention that unless I state otherwise , any emphases that exists always belong to the original author(s). For example, on page 25, I quote the following paragraph:

> In semiconductors, current conduction by holes is as important as electron conduction in general. *It is important to become familiar with thinking of the holes as mobile particles carrying positive charge, just as real as conduction electrons are mobile particles carrying negative charge.* [Hu, 2009, 5]

The italics in this quotation is due to the original author as I have not explicitly written something like "my emphasis" at the end.

Now, on to the heart of the text ...

# CHAPTER 2

# CASE STUDIES

In this chapter I will present some short pointers or case studies from different branches of science. All but one is from the twentieth century and most of them are still relevant today.

I have explained the aims of this chapter in section 1.3 an I will further discuss the methodology of case studies and its problems in section 3.2.

## 2.1 Forcing Conventions

> All mathematicians are familiar with the concept of an *open research problem*. I propose the less familiar concept of an *open exposition problem*. Solving an open exposition problem means explaining a mathematical subject in a way that renders it totally perspicuous. Every step should be motivated and clear; ideally, students should feel that they could have arrived at the results themselves. The proofs should be "natural" [i.e., lack] any ad hoc constructions or *brilliancies*. I believe that it is an open exposition problem to explain forcing.
>
> Timothy Y. Chow [2008]

There is one point about choices in science that I want to dispel early on: every choice in science signifies a serious involvement of values. Not true. Sometimes a choice is made just by convention or historical accident. Let me give an example from mathematics ignoring the technical details.

*Forcing* is a technique in set theory first used by Paul Cohen [1963] (see figure 2.1).

> Mathematicians have often regarded the history of mathematics as the history of great problems: Fermat's Last Theorem, the Riemann Hypothesis, the Continuum Hypothesis. But the history of mathematics is equally the history of great methods (such as the infinitesimal calculus) and their application to a wide range of problems. Forcing is such a method. Although Cohen invented forcing in order to settle two major problems (the independence of the Axiom of Choice and of the Continuum Hypothesis), forcing was quickly applied by logicians to a vast array of independence questions and, in addition, to other questions in set theory. Here the history of a great problem and the history of a great method are thoroughly intertwined. [Moore, 1988]

*THE INDEPENDENCE OF THE CONTINUUM HYPOTHESIS*

BY PAUL J. COHEN*

DEPARTMENT OF MATHEMATICS, STANFORD UNIVERSITY

*Communicated by Kurt Gödel, September 30, 1963*

This is the first of two notes in which we outline a proof of the fact that the Continuum Hypothesis cannot be derived from the other axioms of set theory, including the Axiom of Choice. Since Gödel[1] has shown that the Continuum Hypothesis is consistent with these axioms, the independence of the hypothesis is thus established. We shall work with the usual axioms for Zermelo-Fraenkel set theory,[2] and by Z-F we shall denote these axioms without the Axiom of Choice, (but with the Axiom of Regularity). By a model for Z-F we shall always mean a collection of actual sets with the usual $\epsilon$-relation satisfying Z-F. We use the standard definitions[3] for the set of integers $\omega$, ordinal, and cardinal numbers.

THEOREM 1. *There are models for Z-F in which the following occur:*

(1) *There is a set $a$, $a \subseteq \omega$ such that $a$ is not constructible in the sense of reference 3, yet the Axiom of Choice and the Generalized Continuum Hypothesis both hold.*

(2) *The continuum (i.e., $\mathcal{P}(\omega)$ where $\mathcal{P}$ means power set) has no well-ordering.*

(3) *The Axiom of Choice holds, but $\aleph_1 \neq 2^{\aleph_0}$.*

(4) *The Axiom of Choice for countable pairs of elements in $\mathcal{P}(\mathcal{P}(\omega))$ fails.*

Figure 2.1: Beginning of an article by Paul J. Cohen [1963] which established the independence of the continuum hypothesis and the axiom of choice from standard axiomatizations of set theory using a new technique called forcing which turned out to be very important on its own.

Forcing uses language of the form "q is stronger than p" where q, p are members of a suitable set. For our purposes, it does not matter what this relation is; what matters is that there are two different conventions to symbolize it: Israeli mathematicians write $q \geqslant p$ and the rest of the world write $q \leqslant p$. This little change of sign reverberates through forcing language; for example what is called *minimal* in one convention becomes *maximal* in the other. Set theorist Martin Goldstern explains the situation as follows:

> Traditionally, there are two (contradictory) notations for interpreting a partial order as a forcing notion. A majority of set theorists (including the books by Kunen and Jech) uses the "Boolean" or "downwards" notation, where $q \leqslant p$ means that q is "stronger" than p . . . , citing the universal agreement on the standard order of a boolean algebra or a lattice: A conjunction $p \wedge q$ is traditionally considered to be smaller than its constituents.
>
> The "Israeli" or "upwards" tradition (used not only by Shelah and some of his coauthors but also by Cohen in his original paper) expresses the same concept by $q \geqslant p$ (arguing that q has "more" information than p). [Goldstern, 1998, 72]

It is a testimony to the strength of the Israeli set theory school that their convention lives along the international one. Every graduate mathematics student specializing in set theory develops the ability to read both conventions and translate from one to other since he is bound to run into some set theory papers by Israeli set theorists.

In the above quotation, Goldstern points to the fact that there are different historical influences of each convention. One was chosen to extend the boolean algebra language already in use before the invention of forcing. The Israeli notation, on the other hand, follows

Cohen's initial account and motivates the notation in terms of information. This explains why each convention was created but it does not explain why each camp follows one but not the other. I believe that it boils down to the fact that once a few influential set theorists in close contact chose their convention, it spread to students and new-comers to the field forming an Israeli school. It was just a historical accident that those early adopters where Israeli but not Japanese.

Different conventions (in science or in our daily life) are always a nuisance to work with. Goldstern introduces a naming scheme to ease the pain: the letter that forces the other comes later in the alphabet:

> Whenever two conditions are comparable, the notation is chosen so that the variable used for the stronger condition comes "lexicographically" later. For example, we can have a condition $q$ which is strictly stronger than $p$, but we try to avoid the converse situation. Similarly, a condition called $p_2$ or $p_1'$ is allowed to be either stronger than $p_1$ or incompatible with $p_1$, but not (strictly) weaker. [Goldstern, 1998, 72]

This convention makes the transition easier but Goldstern [1998, 72] notes that it is not foolproof.

There are a few philosophically interesting points about this case.

(1) There are different schools in science and sometimes national boundaries shape them. Naturally, different communities have different needs, interests, and traditions which result in different schools. For example, we have the Polish logic school and the Japanese topology school. Most of the time the science of different national schools are compatible with each other as in the set theory case unless they support competing theories.[3]

(2) Not every choice in science is the result of serious theory-choice or battle of values. It can be by convention or historical accident. This is not to say that there might be different (dis)advantages of different conventions and I will look at such an example in section 3.5. But, really, whether you choose to write $q \geqslant p$ or $q \leqslant p$ for "q forces p" has no bearing on the results. One is a linguistic variant of the other. You get the same theorems in different notational conventions — the set theory is the same. This brings me to a more general point.

(3) The same theory could be expressed in different languages. This is why rhetorical analysis is (necessary but) insufficient method to understand science. For example, it is possible to criticize a theory for some insensitive language it uses but we should not jump to conclusions about the science itself because the same theory might be expressible in a decent language.

---

[3] For example, the English, Germans, and French tended to have different ideas about spontaneous generation controversy [see Farley, 1977]. I touch upon spontaneous generation in section 3.4.

## 2.2 The Zero Article

*"Who did you pass on the road?" the King went on, holding out his hand to the Messenger for some more hay.*
"Nobody," said the Messenger.
*"Quite right," said the King: "this young lady saw him too. So of course Nobody walks slower than you."*
"I do my best," the Messenger said in a sullen tone. "I'm sure nobody walks faster than I do!"
*"He can't do that," said the King, "or else he'd have been here first."*

Lewiss Caroll, *Through the Looking-Glass*, 1871

In this section I will give a simple though interesting example from linguistics which actually reminds me the dialogue from *Through the Looking-Glass* given above. (See also figure 2.2.)

Consider the following list of sentences from from BBC World Service's website [Woodham, 2011]:

- Formal education in Britain begins when children reach the age of five.
- Basketball is more popular in China than football.
- Patience and gratitude are qualities which are rarely observed in the youth of today.
- The education I received was substandard.
- The football played by Liverpool in their last match was awesome.

In the last two sentences, the word "the" is used in order to be specific. On the other hand, no "the" is used in the first three sentences because BBC tells us that "When we are discussing things in general, we normally use zero article with plural and uncountable nouns."

The *zero article* is a theoretical concept widely used in grammar. Similar to the BBC example, the following scenario is quite common in grammar texts: Rather than telling us that *no article is used* in a particular situation, the texts prefer to say that *a/the zero article is used* in that situation: "In generalisations we use *zero article*, but not *the*, with plural or uncountable nouns ..." [Hewings, 1999, 120] Lewis Carroll would have been proud of linguists usage of No-article as if it were Some-article. More examples from grammar texts:

> Zero article is found in certain meanings with plural count nouns. [Broughton, 1990, 89]
>
> The use of nouns of their own without an article is so fundamental in English that we should not regard this merely as 'the omission of the article', i.e. as something negative. We should think of the non-use of the article as something positive and give it a name: *the zero article*, which is usually given the symbol 0. [Alexander, 1988, 65]
>
> When no determiner occurs before the noun, it is useful to say that there is a zero article. [Biber et al., 2002, 67]

In an annotated edition of the Carroll's *Alice* books, Martin Gardner sheds light on the literary use:

Figure 2.2: An illustration by John Tenniel from *Through the Looking-Glass, and What Alice Found There* (1871) which is a novel by Lewis Carroll (Charles Lutwidge Dodgson), the sequel to *Alice's Adventures in Wonderland* (1865). The illustration depicts Alice, the King and the Messenger talking. [Image in the public domain.]

> Mathematicians, logicians, and some metaphysicians like to treat zero, the null class, and Nothing as if they were Something, and Carroll was no exception. In the first *Alice* book the Gryphon tells Alice that "they never executes nobody." Here we encounter the unexecuted Nobody walking along the road, and later we learn that Nobody walks slower or faster than the Messenger. "If you see Nobody come into the room," Carroll wrote to one of his child-friends, "please give him a kiss for me." In Carroll's book *Euclid and His Modern Rivals*, we meet Herr Niemand, a German professor whose name means "nobody." When did Nobody first enter the *Alice* books? At the Mad Tea Party. "Nobody asked *your* opinion," Alice said to the Mad Hatter. He turns up again in the book's last chapter when the White Rabbit produces a letter that he says the Knave of Hearts has written to "somebody." "Unless it was written to nobody," comments the King, "which isn't usual, you know."
>
> Critics have recalled how Ulysses deceived the one-eyed Polyphemus by calling himself Noman before putting out the giant's eye. When Polyphemus cried out, "Noman is killing me!" no one took this to mean that someone was actually attacking him. [Carroll, 2000, 223]

We should clearly add linguists to Gardner's list of "mathematicians, logicians, and some metaphysicians." The only grammatical function of the zero article is to highlight the absence of an article. Strictly speaking, the zero article is not something — there is no article used. It is the lack of an article. But why then, decade after decade, grammar text writers commit to the zero article ontology and students keep learning about the zero article? The reason for keeping the zero article in our ontology is similarly voiced by different writers:

> To our mind zeroes and gaps are far too much in syntactic analyses these days, but this one place where we actually think it may be defensible, because the fact that there is nothing there actually means something — the fact that the speaker doesn't use *the* means that the noun phrase is not definite. It is different from, say, not using an adjective — the fact that a speaker doesn't use *small* in a noun phrase doesn't mean that he thinks the thing referred by the noun question is big. He may just not have anything to say about size. If you don't like this argument, then just think of it as an absence of an article, rather than the presence of a zero article. We will, however, use the article from now on. One of our former colleagues in Manchester, Alan Cruse, has the following example of "meaningful nothingness". You have agreed with a friend that unless she phones you before six o'clock you'll meet up in the Hog's Head for a drink at eight. Now, if your phone does not ring before six, that is a meaningful nothingness, i.e. we would be prepared to let a zero element into our system. If, on the other hand, you have made no such arrangement, then the fact that your phone does not ring has no specific meaning and we would not want to represent as a meaningful element. [Börjars and Burridge, 2010, 172]

> When we choose the option of using the zero article, we are sending a specific message. We are signaling that the noun modified by the zero article is being used to make a generalization or a categorical statement about that noun. [Lester, 2013, 61]

Linguists keep the zero article in their toolbox in order to underline the fact that a generic reference or a reference to a non-specific class of things is made. The zero article is used to express information about a particular grammatical situation. It is a reminder to us that a generic or indefinite noun is used. Therefore it is kept just for *psychological* reasons.

This psychological tactic has been very successful and the class of *zero-marking* employed in linguistics is quite rich: zero pronoun, zero conjunction, zero preposition, zero copula, null morpheme, and so on. All zero-marking do their grammatical function by absence of word, prefix, or suffix. Their actual job is to signal us or to emphasize absence. And being such good reminders, we continue using zero-marking.

I suspect that there is a second virtue in play as well: it makes the relevant points more understandable and memorable for students. I can attest to this fact as a learner of English as a second language: when there is a name for a grammatical function, it is easier for me to remember it. So it is easier for teachers as well to teach the picture of *zero articles* than the picture with *no articles*. But I am just speculating about this second benefit here since I have not seen this aspect of zero-marking mentioned anywhere.

Therefore zero article is used for its psychological and possibly for pedagogical virtues. It draws our attention, highlights a particular use, makes easier to remember and teach. It is for these reasons that linguists forgo faithfulness; that is, they choose to work with the zero article rather than the alternative that there is no such thing; that alternative being the true nature of things.

One might think that the case of zero article is an oddity and in hard sciences one would never knowingly work with fictions as if they are real. Then one would be mistaken as we will see later in this chapter.

The case of zero article shows psychological values in isolation. Such values are usually a factor in all sciences but never as obvious as this case because most of the time there are other values at work as well.

## 2.3 A Tale of Two Rocks

> Abstract. Significant abundances of trapped argon, krypton, and xenon have been measured in shock-altered phases of the achondritic meteorite Elephant Moraine 79001 from Antarctica. The relative elemental abundances, the high ratios of argon-40 to argon-36 ($\geqslant 2000$), and the high ratios of xenon-129 to xenon-132 ($\geqslant 2.0$) of the trapped gas more closely resemble Viking data for the martian atmosphere than data for noble gas components typically found in meteorites. These findings support earlier suggestions, made on the basis of geochemical evidence, that shergottites and related rare meteorites may have originated from the planet Mars.
>
> D. D. Bogard and P. Johnson [1983]

Most meteorites are as old as Earth: 4.5 billion years. But there is a strange group of young meteorites called the *SNCs* (pronounced *snick*) named after the three subclasses of these type of meteorites: Shergotty, Nakhla, and Chassigny which in turn were named for the places where one of their type fell. (See figure 2.3.) SNC meteorites “were similar geochemically to terrestrial basalt and thus were from a parent body that had experienced complex melting


Figure 2.3: A Chinese stamp issued in 2003 depicting a meteorite shower

and crystallization through vulcanism similar to Earth's. But the SNCs were all thought to have crystallized only 1.3 billion years ago, long after the asteroids and the Moon had cooled enough for volcanic activity to end." [Dick and Strick, 2004, 181] SNCs were a great puzzle because they were crystallized at most 1.3 billion years ago and yet they exhibited melting and crystallization similar to the volcanic rocks found on Earth. This seems to rule out asteroids or the Moon as the sources of SNCs as they ceased their volcanic activity much earlier.

> In appearance and composition, the SNCs resemble certain basaltic rocks commonly found on Earth. That was perplexingly odd for meteorites that were supposed to come from asteroids, where conditions are most unplanetlike. On the other hand, the SNCs have some distinctive characteristics, such as their oxygen isotope composition, that set them apart from Earth and from any other meteorite. [Kerr, 1987, 721]

Then where did they come from? There was one suggestion of their origin:

> Rocks as young as the SNC meteorites had to have formed on a geologically active planet, and the most likely planet was Mars. The Mariner 9 and Viking Orbiter images had shown that Mars has enormous volcanos, up to 3 times as tall as Mauna Kea in Hawaii, and most of them could be as young as 1.3 billion years old. In 1979, a number of scientists seriously suggested that the young meteorites might be from Mars, but their ideas were met with great skepticism. [Dasch and Treiman, 1997]

The novel idea was that SNCs were from Mars — an impact on Mars ejected rocks off Mars and they made their way to Earth.

> The young ages of the SNC meteorites led some scientists to suggest that the SNC had to come from a body that was large enough to remain geologically active until at least 1.3 billion years ago, and perhaps as recently as 180 million years ago. The assumption was that the meteorites were blasted off their planet of origin by a large impact. But which planet or asteroid? Because the rocky planets remain hot in proportion to their size, meteoriticists argued that the SNC meteorites must come from a planet larger than the Moon. Earth was out, since several

> SNC meteorites were observed to have blazed through the atmosphere, and their oxygen isotopes are distinctly different from those of the Earth and Moon. Venus was not likely because its thick atmosphere would impede escape of impact ejecta. Mars was the best bet.
>
> The idea was not embraced enthusiastically, especially by scientists who study the dynamics of the impact process. They argued that there was no way of getting a meteorite off Mars without melting the rock, and there was no evidence for impact melting in the Martian meteorites, although the effects of impact short of melting were evident in many SNC meteorites. In fact, they argued, it was not possible to eject a rock from the Moon without melting it. [Taylor, 1999]

The idea that SNCs are from Mars contradicted the widely accepted background assumptions of the scientific milieu. Forget Mars, not even a lunar meteorite was thought to be possible: "The conventional wisdom had been that an impact energetic enough to splatter debris at the escape velocity of the moon (2.5 kilometers per second) would also melt the debris or at least crush it to a powder." [Kerr, 1983, 288] The Martian origin of any meteorite was thought to be highly unlikely. Kerr [1983, 288] mentions a "psychological barrier to the idea that meteorites can originate on large bodies" and considers two scientific developments in 1983 that were influential on the removal of the psychological barrier.

One of the developments involved a meteorite called ALH81005 that was found in 1982 in the Allan Hills region in Antarctica. Toshiko Mayeda and Robert Clayton found in 1983 that the oxygen isotope composition of ALH81005 matched that of lunar rocks brought back from the moon in the Apollo missions. The establishment of the lunar origin of this meteorite opened the floodgates. As Dick and Strick [2004, 181] put it, the "conceptual barrier to accepting the idea of intact escape of a rock from a planetary-sized body had been broken." If it was possible for a lunar meteorite to escape moon, the same could be true even for a Martian one. "This sample [ALH81005] was of historic significance not only because it was the first lunar meteorite, but it became a great piece of evidence in favor of dynamic arguments that fragments of the Moon and Mars could be delivered to the Earth after being ejected from their parent bodies during an impact event." [Righter and Gruener, 2007] The discovery of the lunar origin of ALH81005 is the topic of a newspaper article reproduced here as figure 2.4. The article does a surprisingly good job at explaining the importance of the discovery and the scientific details.

Even though the lunar origin of ALH81005 provided a "psychological boost" [Kerr, 1987, 721] to Martian origin hypothesis of SNCs, there was another development in 1983 that provided more direct support of the hypothesis. One of the SNCs, the meteorite EETA79001 (see figure 2.5), was discovered in 1979 at the Elephant Moraine region of Antarctica. In 1982 NASA scientists Donald Bogard and Pratt Johnson analyzed the noble gas composition of some of the trapped gas in EETA79001. The opening quotation of this section is the abstract part of their paper [1983] in which they announced that the results were consistent with the noble gas composition of the Martian atmosphere which was sampled in NASA's Viking program to Mars. They tentatively stated that "The presence of trapped martian atmosphere in EET79001, if corroborated by further studies on nonnoble gases, would be a particularly strong argument for the idea that this meteorite is from Mars." [Bogard and Johnson, 1983, 221] In an interview

# Scientists say meteorite is likely piece of the moon

Chicago Sun-Times

A meteorite discovered last year in Antarctica is a piece of the moon, a team of scientists, including two University of Chicago researchers, have concluded.

Robert Clayton, a University of Chicago chemist, said Sunday the meteorite might have been knocked off the lunar surface long ago by the impact of a large meteor on the moon.

The finding lends support to the theory that several other rocks found on earth may have come from Mars. It also is the first time scientists have been able to firmly establish the specific origin of a meteorite.

Most meteorites that land on earth are believed to be debris left from the formation of the solar system or parts of asteroids that broke apart.

So the discovery last year in Antartica of the meteorite — about the size of a golf ball — created a sensation among planetary scientists. A preliminary examination of the one-ounce, light green—crusted rock by Brian Mason of the Smithsonian Institution suggested it is similar to rocks from highland areas of the moon.

Rock samples then were distributed, for further confirmation, to scientific teams, which included Clayton and University of Chicago chemist Toshiko Mayeda.

Their contribution was to analyze the rock's composition of isotopes of oxygen, a technique they used in 1973 to identify, for the first time, matter that originated outside the solar system. Planets, asteroids and meteorites have different proportions of the isotopes, and the isotope signature of the lunar meteor matched those of the earth and the moon.

Clayton said further analysis by another researcher, Donald Bogard of the Lyndon B. Johnson Space Center in Houston, showed the material could not have originated on earth. The rock contained large amounts of helium, neon and other elements that result from the bombardment of the solar wind — the continuous flow of gases from the sun. That can occur on the moon but not on earth because of the earth's atmosphere.

The scientists' excitement over the discovery is not so much over possession of a moon rock. After all, more than 800 pounds of lunar rocks were brought back by Apollo astronauts and about 90 percent still have not been studied.

Rather, the lunar meteorite proves rocks ejected from a moon or even a planet can be transported across interplanetary space by natural phenomena.

The biggest question is how the rock survived the tremendous heat generated by the impact with earth's atmosphere and how it remained intact, Clayton said.

Figure 2.4: An article from Spokane Chronicle dated March 16, 1983 reformatted to two columns.

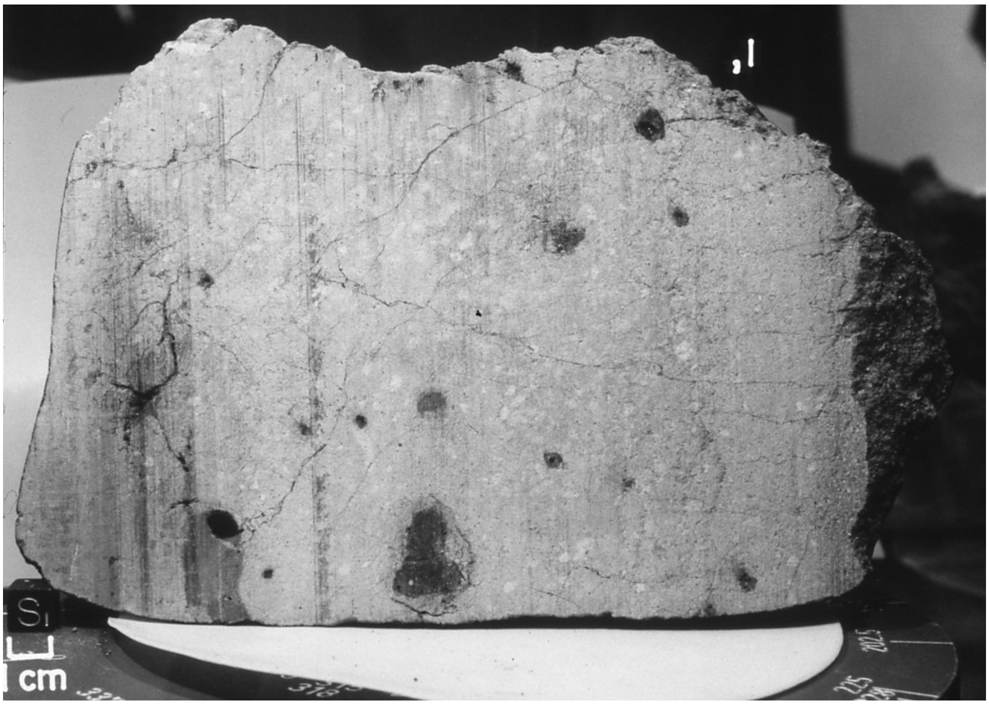



Figure 2.5: Meteorite EETA79001. According to Dasch and Treiman [1997], "This meteorite, which originally weighed nearly 8 kilograms, provided the first strong proof that meteorites could come from Mars. EETA 79001 is an achondrite meteorite, a basalt lava rock nearly indistinguishable from many Earth rocks. This picture shows a sawn face of this fine-grained gray rock (the vertical stripes are saw marks). The black patches in the rock are melted rock — glass — formed when a large meteorite hit Mars near the rock. This meteorite impact probably threw EETA 79001 off Mars and on its way to Antarctica on Earth. The black glass contains traces of martian atmosphere gases." [NASA image S80-37631 in the public domain.]

with Derek W. G. Sears, Bogard reflects on his work of 1982 and the atmosphere of the time:

> In spite of my earlier skepticism, I came to the conclusion that these data really did look like the Mars Viking results. Here was my . . . chance to argue for a meteorite from Mars, and I was still hesitating. I remember that [geophysicist] Jay Melosh said you can't get meteorites off Mars. Theorists had said that a force sufficient for ejection would destroy the rocks. You have to remember this is before the first lunar meteorite, this is 1982, so I was a little nervous[.] . . .
>
> I gave my talk in October 1982. I wrote a paper for *Science*. *Science* couldn't decide what to do with it. . . . The paper got mostly positive reviews, but *Science* really didn't know what to do. They wouldn't accept it, they wouldn't reject it. I gave a very similar talk at the March 1983 meeting of the LPSC. Immediately afterwards a *Science* reporter came up and was enthusiastic about the work and wanted me to submit a paper to *Science*! I said, "You've had it for 4 months now and you won't tell me what you are going to do with it"! They accepted it immediately after that. It was published. [Sears, 2012, 426–427]

Bogard and Johnson's work "brought sudden respectibility, if not credibility, to the suggestion of a Martian origin." [Kerr, 1983, 289] The years 1982–1987 witnessed a change in the status of the Martian origin hypothesis of SNCs. Richard A. Kerr wrote two reports of this field [1983; 1987] in *Science* and the titles of his articles clearly indicate this change: whereas the 1983 article is titled *A Lunar Meteorite and Maybe Some from Mars*, the 1987 one is titled *Martian Meteorites are Arriving*. Kerr [1987, 721] quotes geochemist Michael Drake commenting on the Martian hypothesis: "It's probable, but not proven; it's not likely to be incorrect. But short of going to Mars, no one will be absolutely convinced." Kerr [1987, 721] adds that "the passage of time has made a martian origin an acceptable hypothesis." Further work on SNCs made the hypothesis adopted by the whole scientific community: "With continued study of meteorites and collection of new ones, more were recognized to be of the SNC class, and their Martian origin was more and more widely and certainly accepted." [Dick and Strick, 2004, 182–183] Martian meteorites had arrived and they were indeed here to stay.

What does this episode teach us?

(1) To begin with, the common-sense or background assumptions are not easy to shake off. The belief that a big enough rock cannot survive an impact on Mars intact and escape Mars to reach Earth was so strong that the geochemical evidence was shelved for a time with the expectation that another explanation would come. Agreement with background or common-sense theory is a strong value in science and scientists are hesitant to argue against it.

(2) The "psychological barrier" was only laid aside when the search for an alternative theory consistently failed. Lack of alternatives that kept the background theory still viable eventually caused it to be laid aside. Scientists realized that "no alternative seemed as appealing as the large impact [hypothesis]." [Kerr, 1987, 721]

(3) This episode shows that scientists can hold on to a theory even if seeming contradictory evidence surfaces with the hope that further work can reconcile them. Naive falsifications cannot be found in science.

(4) Naturally, when the Martian origin of SNCs were eventually accepted, the background theory gave way and the new puzzle was to explain how the background theory could fail, that is, how can a rock from Mars reach Earth:

> Although the geochemical case for Martian meteorites has become stronger and stronger, the problem of getting them off Mars remains a major obstacle. Mars's escape velocity (5 kilometers per second) is twice that of the moon. In addition, measurements of shergottite cosmic-ray exposures require that a single fragment ejected from the parent body later shattered to form the individual meteorites. The original object had to be at least 10 meters in diameter. Dynamicists can get gas, liquid, or dust off Mars easily enough, but house-size boulders are another matter — the energy required for escape seems to be always greater than the energy sufficient to destroy such large boulders. [Kerr, 1983, 289]

What started as a background assumption turned out to be false and showing how this could be kept the impact and orbital dynamicists busy for the next thirty years.

(5) It is important to highlight one aspect of this further work. When the background theory was given up and it was accepted that meteorites from Mars can end up in our backyards, scientists did not know how this could be. They did not know any mechanism that can achieve this miracle. The search for such a mechanism started but, most interestingly, not for an *actual* mechanism but rather for a *possible* mechanism. The topic of research was not *how did these rocks from Mars reached Earth intact?* but rather *how could these rocks from Mars reach*

*Earth intact?* As geophysicist Vickery put it: "This is a plausibility argument. Until there is a lot more information about the surface of Mars, this is about as well as we can do." [Quoted in Kerr, 1983, 289]

(6) There is a further point this case can show us. Consider the following quotation detailing some work carried out in 1982–1983, paying closer attention to emphasized words than the technical details:

> Bogard and Johnson found that the shock trapped neon, argon, krypton, and xenon in the same relative abundances as calculated for the rocks of Earth and Mars. Most other meteorites have no trapped gases, and those that do have about ten times as much xenon as the Elephant Moraine shergottite. Even more impressive, the extracted argon had a ratio of argon-40 to argon-36 as high as 1750. *Corrected* for the amount of argon-36 thought to have been produced by cosmic rays, the ratio climbed to 2040. Earth's atmospheric argon has a ratio of 300, and the ratio for the Martian atmosphere is somewhere between 2225 and 3500, depending on who *interpreted* the data.
>
> In a late paper, Richard Becker and Robert Pepin of the University of Minnesota confirmed the *uncorrected* ratio of about 1800, but they based their *correction* on analysis of adjacent, nonglassy rock that had the same chemical composition and *presumably* the same exposure to cosmic rays. Their *correction* raised the ratio to 2400, within the range of reported Martian ratios. How an impact on an airless asteroid could trap Mars-like gases is not clear.
>
> Becker and Pepin's measurement of the nitrogen isotopes trapped in the same meteorite was received a bit more tentatively. The ratio of extracted nitrogen-15 to nitrogen-14, even after a *correction* for nitrogen in the unshocked part of the meteorite, was only +130 per mil in the standard notation of isotopic ratios. The value for nitrogen in the Martian atmosphere is $620 \pm 160$ per mil. To test the Martian origin hypothesis, Becker and Pepin made *another correction* based on the assumption that there was more nitrogen in the minerals that formed the gas-containing glass than found elsewhere. They took as a measure of that excess the relative sizes of the nitrogen-argon ratios in the glass and in the Martian atmosphere. This *second correction* raised the value to +500 per mil, which is within the Martian range. Most listeners viewed the necessity of a second correction as regrettable but took some reassurance from the high nitrogen-argon ratio of the meteorite, which lies between that of Earth and Mars. [Kerr, 1983, 289, emphases added]

There is an old (and hopefully dead) idea in philosophy of science that observation is separate from theory. The claim is that evidence bears on theory but not reverse: evidence stands on its own and weighs on theory. Contrary to the claim, the above short quotation exemplifies how theoretical considerations effect observations. Measurements are not taken in vacuo but they are interpreted and corrected according to theory. To put it another way, evidence is *theory-laden*. I will return to this aspect of evidence/data in section 3.4.

Even though my discussion of this episode ends here, I will follow one further Martian debate in section 2.6.

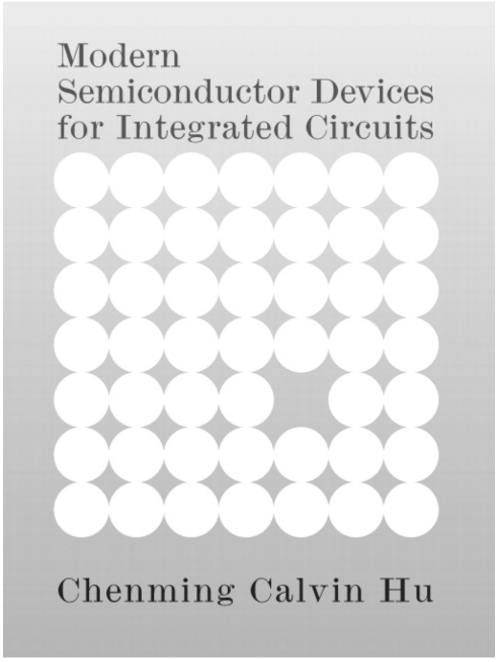



Figure 2.6: The cover of Chenming Calvin Hu's *Modern Semiconductor Devices for Integrated Circuits* [Hu, 2009] designed by Jason Hu. [Courtesy of Chenming Calvin Hu.]

## 2.4 Holes in Semiconductors

A nonexistent object is something that does not exist.

Maria Reicher [2014]

In physics, a complicated many-particle system can present some phenomena which in some respects can be regarded as if it is exhibited by a single particle. A complex behavior of a collection of particles is sometimes best described by introducing an imaginary particle and attributing the behavior in question to this new particle. Such particles are called *quasiparticles*. (Actually, there are two types of such phenomena: *quasiparticles* and *collective excitations*. According to Kaxiras [2003, 68], quasiparticles are related to fermions, and collective excitations are related to bosons; but this distinction is not needed for our purposes.) Quasiparticles are emergent phenomena of a microscopically complex system. So a quasiparticle is an imaginary entity that actually corresponds to some characteristics of a grouping of multiple particles.

> Quasi-particles are no proper kinds of entities at all — they are merely collective effects of (an indeterminate number of) "real" entities, and they must be acknowledged as illusory entities even from a non- or anti-theoretical position . . . [Gelfert, 2003, 246]

I will present a simplified account one type of quasiparticle: *(electron) holes*.[4] In a semiconductor crystal, the *valance band* is the nearly filled electron band. Electrons of this band move around leaving voids behind. It turns out that, both conceptually and mathematically, it is easy to deal with the few number of holes that electrons leave behind rather than the huge number

[4] See [Hu, 2009] or any solid state physics textbook for a faithful account.

of electrons themselves. The cover of Hu's book [Hu, 2009] given in figure 2.6 beautifully depicts the hole concept. White circles represent electrons moving around in the valance band. The missing circle, or the hole, completely specifies the whole electron assembly. The cover also emphasizes the importance of holes in modern semiconductor theory by reserving the first image seen by the readers to them. This importance is reflected in the text as well:

> An alternative way to think of this process is that the hole moves to a new location. It is much easier to think of this second means of current conduction as the motion of a positive hole than the motion of negative electrons moving in the opposite direction just as it is much easier to think about the motion of a bubble in liquid than the liquid movement that creates the moving bubble.
>
> In semiconductors, current conduction by holes is as important as electron conduction in general. *It is important to become familiar with thinking of the holes as mobile particles carrying positive charge, just as real as conduction electrons are mobile particles carrying negative charge.* [Hu, 2009, 5]

Solid state physicists bring these holes to the forefront and treat these holes as if they are real particles. They attribute physical properties like charge and mass to them. Holes *cause* various kinds of physical phenomena. They turn up in all kinds of electronics and they are used in physical explanations. Physicists talk holes, use holes, predict using holes, explain using holes, experiment with holes, and so on. Listening to solid state physicists talk, you would be hard-pressed to discern that these are fictitious particles.

The 1956 Nobel Prize in Physics was awarded to William Shockley, John Bardeen, and Walter Houser Brattain (figure 2.7) "for their researches on semiconductors and their discovery of the transistor effect." Shockley's book *Electrons and Holes in Semiconductors with Applications to Transistor Electronics*, printed in 1950, became the textbook and the primary source of the new field of transistors for at least a decade. The first chapter of this book features a conceptual introduction to holes which is a pleasure to read even today. The title of the book reflects the important experimental role played by the holes. Shockley's preface to this book starts as follows:

> The hole, or deficit produced by removing an electron from the valence-bond structure of a crystal, is the chief reason for existence of this book. Although the hole and its negative counterpart, the excess electron, have been prominent in the theory of solids since the work of A. H. Wilson in 1931, the announcement of the transistor in 1948 has given holes and electrons new technological significance. From the theoretical viewpoint, the hole is an abstraction from a much more complex situation and the achieving of this abstraction in a logical way appears inevitably to involve rather detailed quantum-mechanical considerations. From the experimental viewpoint, in contrast, the existence of holes and electrons as positive and negative carriers of current can be inferred directly by the experimental techniques of transistor electronics so that holes and electrons have acquired an *operational* reality in Bridgman's sense of the word. Furthermore, the new experiments have established the quantitative aspects of the behaviors of holes and electrons with sufficient accuracy for many of the purposes of transistor electronics. [Shockley, 1950, ix]

It is important that Schockley refers to operational reality of the holes. Holes are only *mathematical entities*; they are nothing more than a mathematical trick. But they have turned out to

Figure 2.7: From left to right: John Bardeen, William Shockley, and Walter Brattain in 1948. [Image in the public domain from http://commons.wikimedia.org.]

be so useful and they feature prominently in experiments in which their physical properties can be measured.

Incidentally, holes and other quasiparticles are used by Axel Gelfert [2003] to argue against Ian Hacking's *entity realism*. Gelfert convincingly argues that quasiparticles pass Hacking's *manipulation criterion* (if you can use an entity to manipulate other parts of nature regularly, then it must be real), yet, they do not exist. Hence quasiparticles form a class of counterexamples to entity realism.

To sum up, the main reasons why physicist keep quasiparticles in their bag is mathematical simplicity as well as their usefulness in expressing experimental results.

Mathematical entities will be reconsidered in section 2.11.

## 2.5 Dispensable Mathematics

> But physics is not mathematics. Physicists work by calculation, physical reasoning, modeling and cross-checking more than by proof, and what they can understand is generally much greater than what can be rigorously demonstrated. ... Physicists by their methods can obtain new results whose mathematical underpinning is not obvious.
>
> Joseph Polchinski [2007]

In the last section we saw how physicists sacrifice faithfulness for simpler mathematics: instead of working with electrons, they prefer working with mathematically simple but fictitious holes. But sometimes even mathematics itself is sacrificed for simplicity: in this section

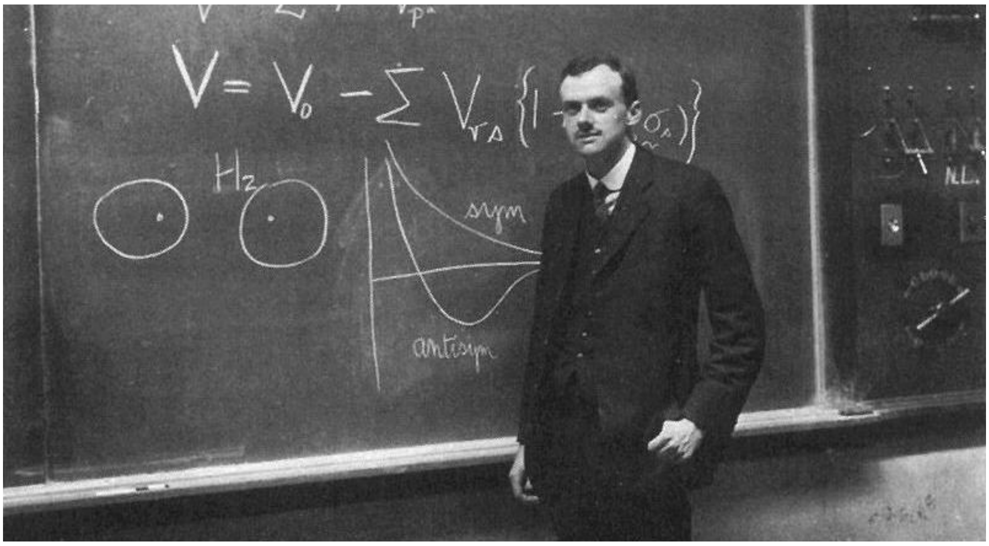


Figure 2.8: Paul Dirac at the blackboard circa 1930, when he published his influential book *The Principles of Quantum Mechanics*. [Image in the public domain from wikimedia.org.]

I will give two examples of physicists deliberately using mathematically false statements to make their theory work.

**Dirac's delta function.** The first example is Dirac's delta function $\delta$.[5] Even though $\delta$ did not originate[6] with Paul Dirac (see figure 2.8), this function bears his name following his use of it in his highly influential book *The Principles of Quantum Mechanics*[7] published in 1930. This is how Dirac introduces $\delta$:

> [W]e introduce a quantity $\delta(x)$ depending on a parameter $x$ satisfying the conditions
>
> $$\int_{-\infty}^{\infty} \delta(x)\,dx = 1$$
> $$\delta(x) = 0 \quad \text{for } x \neq 0.$$
>
> To get a picture of $\delta(x)$, take a function of the real variable $x$ which vanishes everywhere except inside a small domain, of length $\varepsilon$ say, surrounding the origin $x = 0$, and which is so large inside this domain that its integral over this domain is unity. ...Then in the limit $\varepsilon \to 0$ this function will go over into $\delta(x)$.
>
> $\delta(x)$ is not a function of $x$ according to the usual mathematical definition of a function, which requires a function to have a definite value for each point in its domain, but is something more general, which we may call an 'improper

[5] See [Bueno, 2005] for a more detailed philosophical investigation of Dirac's delta function.

[6] For a history of delta function and distributions see [Lützen, 1982] as well as its review [Dieudonne, 1984].

[7] My page references are not to the 1930 edition but rather to the fourth edition published in 1958 which is still in print. There are textual differences between these two editions but none that impacts the philosophical aspects discussed here.

function' to show up its difference from a function defined by usual definition. [Dirac, 1958, 58]

Dirac then uses $\delta$ to give an elegant, powerful, and general account of quantum mechanics. The delta function has a central role in this account. The problem is that $\delta$ is not mathematically well-defined and has contradictory consequences.[8] The contradictory nature of this function is more apparent in the version of it as given in modern physics and engineering texts:

$$\delta(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } x \neq 0 \end{cases}$$

and satisfies the identity

$$\int_{-\infty}^{\infty} \delta(x)\, dx = 1.$$

In this definition, the only non-zero value of the function is at the origin, yet it has a non-zero integral.

So how come physicists use such a contradictory function? Dirac argues that the use of the function is inessential:

> [I]t should be possible to rewrite the theory in a form in which the improper functions appear all through only in integrand [where they are not problematic]. The use of improper functions thus does not involve any lack of rigour in the theory, but is merely a convenient notation, enabling us to express in a concise form certain relations which we could, if necessary, rewrite in a form not involving improper functions, but only in a cumbersome way which would tend to obscure the argument. [Dirac, 1958, 59]

According to Dirac, the delta function is in principle dispensable from his account. At the time he wrote his book, there was no mathematical way of doing so. It was the theory of *general functions* (also called *distributions*) developed in the next decades by Sergei Sobolev and Laurent Schwartz which provided the mathematical foundation. Dirac chose the simplicity of using the delta function even though it was mathematically unsound. Bueno [2005] remarks that it is even now pragmatic to use the delta function rather than its mathematical counterpart distributions:

> In fact, the introduction of the theory of distributions increases hugely the complexity of quantum theory as well as the size of the function space for QM [quantum mechanics]. Not surprisingly, introducing distributions ends up "obscuring the argument", in pretty much the way Dirac warned. There's something to be said for Dirac's strategy after all. [Bueno, 2005, 471]

Dirac [1958, vii] in the preface of his book emphasizes that mathematics is only a tool and his choice to use the delta function is a reflection of this approach. The delta function is useful and leads to a simple account. Thus it is much convenient to employ $\delta$ rather than a mathematically sound form.

Dirac's approach is generally contrasted to that of John von Neumann. Being one of the greatest twentieth century mathematicians, von Neumann's mathematical sensitivities

[8] See [Bueno, 2005, 466].

sided with rigour. The following quotation is from the preface of von Neumann's 1932 book *Mathematische Grundlagen der Quantenmechanik*. I am using the translation given by Smoryński [2012, 179–180].

> The object of this book is the standard, and, so far as possible and reasonable, mathematically objection-free presentation of modern quantum mechanics, which in the course of recent years has acquired in its essential parts an expectedly definitive form: the so-called "theory of transformations" …
>
> In several discourses, such as in his recently published book, Dirac has given a presentation of quantum mechanics, which … is scarcely to be outdone in brevity and elegance. Thus it is perhaps appropriate to supply here a few arguments for our methodology, which is essentially different from that named.
>
> The methodology of Dirac referred to, which because of its transparency and elegance has inundated a large portion of the quantum mechanical literature, in no way comes up to the requirement of mathematical rigour — also not, if moreover this is to be reduced naturally and fairly, to the usual norm of theoretical physics. So, for example, in consequence of holding on to the fiction that every self-adjoint operator can be brought into diagonal form, which in fact is not the case for these operators, the introduction of "improper" functions with self-contradictory properties is necessitated. Such an inclusion of mathematical "fictions" is inevitable under the circumstances, if it is only a matter of numerically calculating the result of an intuitively defined experiment. This would be no objection, if these concepts which are unsuitable in the framework of analysis were really essential for the new physical theory. Just as Newtonian mechanics first gave rise to the development of an infinitesimal calculus unquestionably self-contradictory in its form at the time, the quantum mechanics would suggest a new construction of our "analysis of infinitely many variables" — i.e. the mathematical apparatus would have to be changed, not the physical theory. That is, however, in no way the case; rather it should be shown, that the theory of transformations can also be mathematically unobjectionably founded in just as clear and standard manner. Thereby it is required that the correct construction not consist of a mathematical precisioning and explication of Dirac's methods, but rather that it makes necessary from the outset a different approach, namely the dependence on Hilbert's spectral theory of operators. [von Neumann, 1932, 1–2]

These issues are also connected to more recent physics:

> An ongoing debate in the foundations of physics concerns the role of mathematical rigor in theorizing. The contrasting views of von Neumann and Dirac provide interesting and informative insights concerning two sides of this debate. Von Neumann's contributions often emphasize mathematical rigor and Dirac's contributions emphasize pragmatic concerns. … The entry quantum field theory provides an overview of a variety of approaches to developing a quantum theory of fields. The purpose of this article is to provide a more detailed discussion of mathematically rigorous approaches to quantum field theory, as opposed to conventional approaches, such as Lagrangian quantum field theory, which are

> generally portrayed as being more heuristic in character. [Kronz and Lupher, 2012]

I will not investigate these matters further here.

**The sum of natural numbers.** The second example of unsound use of mathematics I want to mention can be summed up by the identity:

$$1 + 2 + 3 + 4 + \cdots + n + \cdots = -\frac{1}{12}.$$

String theorists are lately using the above preposterous identity. There is no misprint here. On the left hand side of the identity you have the sum of all natural numbers and on the right hand side you have *negative* $1/12$. There are many mathematical methods (e.g. zeta function regularization) to assign finite values to divergent series. But this does not change the fact that the series in question is divergent.[9]

Tong [2009, 39], in the first edition of his book, uses the word "heuristic" for the identity and in the third revised edition [Tong, 2012, 42] he writes that "it's actually a very useful trick for getting the right answer."

Unfortunately, the related physics is beyond my comprehension, and I refer the reader to [Tong, 2012, chapter 4] for a technical exposition. But even without understanding the technical details, what should be clear is that physicists are willing to use suspect (and in this case seemingly absurd) mathematics if it leads to a simple theory.

**Philosophy of mathematics.** I cannot help but make a quick digression about philosophy of mathematics. Consider the following version of the indispensability argument:[10]

> Quine and Putnam have argued that the indispensability of mathematics to empirical science gives us good reason to believe in the existence of mathematical entities. According to this line of argument, reference to (or quantification over) mathematical entities such as sets, numbers, functions and such is indispensable to our best scientific theories, and so we ought to be committed to the existence of these mathematical entities. . . . Moreover, mathematical entities are seen to be on an epistemic par with the other theoretical entities of science, since belief in the existence of the former is justified by the same evidence that confirms the theory as a whole (and hence belief in the latter). This argument is known as the Quine-Putnam indispensability argument for mathematical realism. [Colyvan, 2015]

If there is an inference from the use of mathematics in physics to the truth or reality of mathematics, then it would be true that the sum of all natural numbers is equal to $-\frac{1}{12}$. Such an inference is absurd. If anything, *good* mathematics is dispensable in physics. Physicists use mathematics as they see fit and this has got nothing to do with the ontological status of mathematical objects.

---

[9] See https://en.wikipedia.org/wiki/Divergent_series.

[10] There are different versions of the indispensability argument and here I consider the one that serves my purpose. Putnam [2012] discusses different versions and in particular he differentiates his indispensability argument from the one attributed to him by Colyvan quoted below.

Physicists treat mathematics as a toolbox of useful techniques and do not hesitate to take the tools suitable for the job at hand disregarding others.

## 2.6 Martian Life

> What I would like to do this afternoon is lead you through our story, which is a bit of a detective story, on why we think we have found evidence for past life on Mars. ... We believe, we interpret that these are indeed microfossils from Mars. They are extremely tiny, the longest one is about 200 nanometres, this is very high magnification. One of the techniques that we're using, by the way, is high-resolution scanning electron microscope. We're looking at rocks and minerals at a scale that has really not been used before. These are extremely high-magnification, high-resolution pictures. Next slide please.
>
> David McKay, NASA Press conference, 7 August 1996

*Has there ever been life on Mars?* One might think that the method to answer this question is straightforward: Look at Martian rocks until you find evidence of past life. Of course the availability of Martian rocks is a problem, though not a philosophically interesting one. The more salient problem is to decide what counts as evidence of past life. This problem became acute in August 1996 when a group of scientists announced that they have found such evidence in a Martian rock called ALH84001.

ALH84001 is a meteorite of size $17 \times 9.5 \times 6.5$ centimetres and weighs 1.9 kilograms. A comet or asteroid impact 16 million years ago on Mars ejected ALH84001 off Mars and it landed on Earth about 15,000 years ago. It was found on the white snow of the Allan Hills area of Antartica (hence its name) in 1984 and its Martian origin was established in 1993 by analysing isotopic composition of oxygen it contains. In the 16 August 1996 issue of the *Science* magazine, David S. McKay and his co-authors tentatively claimed that some features of ALH84001 point to "fossil remains of a past martian biota". Their claim made big news and led to a lively scientific debate that still goes on today. What features of ALH84001 suggest past life?

> The lines of evidence which indicate possible biogenic activity in the Martian meteorite ALH84001 are: (1) the presence of carbonate globules which had been formed at temperatures favorable for life, (2) the presence of biominerals (magnetites and sulfides) with characteristics nearly identical to those formed by certain bacteria, (3) the presence of indigenous reduced carbon within Martian materials, and (4) the presence in the carbonate globules of features similar in morphology to biological structures. [Gibson et al., 2001, 16]

> David McKay's team gleaned four lines of evidence from this meteorite, which collectively convinced them that Mars once supported microbial ecosystems. The carbonate minerals in ALH-84001 (1) resemble terrestrial deposits formed where bacteria are active, (2) contain distinctive grains of the iron oxide mineral magnetite that compare closely to magnetite crystals formed inside bacterial cells, (3) preserve complex organic molecules thought to be derived from bio-molecules, and (4) harbor tiny round and rodlike structures interpreted as microfossils. [Knoll, 2003, 229]

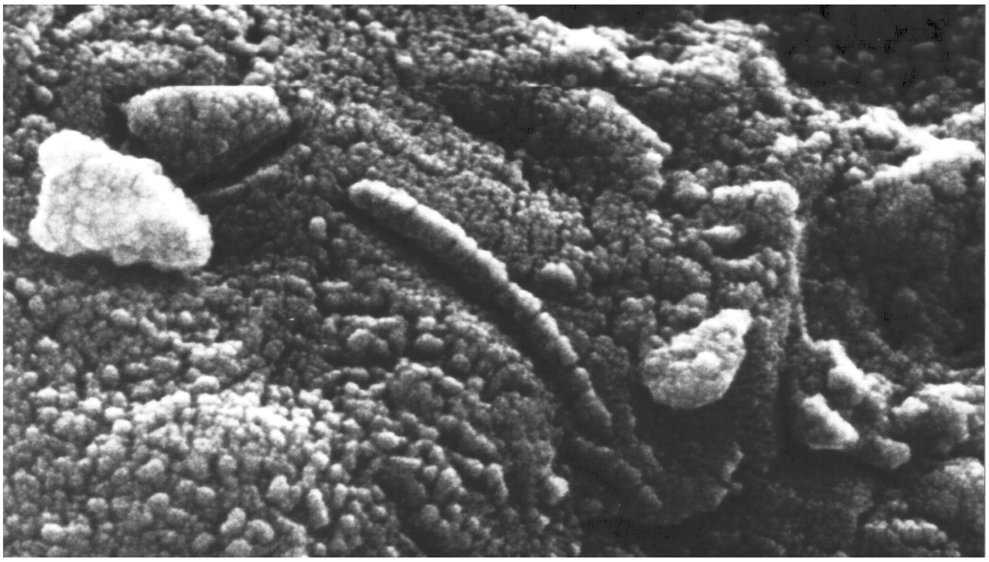


Figure 2.9: A scanning electron microscope image of the meteorite ALH84001. [NASA image S96-12609 in the public domain.]

A close-up photograph of one of the mentioned (possible) microfossils is in figure 2.9. The rod in the centre of this image is 0.2 micrometers long and resembles some very small bacteria found on Earth.

The scientific debate revolves around whether the best explanation of the above quoted evidence is terrestrial or Martian. The analysis and discussion of ALH84001 is still continuing and the jury is still out on Martian life. My aim here is not to weigh on evidence, but rather to highlight an interesting feature of data interpretation. There are two levels when dealing with data: looking at individual trees and looking at the whole forest. The first level is exemplified by the following view:

> This brings us to the crux of astropaleontological interpretation. We can accept the morphological or chemical patterns in rocks as biological only if they make sense in terms of known biological processes and are unlikely to be made by purely physical mechanisms. That's the rule on Earth ... and it is the rule elsewhere in the solar system. [Knoll, 2003, 232]

This is a principle that says that we should give precedence to physical and chemical processes over biological ones when trying to explain features of rocks. As I defined in section 1.2 this is a *parsimony* principle. Such principles tell us to attribute simple properties to nature. Here physical and chemical processes are taken to be simpler than biological ones and we are asked to prefer them.

The second level is exemplified by the reason why McKay et al. [1996] accepted the meteorite ALH84001 as evidence for past life on Mars:

> It is possible that all of the described features in ALH84001 can be explained by inorganic processes, but these explanations appear to require restricted conditions ... None of these observations is in itself conclusive for the existence of past life. Although there are alternative explanations for each of these phenomena taken individually, when they are considered collectively, particularly in view of


Figure 2.10: A comic from www.abstrusegoose.com/463. [Image licensed under Creative Commons Attribution-Noncommercial 3.0 United States License.]

> their spatial association, we conclude that they are evidence for primitive life on early Mars. [McKay et al., 1996, 929]

They also repeated their position five years later:

> Each of these phenomena could be interpreted as having abiogenic origins but the unique spatial relationships indicated that, collectively, they recorded evidence of past biogenic activity within the meteorite. [Gibson et al., 2001, 16]

While they accept that each individual feature of ALH84001 can be non-biological, the totality of evidence is best explained by the biological hypothesis. This is what I called the value of *(theoretical) simplicity* in section 1.2. What simplicity mandates is to look at all evidence and come up with the simplest theory that explains them; not simplicity of *nature* as in parsimony; but simplicity of the *theory*. "From the beginning of work on the paper, the team members realized that none of their lines of evidence was conclusive by itself; all had ambiguities that allowed for an abiotic explanation as readily as a biogenic one." [Dick and Strick, 2004, 186] McKay's team do not argue that their theory is parsimonious, rather they argue that the biogenic theory is the simplest theory that explains the overall evidence. They do not consider the evidence in isolation but highlight the importance of all of evidence seen together.

> These lines of evidence were not simply to be considered in an additive fashion, they argued; because so much of the independently suggestive molecules all existed in the carbonate globules or their immediate vicinity, the presumption of all having been caused by biogenic activity in that locale was strengthened in a synergistic way. This "spatial association" argument was important: a large number of observers were willing to dismiss the case out of hand based on each of the lines considered separately because in not one of those cases had the team shown the biogenic explanation to be significantly more persuasive than one or more abiotic explanations. Many skeptics who said they still kept an open mind on the question said it was the spatial association argument that gave them pause. [Dick and Strick, 2004, 187–188]

Even though the overall simplicity argument for the Martian life seemed promising, the next few years show that the scientific community at large was against it. This was partly due to the fact that re-evaluation and further analysis of the four types of evidence resulted in strengthening of abiotic explanations for at least three types [see Kerr, 1998]. Also, scientific community mostly weighed on the side of parsimony rather than simplicity. Reporting on a conference on Martian Meteorites in 1998, Kerr writes that


Figure 2.11: Tunisian stamp issued in 1964 commemorating African micropaleontology conference. *Micropaleontology* is the branch of palaeontology that studies morphology and characteristics microfossils. After the ALH84001 debate, the norms in micropaleontology regarding the size limits of very small microorganisms and the evidence of life were revised.

> Even 2 years ago, many researchers were unimpressed with that holistic argument. "I never bought the reasoning that the compounding of inconclusive arguments is conclusive." says petrologist Edward Stolper of the California Institute of Technology in Pasadena. And it was clear at the workshop that now, as pieces of the argument weaken, it is losing its grip over the rest of the community. [Kerr, 1998, 1400]

Though there is a consensus, the debate about the Martian life is yet to end and there is some further work that gave "new vigor" to the debate. [Dick and Strick, 2004, 198] Astrobiology is live and kicking. (See figure 2.10.)

The current astrochemical and astrobiological research reflects the values parsimony and simplicity at work. While scientists look for chemical processes that can explain the individual features, they also change perspective to look at all evidence as well.

What are the philosophically interesting aspects of the Martian life controversy?

(1) The first aspect is obviously the clash between simplicity and parsimony. The clash between these two values is not unique to this episode and we will see it again. I will analyze the interaction of these two values in section 4.1.

(2) There is one common idea expressed by both sides of the debate and that is how fruitful the debate has been. One might think that before the debate started we did not know if there was Martian life and we still do not know — there has been no progress. Actually, the scientific progress has been tremendous. The understanding of signs of life, techniques, and technology has greatly increased. The debate has catalyzed a flurry of research that improved the biochemical knowledge and methods available. "Whatever the outcome on nanobacteria per se, the Mars meteorite claim does seem to be driving crucial parts of the science of exobiology forward." [Dick and Strick, 2004, 195]

(3) There is an another feature of science this episode show us: the role of risk in theory

appraisal. The more risky or important the hypothesis is, the more energy spent to evaluate it. The Martian life hypothesis is so contentious that ALH84001 caused a huge controversy and its features were and still are examined meticulously. But it turns out that scientists did not show the same thorough and careful attention to detail analyzing "fossils" from Earth — some of these were accepted hastily as signs of past life. The norms of evaluating signs of life has changed after ALH84001 debate and scientists are using the new norms to re-evaluate the past attributions of life. [Dick and Strick, 2004, 199–200] (See figure 2.11.) This shows us that various pragmatic values such as importance has a bearing on the acceptance of hypotheses.

## 2.7 Nuclear Models

> The essence of knowledge is generalisation.
>
> Hans Reichenbach [1951, 5]

Scientific theories have their scope or domain of use: they are applicable in certain areas and this changes in time. The hope is that a theory widens its domain but sometimes the opposite happens. For example, once thought to be universal, Newtonian mechanics lost its generality first to electrodynamics and then to twentieth century physics. If there are related theories with different domains, the hope is to find a theory that subsumes them.

Sometimes, this hope takes form as trying to find a general formula. Before 1900, there were different formulas (Wien's law, Rayleigh-Jeans law) for black-body radiation that worked for different ranges of wavelengths and there was a search for a general formula that worked at all wavelengths. It was Max Planck's early quantum work in 1900 resulting in the Planck's law (figure 2.12) which settled the issue. See [Segrè, 2007, 66–77].

Sometimes it is two distinct theories that scientists want to unify. For example, there are a number of technical and conceptual incompatibilities[11] between quantum theory and general relativity and there are different research programs[12] trying to merge them. The motivation for this activity is the importance given to generality and unification by the physicists:

> [W]e can happily maintain a schizophrenic attitude and use the precise, geometric picture of reality offered by general relativity while dealing with cosmological and astrophysical phenomena, and the quantum-mechanical world of chance and intrinsic uncertainties while dealing with atomic and subatomic particles. Clearly, this strategy is quite appropriate as a practical stand. But it is highly unsatisfactory from a conceptual viewpoint. Everything in our past experience in physics tells us that the two pictures we currently use must be approximations, special cases that arise as appropriate limits of a grander theory. That theory must therefore represent a synthesis of general relativity and quantum mechanics. This would be the quantum theory of gravity. [Ashtekar, 2005, 2064]

It is not always two different theories that needs to be married, rather a fragmented theory itself. The following paragraph is from Robert B. Leighton's classical *Principles of Modern Physics* published in 1959:

---

[11] See [Rickles, 2008].

[12] See [Ashtekar, 2005] for a history of the efforts to join quantum and relativity theories and in particular page 2072 for a list of different programs.

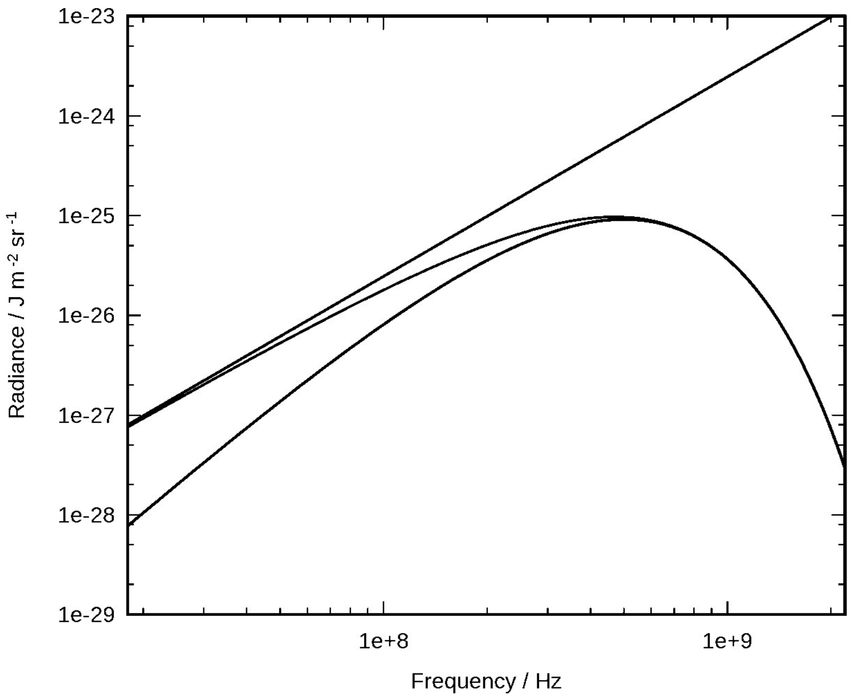



Figure 2.12: Log-log plots of radiance vs. frequency for Rayleigh-Jeans law (top, straight looking), Wien's law (bottom, reverse U shaped), and Planck's law (middle, approaching one of the others on each side) for a black body at 8 mK temperature. [Image adapted from https://en.wikipedia.org/wiki/File:RWP-comparison.svg by sfu licensed under Creative Commons Attribution-Share Alike 3.0 Unported.]

> In nuclear physics we seem at present still to be rather far from a simple complete nuclear theory; for this reason our concepts of nuclear structure are rather crude and the range of applicability of any given model of the nucleus is correspondingly rather restricted. Indeed, it sometimes has seemed that a separate model is needed to describe each nuclear property, and models which appear quite satisfactory within their designed range of applicability sometimes appear to contradict one another when used to describe certain other nuclear phenomena. In this chapter we shall outline very briefly some of the nuclear models that have been applied with more or less success toward the description of various features of nuclear structure. [Leighton, 1959, 595–596]

Even though particle physics has come a long way since 1959, this problem about nucleus has not been resolved today and there are a multitude of models. According to nuclear physicist/chemist Romualdo de Souza, the difficulty of giving a nuclear model stems from two principal problems:

(1) There is no exact mathematical expression that accounts for the nuclear force, unlike the atomic case, for which the electromagnetic force is well-defined by Coulomb's Law.

(2) There is no mathematical solution to the many-body problem, a limitation shared by both nuclear and atomic systems. [de Souza, 2014, Section 8]

There are a number of nuclear models but for our purposes it is enough to consider two of them. De Souza introduces the models as follows:

> The starting point for theoretical models of the nucleus treats the problem from two divergent perspectives: At the macroscopic extreme is the Liquid Drop Model, which examines the global properties of nuclei, such as energetics, binding energies, sizes, shapes and nucleon distributions. This model assumes that all nucleons are alike (other than charge). In contrast the Shell Model is designed to account for the quantal properties of nuclei such as spins, quantum states, magnetic moments and magic numbers. The basic assumption of the Shell Model is that all nucleons are different, i.e. nucleons are fermions and must occupy different quantum states, as is the case for atoms.
>
> The idealized goal of theoretical nuclear physics is to combine these two concepts into a Unified Model that will describe both the macroscopic and microscopic aspects of nuclear matter in a single comprehensive framework. [de Souza, 2014, Section 8]

The liquid drop model takes nucleus to be similar to a drop of a uniform incompressible liquid: "the chemical analogy would be a droplet of composed of nonpolar molecules such as $CCl_4$ or isopentane held together by Vander Waal's attraction" [de Souza, 2014, Section 8]. This analogy allows liquid physics to be adapted for nucleus. For example, just like a liquid drop, nucleons on the surface is different from those in bulk of the liquid and considered to have something like a surface tension. Nucleus would then have a volume term and a surface term. The binding energy is analogous to the mass of the drop. The model carries over other concepts as well, but of course there are "significant differences between a classical liquid drop and a nucleus which must be accounted for in the model" [de Souza, 2014, Section 8].

The shell model, on the other hand, is analogous to the atomic shell model which describes distribution of electrons of an atom or molecule. "The approach is analogous [to] that for the hydrogen atom model for periodic behavior in chemistry. The principal difference is that for atoms the electron-electron force is repulsive, whereas for nucleons the force is attractive." [de Souza, 2014, Section 9] This analogy makes use of energy quantization similar to the atomic energy levels. Just as filled shells results in greater stability in an atom, additional stability of some nucleus is explained in terms of shell closures at certain "magic" nucleon numbers.

These models are philosophically interesting for at least the following reasons.

(1) Both models serve well to a degree to enlighten nuclear structure but these approaches have their limits. The importance of scope comes into play at this level of models as physicists try to improve each model. Some of these shortcomings can be overcome by modifying or refining the models; see for example [de Souza, 2014, Section 9, p. 12] and [Leighton, 1959, 603–605]. The hope is to eventually give a unifying model of nuclear structure. This may be a completely new model or an improvement of a current model. Until then, physicists work with contradictory, insufficient, and cramped models. Lessons learned from these models may directly or indirectly help to form a general theory.

(2) Some models feature collective excitations (related to quasiparticles) but I will not touch on this topic as I have dealt with this topic in section 2.4.

(3) The models show an important and widespread methodology in physics: employing analogies. Physicists apply successful models of the past to shed light on novel situations. This may require analogical reasoning. For example, in the liquid drop model, the nucleus is considered *as if* it is a drop of a liquid. This shows us that physicist do *not* always take their theories literally and they show great flexibility in their conceptualization of their subjects if

it is useful to do so. I will investigate analogies in section 3.3.

(4) Finally, let me mention in passing that some of these models are semi-empirical in nature: some terms in these models are experimentally determined. Physicists ideally prefer to derive their parameters from theory rather than to measure them experimentally. One of the reasons why physicists look for an improvement of the standard model is the large number of parameters in it. The hope is that a grand unified theory will make these parameters predicted away. But as long as they lack a theoretical way of doing so, they use the models with their semi-empirical parameters. This value of "parameters from theory" is a prevalent value in theoretical physics though I will not detail further here.

## 2.8 Cycling and Helmets

Protect your brain, save your life.

From the NHTSA pamphlet *Kids and Bicycle Safety*

There is currently a scientific controversy about the role of helmets in bike safety.[13]

> With regard to the use of bicycle helmets, science broadly tries to answer two main questions. At a societal level, "What is the effect of a public health policy that requires or promotes helmets?" and at an individual level, "What is the effect of wearing a helmet?" Both questions are methodologically challenging and contentious. [Goldacre and Spiegelhalter, 2011]

To begin with, it may turn out that the protective value of bike helmets is not much. There are a number of statistical analysis of the effectiveness of helmets with different conclusions. Goldacre and Spiegelhalter [2011] claim that the superior analysis points to the result that helmets are at best slightly effective. They mention a paper that "concludes that the effect of Canadian helmet legislation on hospital admission for cycling head injuries seems to have been minimal."

The statistical analysis is one way to judge the effectiveness of helmets. Case-control studies, on the other hand, have demonstrated shown that wearing helmets are more likely to protect from a head injury:

> Such findings suggest that, for individuals, helmets confer a benefit. These studies, however, are vulnerable to many methodological shortcomings. If the controls are cyclists presenting with other injuries in the emergency department, then analyses are conditional on having an accident and therefore assume that wearing a helmet does not change the overall accident risk. There are also confounding variables that are generally unmeasured and *perhaps even unmeasurable*. People who choose to wear bicycle helmets will probably be different from those who ride without a helmet: they may be more cautious, for example, and so less likely to have a serious head injury, regardless of their helmets. [Goldacre and Spiegelhalter, 2011, my emphasis]

It is interesting that the authors suggest that the issue can be too complicated for measure. But it becomes even more complicated when further factors are considered:

---

[13] I will be closely following [Goldacre and Spiegelhalter, 2011] in this section.

> People who are forced by legislation to wear a bicycle helmet, meanwhile, may be different again. Firstly, they may not wear the helmet correctly, seeking only to comply with the law and avoid a fine. Secondly, their behaviour may change as a consequence of wearing a helmet through 'risk compensation', a phenomenon that has been documented in many fields. One study — albeit with a single author and subject — suggests that drivers give larger clearance to cyclists without a helmet. [Goldacre and Spiegelhalter, 2011]

Possibly, cyclists with helmets are more daring and more likely to be involved in accidents. Also, drivers pass cyclists more closely if they wear helmets. (See also figure 2.13.)

Since cycling is a health benefiting activity, there is a population health aspect of the discussion as well. One writer[14] lists these consequences of compulsory helmet wearing as follows:

- Discourages cycling because people will prefer to not ride rather than wear a helmet.
- Discourages cycling because it promotes an idea of cycling as inherently dangerous.
- Destroys the possibility of municipal bike sharing/rental programs.
- Helmeted cyclists more likely to be struck by motorists.
- Fewer cyclists = more drivers = more global warming, toxic pollution.
- Fewer cyclists = more drivers = more energy use.
- Fewer cyclists = more dangerous for the cyclists who remain (motorists less used to expecting cyclists on the road).

This list suggests that making helmets compulsory to cyclists can backfire, making cycling a less frequent and dangerous activity.

It seems common-sense that wearing helmets is good for cyclists. But science can question the common-sense view. There is anecdotal evidence[15] that helmets "save lives". Interestingly, the anecdotal evidence for usefulness of helmets seems more to come from cultures without a strong cycling background. For example, the common view in Netherlands (see figure 2.14) is that helmets are ineffective or unnecessary.[16] The role of anecdotes in science is something that I will look at more closely in section 4.5. For now, I want to emphasize that what anecdotes tell can be different from the scientific version, and whether helmets are needed or not cannot be settled by common-sense.

The surprising thing about the helmet debate is the simplicity of the problem. One might think that surely scientists can decide on the helmet question by looking at the evidence. But it seems that evidence is not enough to decide. The issue is too complex to be decided on evidence — and I have not even mentioned freedom of choice aspects of it. More surprisingly, Goldacre and Spiegelhalter claim that no amount of evidence might be enough:

---

[14] From http://bicycleaustin.info/laws/helmet-laws-bad.html. Retrieved on February 25, 2015.

[15] E.g. see the news story [Bristol Post, 2014] which claims that "Wearing a cycle helmet was a choice that saved one 12-year-old boy's life."

[16] See for example http://helmetfreedom.org/1052/dutch-courage/ or http://study-abroad-blog-amsterdam-ss.ciee.org/2011/03/bike-helmets-the-dutch-have-a-different-philosophy-ciee-amsterdam.html. Accessed on May 1, 2015.

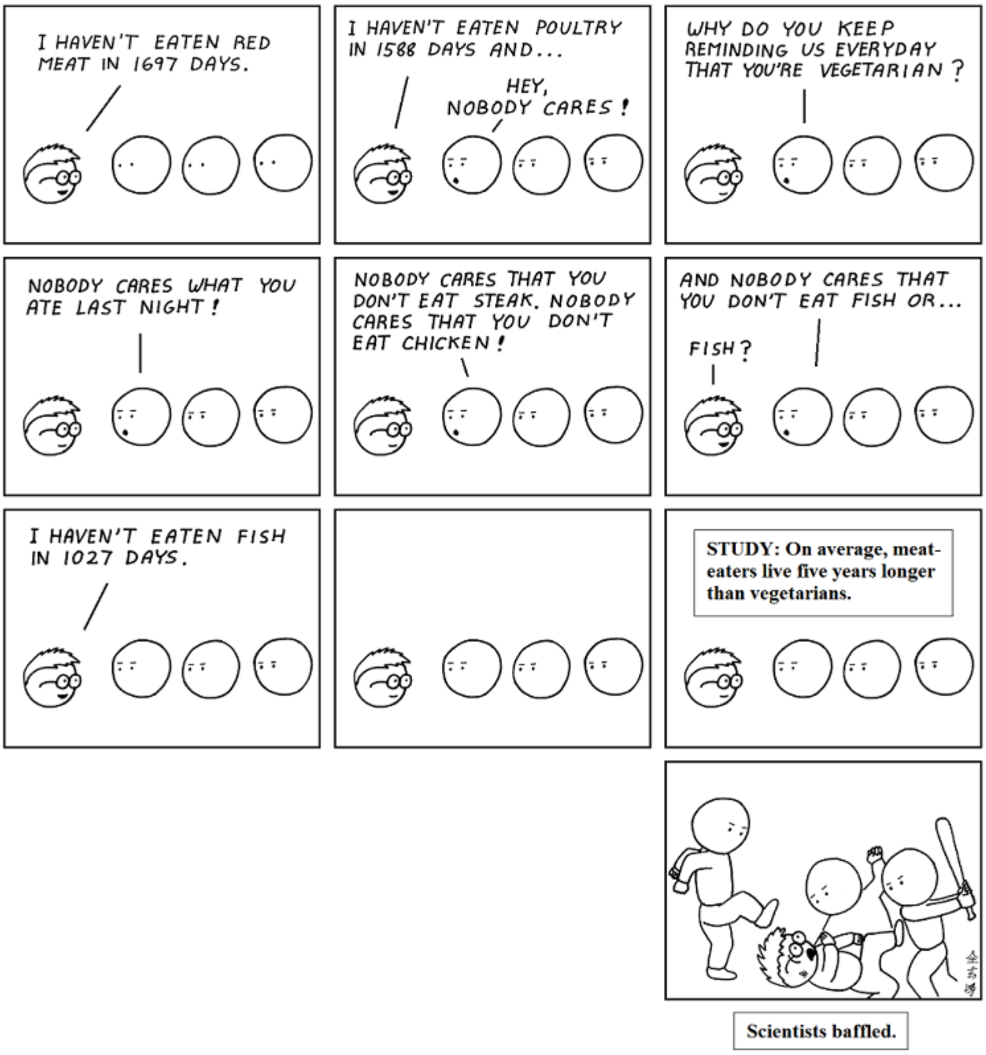



Figure 2.13: A comic from www.abstrusegoose.com/488 which shows confounding variables at work. [Image licensed under Creative Commons Attribution-Noncommercial 3.0 United States License.]

Figure 2.14: A cyclist in Amsterdam. Netherlands has a very strong cycling culture and no mandatory helmet requirement. [Photograph by Alfredo Borba licensed under Creative Commons Attribution-Share Alike 4.0 International.]

> Standing over all this methodological complexity is a layer of politics, culture and psychology. Supporters of helmets often tell vivid stories about someone they knew, or heard of, who was apparently saved from severe head injury by a helmet. Risks and benefits may be exaggerated or discounted depending on the emotional response to the idea of a helmet. For others, this is an explicitly political matter, where an emphasis on helmets reflects a seductively individualistic approach to risk management (or even 'victim blaming'), while the real gains lie elsewhere. It is certainly true that in many countries, such as Denmark and the Netherlands, cyclists have low injury rates, even though rates of cycling are high and almost no cyclists wear helmets. This seems to be achieved through interventions such as good infrastructure, stronger legislation to protect cyclists, and a culture of cycling as a popular, routine, non-sporty, non-risky behaviour.
>
> In any case, *the current uncertainty about any benefit from helmet wearing or promotion is unlikely to be substantially reduced by further research.* Equally, we can be certain that helmets will continue to be debated, and at length. *The enduring popularity of helmets as a proposed major intervention for increased road safety may therefore lie not with their direct benefits – which seem too modest to capture compared with other strategies — but more with the cultural, psychological and political aspects of popular debate around risk.* [Goldacre and Spiegelhalter, 2011, emphases mine]

This is a case of what philosophers call *underdetermination of theory by evidence*: no amount of evidence can decide between the alternative theories. (I will examine underdetermination in section 5.4.) More importantly, helmet issue is embedded in many different cultural, social, and political contexts that it is impossible to untangle without taking these under consideration. Whether or not this context dependency can be considered to be a bias will be the topic of section 3.5.

## 2.9 Monkey Business

> One day when [the bonobo] Kanzi was visiting [the chimpanzee] Austin, he wanted some cereal that had been prepared specifically for Austin. He was told, "You can have some cereal if you give Austin your monster mask to play with." Kanzi immediately found his monster mask and handed it to Austin, then pointed to Austin's cereal. When told "Let's go to the trailer and make a water balloon," Kanzi went to the trailer, got a balloon out of the backpack, and held it under the water faucet. He needed help fitting it on the faucet and filling it with water, but he had clearly understood the sentence. Even sentences with general terms such as "it" or "this" were easy for Kanzi. For example, sentences like "If you don't want the juice put it back in the backpack" were readily responded to, as were sentences like "Get some water, put it in your mouth."
>
> Savage-Rumbaugh et al. [1998, 67]

The concept of *theory of mind* (ToM) was first introduced by Premack and Woodruff [1978]. Other names for ToM are *mindreading* and *mentalizing*. ToM is the cognitive capacity of attributing mental states to self and others. These mental states include perceptions, feelings, sensations, emotions, beliefs, desires, hopes, goals, doubts, intentions, knowledge, and so on.

The word *theory* is used in ToM because early writings on ToM assumed the theory-theory view [see Weiskopf, 2011]. The name ToM stuck and nowadays it is continued to be used by all, even by the opponents of the theory-theory view.

ToM is useful in predicting, explaining, and justifying behaviour of others. If you know what is in some other agent's mind then you can coordinate your behaviours accordingly.

There is a diverse field of research in ToM.[17] One of the main research areas is to determine the extent and type of ToM in children (including autistic children) and animals. The most studied animals in this regard are primates, followed by dolphins, birds, and dogs. I will concentrate on the problem of finding ToM in the great apes (see figure 2.15).

Michael Tomasello, Joseph Call, and Brian Hare's research group in Max Planck Institute for Evolutionary Anthropology, Department of Developmental and Comparative Psychology in Leipzig conducted some interesting experiments on chimpanzees in 2000s. They write in [Tomasello et al., 2003b] that in 1990s they were sceptical of attributing ToM to chimpanzees, but their views changed after the mentioned experiments. Let us look at the experiment given in [Call, 2001]. Daniel J. Povinelli and Jennifer Vonk summarize this food competition experiment as follows:

> A subordinate and dominant were positioned on either side of an empty room from each other, temporarily prevented from entering by doors which could

[17] Some ToM related questions include: "How do people execute this cognitive capacity? How do they, or their cognitive systems, go about the task of forming beliefs or judgements about others' mental states, states that aren't directly observable? Less frequently discussed in psychology is the question of how people self-ascribe mental states. Is the same method used for both first-person and third-person ascription, or entirely different methods? Other questions in the terrain include: How is the capacity for ToM acquired? What is the evolutionary story be hind this capacity? What cognitive or neurocognitive architecture underpins ToM? Does it rely on the same mechanisms for thinking about objects in general, or does it employ dedicated, domain-specific mechanisms? How does it relate to other processes of social cognition, such as imitation or empathy?" [Goldman, 2012]

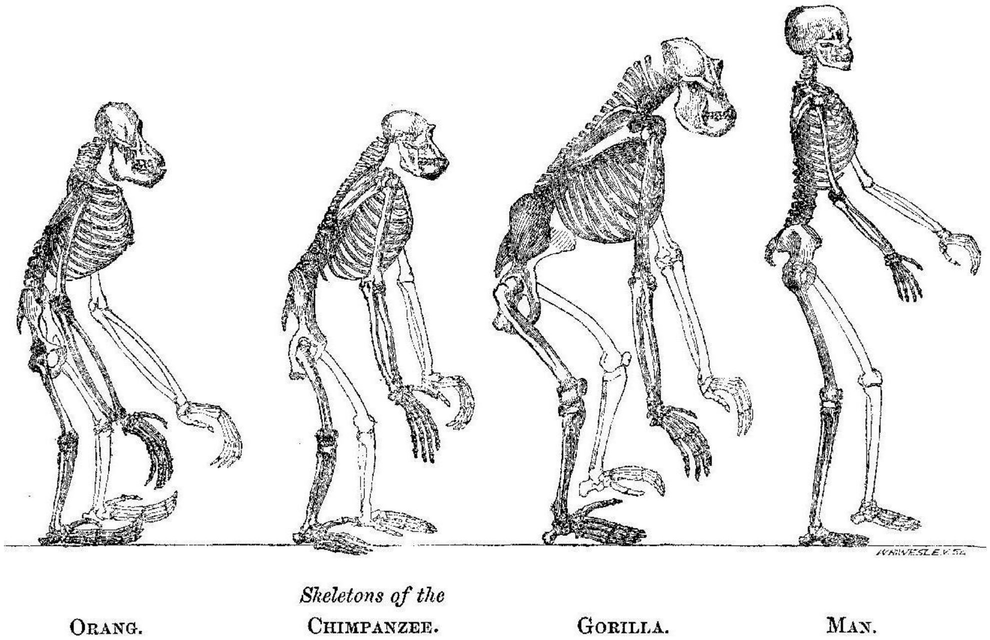



Figure 2.15: Part of the frontispiece of Thomas Henry Huxley's *Evidence as to man's place in Nature*, 1863. The text under the illustration reads "Photographically reduced from Diagrams of the natural size ... drawn by Mr. Waterhouse Hawkins from specimens in the Museum of the Royal College of Surgeons." Most language and ToM research has involved the great apes whose skeletons are depicted in this illustration though there is some research carried on other animals as well.

> be opened: either slightly, to let them look into the room, or all the way, to let them enter. An experimenter placed food into one of two cups between them. The subordinate's door was opened first, giving him or her a head start. When the subordinate, but not the dominant, was allowed to observe the baiting, the subordinate frequently approached the food; when both the subordinate and dominant observed the food being hidden, the subordinate was less likely to approach the food. [Povinelli and Vonk, 2003]

Some clever variations of this experiment were also carried out by the Leipzig group but I will refrain from describing them. The important point is that, according to the Leipzig group, these results show that the subordinate seems to know what the dominant sees and stays away from the food that the dominant saw being planted. They argue in [Tomasello et al., 2003b] that ToM is not a black or white property; it has colours and shades. Even though the chimpanzees do *not* have a "full-blown, human-like" ToM, they have some what we might call "visual ToM". According to the Leipzig group,

> (1) Chimpanzees know what other individuals do and do not see. Thus, they reliably follow the gaze direction of others; they do this around barriers and past distracters (which suggests that they are doing much more than just using head direction as a discriminative cue); and they reliably use information about what conspecifics can and cannot see in a food competition situation ...
>
> (2) Chimpanzees can recall what a conspecific has and has not seen in the imme-

diate past, and this recall is associated with specific individuals. Chimpanzees integrate this recall about what specific individuals have and have not seen into their behavioural decision making in competitive situations.

> For some theorists, this profile is just another way of saying that chimpanzees know what others know. [Hare et al., 2001, 149, references removed]

This conclusion is challenged by Povinelli and Vonk [2003, 2004]. They say that the experiment does not establish that the subordinate understands the connection between seeing and knowing because it is possible that the chimpanzee mind has a behavioural abstraction rule that says

- Don't go after the food if that dominant has oriented towards it.

Perhaps the subordinate is automatically following such a rule without knowing what the dominant sees. In view of this possible lower level explanation, the higher level explanation of the Leipzig group cannot be accepted.

Povinelli and Vonk consider another experiment to test visual ToM in chimpanzees. Ironically, Andrews [2005, 531] gives a lower level explanation for this experiment. Andrews [2005, 532] suggests an even more devious experiment. But a little bit of thought also shows that that experiment is open to low level explanations. I will not examine these experiments here because their details are unimportant for my purposes. Rather, I want to make the general claim that no individual experiment can test ToM in animals.

Consider an experiment[18] in which the only observable factor is a behaviour $A$ of an animal. No such experiment can establish ToM because one can always find lower than ToM level explanations of the behaviour $A$. For example, one can claim that there is a behavioural rule to the effect that under suitable conditions, the animal performs $A$. No doubt, depending on the particular $A$, more ingenious lower level explanations than the one I proposed can be found.

One of the fundamental principles employed in discussions of animal psychology is *Morgan's Canon* introduced by the 19th century comparative psychologist C. Lloyd Morgan. Here is its original formulation and a modern paraphrase of Morgan's Canon:

> In no case is an animal activity to be interpreted in terms of higher psychological processes if it can be fairly interpreted in terms of processes which stand lower in the scale of psychological evolution and development. [Morgan, 1894, 53]
>
> In no case may we interpret an action as the outcome of the exercise of a *more sophisticated* psychological faculty, if it can be interpreted as the outcome of one which is *less sophisticated*. [Fitzpatrick, 2008, 227]

What exactly Morgan meant by this principle is still disputed. But as it is employed today, it is a principle that tells us to favour hypotheses that refer to lower cognitive capacities to those that refer to higher capacities.

If all we have to go on is the behaviour $A$ of an animal, then we do not know what goes on in its mind. Morgan's canon implies that we can consistently claim that for that behaviour to occur, no ToM is needed. Daniel C. Dennett acknowledges that "in principle a lowest-order story can always be told of any animal behavior" but he downplays this idea as follows:

[18] Note that here I am talking about experiment types, not tokens. That is, you are allowed to repeat the same experiment over and over again. As long as you elicit behaviours $A_1, A_2, \ldots, A_n$ which are somehow similar, you can find a common behavioural generalization of these.

> Lloyd Morgan's canon of parsimony enjoins us to settle on the most killjoy, least romantic hypothesis that will account systematically for the observed and observable behavior, and for a long time the behaviorist creed that the curves could be made to fit the data well at the lowest level prevented the exploration of the case that can be made for higher-order, higher-level systematizations of the behavior of such animals. The claim that in principle a lowest-order story can always be told of any animal behavior (an entirely physiological story, or even an abstemiously behavioristic story of unimaginable complexity) is no longer interesting. It is like claiming that in principle the concept of food can be ignored by biologists — or the concept of cell or gene for that matter — or like claiming that in principle a purely electronic-level story can be told of any computer behavior. [Dennett, 1987, 246–247]

Dennett is right in saying that most of the time explanations of different levels do not compete with each other. I can explain why a rectangular wood block does not fit through a smaller hole either geometrically or at a physical level by referring to the electromagnetic forces, etc. Clearly, both explanations complement each other and we can accept both. But Dennett is mistaken to think that explanation of a behaviour is similarly possible at different levels. For example, in the food competition experiment discussed above, any lower level explanation of the subordinate's behaviour is tantamount to saying that the chimpanzee has no visual ToM. These are rival explanations that cannot be simultaneously accepted. Yet the experiment accommodates both.

You may think that there is a way out of the conundrum: perform many different types of experiments as you can. Surely then you can decide which lower or higher level hypothesis is the right one. There are three problems with this reasoning.

First, it is really hard to come up with novel ways of testing ToM in animals. It is not really helpful to come up with more elaborate or controlled experiments. The more elaborate the experiment becomes, that is, the more variables you have, the more wiggle room for interpreting the results. According to Povinelli and Vonk,

> It is tempting to think that we can remedy these failings of the current line of experiments by simply implementing more or better controls. However, the problem is not the ingenuity of the experimenters; it is the nature of the experiments. Techniques that pivot upon behavioral invariants (looking, gazing, threatening, peering out the corner of the eye, accidentally spilling juice versus intentionally pouring it out), will always presuppose that the chimpanzee (or other agent) has [low level] access to the invariant [category of behaviour], thus crippling any attempt to establish whether a mentalistic coding is also used. The sobering point is that no experiment in which the theory-of-mind coding derives from a behavioral abstraction will suffice. Control will chase control with no end in sight, leaving only our intuitions, hopelessly contaminated by our folk psychology, to settle the matter. [Povinelli and Vonk, 2003, 159]

Moreover, it is harder to expect the animals to cooperate in such nuanced experiments.

Second, not everyone agrees with the validity of the results. For example, the Leipzig group has criticized some experiments a number of times for their lack of "ecological validity" [Tomasello et al., 2003a, 239]. In some experiments that chimpanzees interacted with human experimenters, they failed to show any ToM. But there might be good reasons to think that the results failed just because the experimenter was a human. In natural conditions, chimpanzees do not use their (alleged) ToM to attribute mental states to humans which must appear all-knowing to them. What one should test is not chimpanzee's ToM of humans in laboratory conditions, but rather chimpanzee's ToM of chimpanzees in natural conditions. At least that's how their argument goes. To repeat the point, the results of an experiment can be rejected for various reasons, one being ecological validity.

Finally, and this is the most important one, a finite number of experiment types are not enough to decide if a high level hypothesis is true. The reason is that the alternative of this hypothesis is not necessarily a single low level hypothesis. You can pick a different low level explanation for different experiments. Morgan's Canon does not tell us to pick a low level explanation that works for *all* experiments, rather, it tells us to pick a low level explanation for *each* experiment.

It seems that we are left with the puzzling result that one can never empirically confirm that an animal has (visual or other) ToM. In order to solve this puzzle, we need to take a closer look at values at work.

Say we have two competing theories. One is a simple theory that depicts nature to be a complex thing. The other is a complex theory that depicts nature to be simple. Moreover, all empirical evidence supports them equally. If this is all you know about these theories which one would you choose? Think about it before you go on reading. Which one would you prefer?

The answer is: we cannot choose between them given just this information. It is a trick question. Complexity and simplicity are not well-defined concepts. Complex/simple in which regard? How much simple/complex? What is the domain/context of these theories? There are numerous parameters that effects the choice.

The choice between low and high level theories of behaviour is analogous to the trick choice. Low level theories attribute simple properties to nature, but in a sense they are complex theories because for each different behaviour you need to come up with a different low level mechanism. On the other hand, by definition, a high level theory finds sophistication in animal behaviours, but it unifies different behaviours under one capacity, for example ToM. Both have their advantages and disadvantages. Scientists may consider "changing levels of explanation and description in order to gain access to greater predictive power or generality — purchased, typically, at the cost of submerging detail and courting trivialization on the one hand and easy falsification on the other." [Dennett, 1987, 239]

The more different and incompatible low level assumptions one needs to accept in order to make sense of an animal's behaviour, the more attractive a high level hypothesis that subsumes all of the low level assumptions becomes. Especially so, if the low level assumptions are ad hoc ones made to avoid a high level hypothesis. So it is possible to accept a high level hypothesis if enough empirical results make the collection of relevant low level explanations too bloated.

It is interesting philosophically to see how empirical data adjudicates between rival hypotheses. As I have argued above, one can never empirically confirm that an animal has (visual or other) ToM. But experiments do matter in a roundabout way. Enough experimental results can make one hypothesis look simpler or more parsimonious and so more favourable than the others.

So do chimpanzees have visual ToM? We do not know yet. We should continue to gather more empirical data. I think that any such data can be made to fit theories that accept or deny ToM. But the data can render one theory more attractive in view of some values like

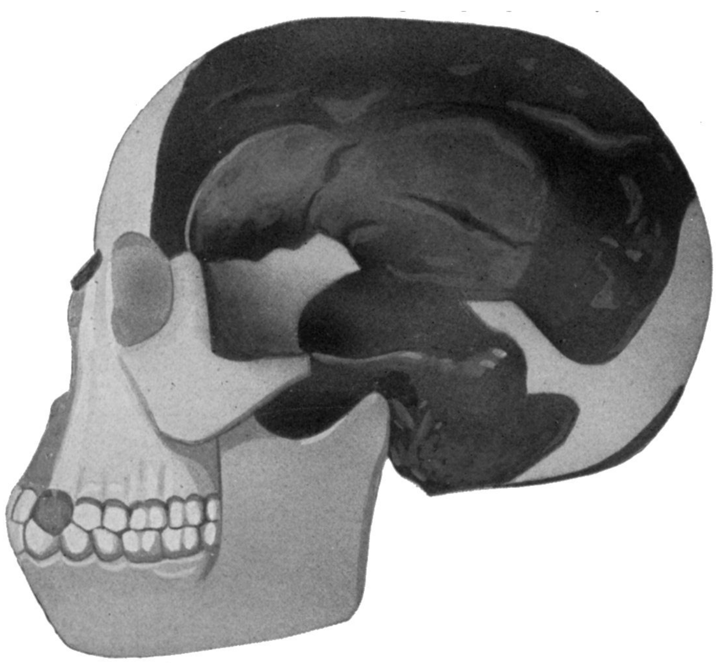


Figure 2.16: A reconstruction of the Piltdown skull from J. Arthur Thomson, *The Outline of Science*, 1922. [Image in the public domain from http://commons.wikimedia.org.]

simplicity, fruitfulness, etc. So the decision can be made eventually by some values different than evidential support.

Quite possibly, the ToM discussions will continue until neuroscience takes over in the future. Once the neural correlates of relevant behaviours are found, there will be no more doubts about the mechanisms of those behaviours. Until then, the controversy might continue.

## 2.10 Scientific Fraud

> The scientific fraud committed by Diederik Stapel, which came to light in Tilburg in early September 2011, sent shock waves across the academic world in the Netherlands and internationally. Ultimately, trust forms the basis of all scientific collaboration. If, as in the case of Mr Stapel, there is a serious breach of that trust, the very foundations of science are undermined.
>
> From *Stapel Investigation Report* [Committee, 2012, 5]

No one would think that fraud and misconduct should be a part of science, or any enterprise for that matter. Nevertheless, there is misconduct in science and looking at it can teach us something about science. In this section I will look at the Piltdown Man case which is arguably the most well-known scientific fraud in history and a more recent Stapel case.

**Piltdown Man.** The skull fragments collected in 1912 by the amateur archaeologist Charles Dawson and the eminent palaeontologist Arthur Smith Woodward of the British Natural History Museum from a gravel pit at Piltdown, East Sussex, England had the remarkable characteristics of both man and ape. Whereas the jaw was ape-like, the upper skull was human. (See figure 2.16.) This sensational find was indeed too good to be true. Scientific community debated the significance of the skull until it was established in 1953 to be a forgery by comparing fluorine content of the bones. As Stephen Jay Gould puts it: "The old anomaly — an apish jaw with a human cranium — was resolved in the most parsimonious way of all. The skull *did* belong to a modern human; the jaw was an orangutan's." [Gould, 1980, 111] The bones were stained and the teeth were filed to make it more believable. The details of *how*, *who*, and *why* of the fraud are still debated to this day. The philosophically remarkable side of this story is the scientific controversy it caused. How did such a fraud, and in hindsight an obvious one, escape the scientific community for 40 years?

There were those who believed that the skull fragments of two different animals were accidentally brought together and mixed in the deposit. Gould [1980, 115] gives the example of the German anatomist and physical anthropologist Franz Weidenreich who wrote in 1940s that the Piltdown man "should be erased from the list of human fossils. It is the artificial combination of fragments of a modern human braincase with orangutanlike mandible and teeth." On the other hand, many British scientists argued for the authenticity of the Piltdown man and they believed that the Piltdown man belonged to the earliest species of humans in Europe. Gould [1980, 115] quotes the Scottish anatomist and anthropologist Sir Arthur Keith who retorted: "This is one way of getting rid of facts which do not fit into a preconceived theory; the usual way pursued by man of science is, not to get rid of facts, but frame theory to fit them." There are a couple of reasons why this controversy lasted for 40 years.

(1) Whereas Europe was blessed with hominid fossils, their tools and art, England was, much to its chagrin, deficient. Considering the political environment of the First World War and the interbellum, the Piltdown man was what the doctor ordered for the British. Piltdown bolstered the national pride of the British by establishing that the birthplace of human race was England. A report in the *Nature* magazine in December 19, 1912 quotes Arthur Smith Woodward downplaying the importance of continental hominids: "the Neanderthal race was a degenerate offshoot of early man and probably became extinct, while surviving modern man may have arisen directly from the primitive source of which the Piltdown skull provides the first discovered evidence." Woodward [1948] even went on to call his book on the Piltdown Man *The Earliest Englishman*. It is no surprise that

> the three leading lights of British anthropology and palaeontology — Arthur Smith Woodward, Grafton Elliot Smith, and Arthur Keith — had staked their carriers on the reality of Piltdown. (Indeed they ended up as two Sir Arthurs and one Sir Grafton, largely for their part in putting England on the anthropological map.) [Gould, 1980, 114]

Other British scientists followed suit and defended the Piltdown man. (See figure 3.3 on page 89.) Here we have both *nationalism* and *influence of authority* at work. Another contributing factor might have been the "all familiar racial views among white Europeans". [Gould, 1980, 117]

(2) Gould also blames another cultural bias guiding the British:

> At that time, many leading paleontologists maintained an a priori preference largely cultural origin, for "brain primacy" in human evolution. The argument rested on a false inference from contemporary importance to historical priority: we rule today by virtue of our intelligence. Therefore, in our evolution, an enlarged brain must have preceded and inspired all other alterations of our body. We should expect to find human ancestors with enlarged, perhaps nearly modern, brains and a distinctly simian body. [Gould, 1980, 116]

We now know that human evolution happened in the opposite way, that is, early hominids that stood erect had small brains. If we look at the Piltdown skull with our cultural background then we would not be tricked by the fraud. But with the "brain primacy" culture, Piltdown skull is something to be expected. Gould [1980, 117] quotes Grafton Elliot Smith writing in 1924:

> The outstanding interest of the Piltdown skull is in the confirmation it affords of the view that in the evolution of Man the brain led the way. It is the veriest truism that Man has emerged from the simian state in virtue of the enrichment of the structure of his mind. ...The brain attained what may be termed the human rank at a time when the jaws and face, and no doubt the body also, still retained much of the uncouthness of Man's simian ancestors. In other words, Man at first ...was merely an Ape with an overgrown brain. The importance of the Piltdown skull lies in the fact that it affords tangible confirmation of these inferences.

There is a myth that, as Gould [1980, 115] puts it, "facts are 'hard' and primary and that scientific understanding increases by patient collection and sifting of these objective bits of pure information." In truth, evidence is seen and interpreted in the light of theory, that is, using the philosophical jargon, evidence is *theory-laden*. The fraud of Piltdown is obvious for us in hindsight because the brain primacy theory is long gone. But with their cultural and theoretical background, the Piltdown man was welcomed. I will return to the relation of theory and evidence in section 3.4.

**Case of Diederik Stapel.** Do littered public places bring out racist tendencies in people? Do individuals consume more candy from a bowl in front of them if the bowl has the word "kapitalisme" printed on it? Diederik Stapel was a star of social psychology who had investigated such topics when it became known in September 2011 that he committed scientific fraud for years by fabricating/manipulating data. When Stapel's fraud came to light, three universities where he had worked — Amsterdam, Groningen and Tilburg — formed committees to investigate his work. (See figure 2.17.) The final report of the joint committee [Committee, 2012] is also available in English. The human side of the story (which I will omit here) is well-told by Yudhijit Bhattacharjee [2013] in an article in The New York Times. What interests me here is rather what the joint committee report tells about the social psychology in general.

Like the Piltdown Man, the Stapel case is puzzling for its endurance: How can the continued fraud of Stapel can go undetected for over fifteen years? He had co-authors, doctorate students, editors, referees, and other social psychologists reading his papers (55 of which are retracted today, see figure 2.18), meanwhile he presented talks, gave seminars, discussed his works with colleagues, and so on. Stapel was well-known and well-regarded in social psychology. How can his fraud go undetected for such a long time? This question is also asked by the joint committee:

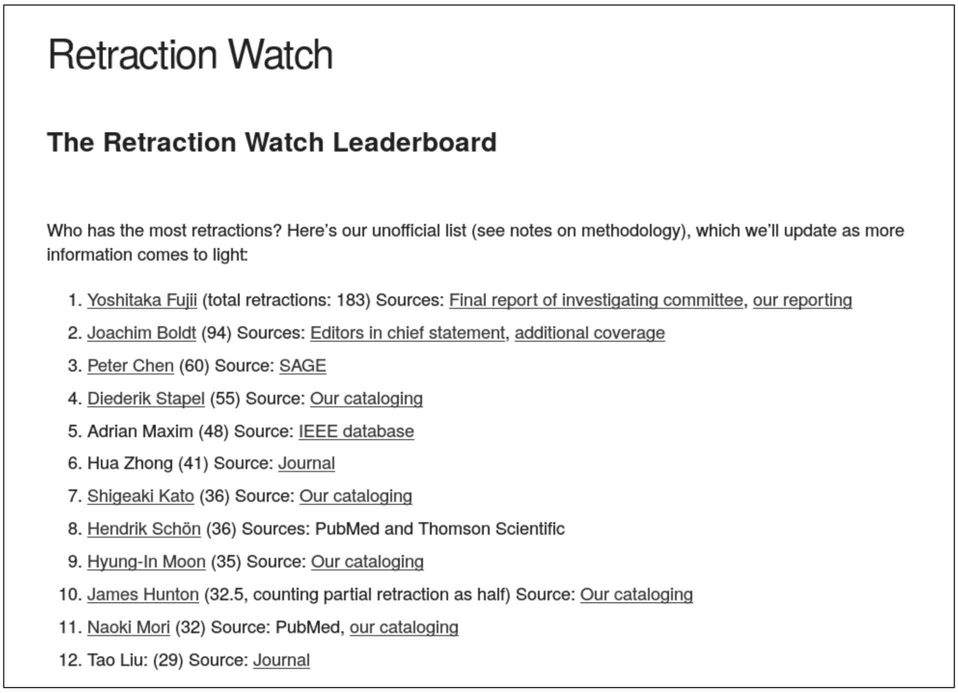

Figure 2.17: Dutch stamps issued in 1964 celebrating the 350th anniversary of The University of Groningen where Stapel was a full professor from 2000 until 2006. The Amsterdam committee investigated publications that appeared in the period from 1993 to 1999; the Groningen committee investigated the publications dating from 2000 to 2006; and the Tilburg committee investigated the publications from 2007 to 2011. They report more than 50 fraudulent publications (co-)authored by Stapel that goes a long way back: "The Committees' findings show that fabrication of data in one form or another started before the Tilburg period. The first publication in Groningen in which fraud has been proven is from 2004, and the first publication where evidence of fraud was found dates back to 2001. During the Amsterdam period the first publication where evidence of fraud was found was from 1996." [Committee, 2012, 31]

> the urgent question that remains is why this fraud and the widespread violations of sound scientific methodology were never discovered in the normal monitoring procedures in science.
>
> The data and findings were in many respects too good to be true. The research hypotheses were almost always confirmed. The effects were improbably large. Missing, impossible, or out-of-range data are rare or absent. Highly conspicuous impossible findings went unnoticed. ...
>
> Virtually nothing of all the impossibilities, peculiarities and sloppiness mentioned in this report was observed by all these local, national and international members of the field, and no suspicion of fraud whatsoever arose. [Committee, 2012, 53]

The committee has an explanation for this surprising picture: the social psychology community did not discover Stapel's fraud because sloppy and suspect work is widespread in this field and that is why Stapel's fraud did not stick out.

> A 'byproduct' of the Committees' inquiries is the conclusion that, far more than was originally assumed, there are certain aspects of the discipline itself that should be deemed undesirable or even incorrect from the perspective of academic standards and scientific integrity.
>
> In the case of the fraud committed by Mr Stapel, the critical function of science has failed on all levels. Fundamental principles of scientific method have been ignored, or set aside as irrelevant. In the opinion of the Committees this has contributed significantly to the delayed discovery of the fraud. [Committee, 2012, 54]

By "the critical function of science" the committee refers to the practice of the peers of Stapel. There are a number of problems with the refereeing process, for example: "Not infrequently



Retraction Watch

**The Retraction Watch Leaderboard**

Who has the most retractions? Here's our unofficial list (see notes on methodology), which we'll update as more information comes to light:

1. Yoshitaka Fujii (total retractions: 183) Sources: Final report of investigating committee, our reporting
2. Joachim Boldt (94) Sources: Editors in chief statement, additional coverage
3. Peter Chen (60) Source: SAGE
4. Diederik Stapel (55) Source: Our cataloging
5. Adrian Maxim (48) Source: IEEE database
6. Hua Zhong (41) Source: Journal
7. Shigeaki Kato (36) Source: Our cataloging
8. Hendrik Schön (36) Sources: PubMed and Thomson Scientific
9. Hyung-In Moon (35) Source: Our cataloging
10. James Hunton (32.5, counting partial retraction as half) Source: Our cataloging
11. Naoki Mori (32) Source: PubMed, our cataloging
12. Tao Liu: (29) Source: Journal

Figure 2.18: The website http://retractionwatch.com/ is dedicated to documenting retractions in scientific journals. Retractions are important for scientific work (and they are interesting) but unfortunately not enough publicized — this is why *Retraction Watch* matters. The above image is a partial screen capture (taken on October 15, 2015) of *Retraction Watch's* top part of *The Retraction Watch Leaderboard* which is a list of people with most retractions. Diederik Stapel and physicist Jan Hendrik Schön whom I write about in section 4.3 are quite high in the list. (See also *Retraction Watch's* sister-site https://embargowatch.wordpress.com/ for another interesting and oft-neglected aspect of science publishing.)

reviews were strongly in favour of telling an interesting, elegant, concise and compelling story, possibly at the expense of the necessary scientific diligence. It is clear that the priorities were wrongly placed.” [Committee, 2012, 53] Another point made by the committee is:

> Time and again journals and experienced researchers in the domain of social psychology accepted that Mr Stapel’s hypotheses had been confirmed in a single experiment, with extremely large effect sizes. ... However, there was usually no attempt to replicate, and certainly not independently. The few occasions when this did happen systematically, and failed, were never revealed, because this outcome was not publishable. [Committee, 2012, 54]

The “outcome was not publishable” because of *publication bias*: in some fields it is much harder to publish negative results. Publication bias can have serious consequences, see section 2.13.

The committee has also found *verification bias* in Stapel’s work.

> Verification bias is not the same as the ‘usual’ publication bias, which is the phenomenon in which negative or weak findings that do not clearly confirm the theoretical expectations, if at all, but were obtained in soundly executed research, do not find their way into the journals, unlike ‘positive’ results. Verification bias refers to something more serious: the use of research procedures in such a way as to ‘repress’ negative results by some means. [Committee, 2012, 48]

The committee’s list of verification bias in Stapel’s work is important and educative, so I have reproduced it as appendix A on page 193. If this bias is not contained to Stapel’s own work, and indicative of wider problems in social psychology as the report suggests, then the whole field becomes suspect.

> [Stapel’s case] involved a more general failure of scientific criticism in the peer community and a research culture that was excessively oriented to uncritical confirmation of one’s own ideas and to finding appealing but theoretically superficial ad hoc results. ...
>
> The Committees were forced increasingly to the conclusion that, even in the absence of fraud in the strict sense, there was a general culture of careless, selective and uncritical handling of research and data. The observed flaws were not minor ‘normal’ imperfections in statistical processing, or experimental design and execution, but violations of fundamental rules of proper scientific research with a possibly severe impact on the research conclusions. The Committees are of the opinion that this culture partly explains why the fraud was not detected earlier. [Committee, 2012, 47]

The committee highlights the role of theory in their report. The researches in this field come up with theory first and devise experiments that reinforces the theory. This extreme version of theory-ladenness was also mentioned by Stapel himself in the year 2000:

> The freedom we have in the design of our experiments is so enormous that when an experiment does not give us what we are looking for, we blame the experiment, not our theory. (At least, that is the way I work). Is this problematic? No. [Quoted in Committee, 2012, 47]

I will end my discussion of the Stapel case here, though there is more to be learned from the committee report.

There is much common between the two cases of fraud discussed in this section. Both present a paradox at first sight: how can they go undetected for so long? The answer lies in their respective research cultures and theoretical backgrounds. These frauds might be obvious in hindsight, but, as I have discussed above, they were not so in their own milieu.

## 2.11 Quarks

> At this moment his dream ended and Pinocchio opened his eyes and awoke. But imagine his astonishment when upon awakening he discovered that he was no longer a wooden puppet, but that he had become instead a boy, like all other boys.
>
> Carlo Collodi, *The Adventures of Pinocchio*

The history of particle physics presents a very rich and fascinating playground for the philosopher of science. In this section, I will look at a section of it, namely the period around the invention of quarks.

**Mathematical entities.** The concept of mathematical entity mentioned in section 2.4 is quite interesting. A *mathematical entity* is a mathematical object/construct which is thought not to exist in the physical sense, though for all intents and purposes, its role in physics is not different than physical objects. The electron holes are a case in point: they do not exist, yet they are as prominent as electrons in solid state physics. While physicists acknowledge the non-existence of a mathematical entity, they behave as if it is as real as the Eiffel Tower. Mathematical entities feature in all kind of interactions, experiments, explanations, predictions. They are causally efficacious and they can have physical properties like mass, charge, and so on.

At first sight mathematical entities present a puzzle: How can physicists acknowledge their non-existence and behave as if they exist? Simply because it is useful to do so. The language of science is not always taken literally. Scientists sometimes talk in a roundabout way if it serves them. Also just because a theory has an ontological commitment and physicists use that theory does not imply that they have the same commitment. Acceptance and use of a theory does not necessarily imply its literal construal or ontological commitment. [cf. van Fraassen, 1980, 11]

The status of mathematical entities can change in time. It is possible for a mathematical entity to lose its usefulness and be discarded. But the dream of every mathematical entity is one day to become a real entity in the eyes of its makers. The annals of physics is full of long forgotten entities and it takes a special entity to bridge the gap. In this section, I will tell the tale of one group of such entities, namely *quarks*. If Pinocchio (figure 2.19) of this story is the quark, then Geppetto is surely the American physicist Murray Gell-Mann who received the 1969 Nobel Prize in physics for his work on the theory of elementary particles.

**The Eightfold Way and quarks.** The story starts in mid-1950s which was an important time for particle physics.

> We can see that until 1953 new particles and interactions were discovered primarily from the investigations of cosmic radiation. The explanation for this


Figure 2.19: An illustration of Pinocchio from the 1901 large print edition by the Italian artist and illustrator Carlo Chiostri (1863–1939). [Image in the public domain.]

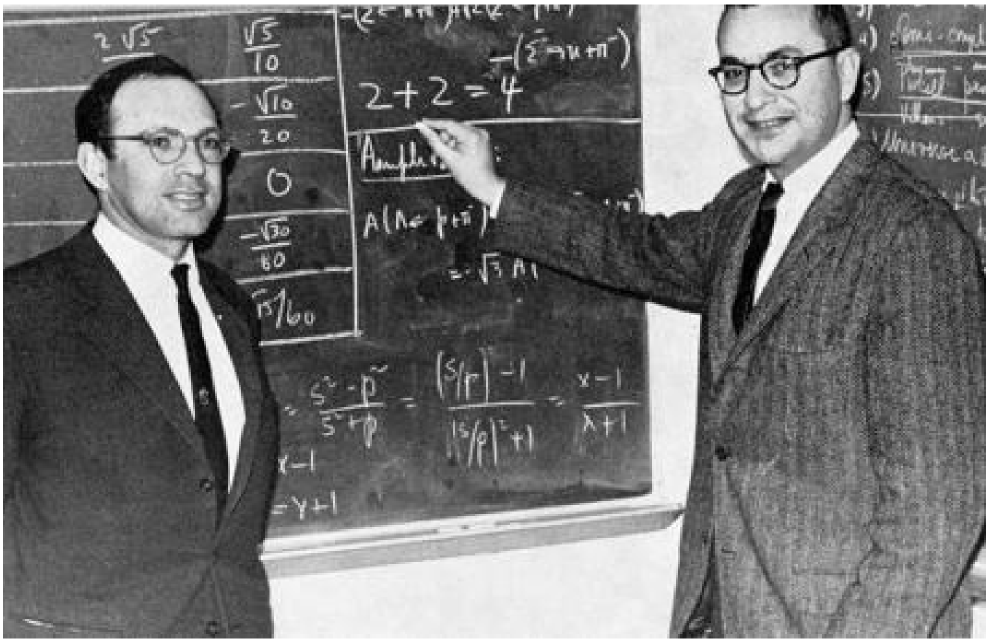



Figure 2.20: Murray Gell-Mann (right) and Yuval Ne'eman (left), co-discoverers of the *Eightfold Way* in 1964. [Source unidentified.]

> is very simple. In the interaction processes of the elementary particles, we are dealing in most cases with very large energies, several GeV (giga electron volts). It was only in the second half of the twentieth century that it became possible to produce such energies in the laboratory. Indeed, ... after the mid 1950s the focus of experimental particle physics shifted to investigations of those phenomena that could be created by using large accelerators. [Simonyi, 2012, 529]

In late 1950s and early 1960s, new particles and resonances were discovered frequently to the chagrin of theorists who lacked an adequate theory that clarified this picture. This was a chaotic time for particle physics. There were dozens of different proposals to classify particles. One successful approach was Gell-Mann's (independently proposed by Yuval Ne'eman, see figure 2.20) *Eightfold Way* theory that organized different subatomic particles into neat geometrical diagrams. One of them is reproduced as figure 2.21.

The Eightfold Way is a continuation of a theme well-known in the history of science: mathematical or geometrical organization of nature.

Probably the most well-known episode in this theme is Mendeleev's periodic table that led to discoveries of new elements. What Mendeleev did for elements were repeated for particles by the Eightfold Way. In a geometrical arrangement of heavier baryons, known as the *baryon decuplet*, one spot was missing. Gell-Mann predicted that a yet undiscovered particle $\Omega^-$ of such properties existed. The discovery of $\Omega^-$ in 1964 was a huge success of the Eightfold Way. But the Eightfold Way did not resolve all questions about the nature of subatomic particles:

> In his Scientific American article in 1957, Gell-Mann had declared that the present understanding of the subatomic particles was like that of the elements after Mendeleev had crafted his periodic table. Mendeleev had charted the similarities, arranged the elements into groups. But an understanding of why the elements lined up just so hadn't come until the invention of atomic theory. Only then was

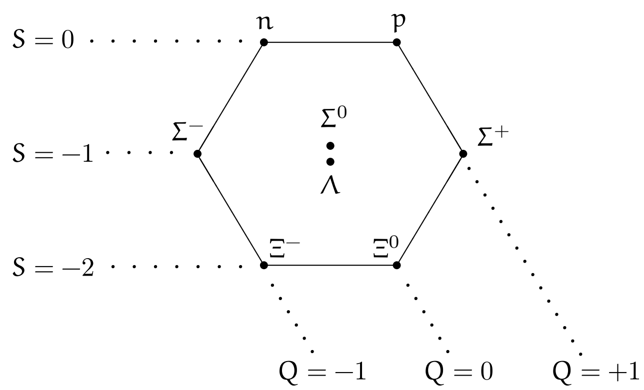



Figure 2.21: The diagram known as the *baryon octet* is the group of lightest baryon particles including the neutron $n$ and the proton $p$ arranged in an hexagon with two particles at the center. The *electric charge* of the particles on a sloping diagonal are like. Those on the same horizontal line are particles of like *strangeness*. For example, the particles $n$, $\Sigma^0$, $\Lambda$, $\Xi^0$ have the same (neutral) charge and the particles $\Xi^-$, $\Xi^0$ have the same strangeness. [Diagram adapted from Griffiths, 2008, 35]

> it established that each successive element had one more proton in its nucleus, hence one more counterbalancing electron in its shell, giving atoms their different chemical characters. Now an ordering principle had to be found for the subatomic particles. It was clear that there were deep patterns, but no one had yet glimpsed the underlying mechanism. [Johnson, 2000, 167]

Just as Mendeleev had no idea why elements were arranged so in his periodic table, physicists did not know what lie behind the Eightfold Way. It was one thing to fit the particles into neat patterns, but another thing to understand why they fit those patterns. There was a hectic search for an underlying theory. As Gell-Mann [1964] starts his quark article, "we are tempted to look for some fundamental explanation of the situation." His answer, also independently proposed by George Zweig (who used the name "ace" rather than "quark"), was the *quark model* put forward in that article in 1964. Quarks were introduced as elementary particles that combined in pairs or trios to form composite particles including mesons and baryons. The properties of these composite particles (including the Eightfold Way) was explained using the quark model.

**Clashing values.** Gell-Mann actually offers two models in his 1964 article the second of which is the quark model. The first unnamed model consists of four particles that have one advantage over quarks: they have integer charge. It was believed at the time that any particle has charge equal to an integer multiple of the charge of an electron. The unnamed model preserves this status quo. But after introducing the first model, Gell-Mann writes that "A simpler and more elegant scheme can be constructed if we allow non-integral values for the charges" and goes on to introduce quarks. This paper demonstrates different values clashing: one model has integer charges, the other has simplicity and elegance. Gell-Mann sides on simplicity and elegance, but as we shall see below, not everyone shared his views.

Values frequently clash in science, but it does not have to be grandeur. Here we see an example of a mini-clash in a single page of Gell-Mann's article. But in this case, the clashing values were not contained there, but occupied the physics community for years to come. The controversy revolved around fractional charge.

**Fractional charges.** Norwood Russell Hanson [1963, 153–159] gives a short history of electricity and particle physics until 1930s in which he tells about two ideas ingrained at the time: (1) There are two elementary particles: electron and proton. (2) There is a basic unit of charge and every charged body has an integer multiple of it.

> Since the proton and the electron came to be thought of not only as carrying the charge, but also virtually as being the charge, the very conception of a third particle beyond the proton and the electron seemed insupportable.
>
> It is this profound conceptual resistance, built into the structures of classical electrodynamics and elementary particle theory, which must be appreciated in order to understand why physicists like Dirac, Blackett, Skobeltzyn, Pauli, Oppenheimer, Anderson, Bohr and Rutherford struggled so hard to avoid such a supposition [about the existence of a third particle]. [Hanson, 1963, 159]

Hanson goes on to tell (the philosophically interesting story of) how the first assumption was overcome in early 1930s and new particles were accepted.[19] But whereas the first assumption was dispensed with, the second one was even more entrenched in the process as each new particle found had integer charge. By 1960s, the integer charge had a perfect score and no doubt it was quite hard to dispense with.

The integer charges can be seen as a case of *conservation of style*. Physicists tend to keep the style or format of the theories that works. They even use this as a tool of discovery/construction: To develop a new theory, they adapt the successful old theories for the new circumstances. The prevalence of analogies in science is a testimony to conservation of style. Physics did well with integer charge; there was no good reason to break this tradition. Besides, no fractionally charged particle was ever found.

The fractional charges were quite frowned upon at the time. In 1962, Yuval Ne'eman and Haim Goldberg also entertained the idea of fractionally charged particles but they were sceptical:

> No one had ever seen charge come in fragments. It was hard enough trying to sell the Eightfold Way. Arguing that it was erected on a scaffolding of fractionally charged subparticles would be a public relations disaster. SU(3) triplets, they decided for the time being, must be little more than a curious mathematical accounting device. [Johnson, 2000, 203]

Zweig, talking about his CERN preprint on aces, recalls that:

> The reaction of the theoretical physics community to the ace model was generally not benign. Getting the CERN report published in the form that I wanted was so difficult that I finally gave up trying. When the physics department of a leading University was considering an appointment for me, their senior theorist, one of the most respected spokesmen for all of theoretical physics, blocked the appointment at a faculty meeting by passionately arguing that the ace model was the work of a "charlatan." The idea that hadrons, citizens of a nuclear democracy,

[19] Hanson concentrates on the discovery of the positron and I will mention this in section 3.4.

> were made of elementary particles with fractional quantum numbers did seem a bit rich. [Zweig, 1980]

Such hostility against fractional charges meant that quarks were either completely ignored or at best considered to be only mathematical entities for years to come.

**Quarks as mathematical entities.** Gell-Mann concludes his quark paper as follows:

> It is fun to speculate about the way quarks would behave if they were physical particles of finite mass (instead of purely mathematical entities as they would be in the limit of infinite mass). ...A search for [quarks] at the highest energy accelerators would help to reassure us of the non-existence of real quarks. [Gell-Mann, 1964, 215]

The view of quarks as "purely mathematical entities" was the common view at the time:

> Thus all mesons were accounted for by quark-antiquark combinations and all baryons could be viewed as three-quark structures. Most physicists have regarded this viewpoint as a mnemonic for SU(3) symmetry considerations, rather than viewing quarks as real physical objects. The strong belief that all true physical variables should be experimentally measurable was at the heart of this refusal to accept quarks as physical building blocks, because of their fractional electric charges and wrong spin-statistics relations. Nonetheless it won popularity because the quark model seemed to be the natural way to explain SU(3) representations, i.e. why representations other than 1, 8 and 10 have not been observed in hadron physics. [Horn, 2015, 109–113]

The co-inventor of quarks, Zweig, was a little more optimistic about the reality of quarks (aces). He finishes his paper as follows:

> Perhaps the model is valid inasmuch as it supplies a crude qualitative understanding of certain features pertaining to mesons and baryons. In a sense, it could be a rather elaborate mnemonic device.
>
> There is also the outside chance that the model is a closer approximation to nature than we may think, and that fractionally charged aces abound within us. [Zweig, 1964]

Zweig (along with Yoichiro Nambu) went on to become one of the few defenders of quarks as concrete entities for years to come [Johnson, 2000, 234,239,284,285]. The situation was not helped by the elusiveness of quarks: "By the summer of 1966, almost twenty experiments had failed to turn up a single one" [Johnson, 2000, 243] and they would keep evading the experimenters. In the meantime, Gell-Mann was not bothered by this failure:

> In a funny way, not finding quarks bolstered his [Gell-Mann's] argument that they were "mathematical," either an abstract accounting device or some philosophically maddening entity somehow trapped forever within the baryons and mesons. Whether quarks could be found in the laboratory or in outer space was, he kept insisting, irrelevant — as long as they helped make sense of the symmetries. ...
>
> In the summer of 1966, Gell-Mann was asked, during a discussion at the Ettore Majorana summer school in Sicily, whether he agreed that "the best feature about quarks is their name." "Yes," Murray replied. "The whole idea, as far as I introduced it (and I still think it is right), is that they are a useful mathematical

> creation in order to express the commutation rules of the currents and the approximate symmetry properties of the particle states. Maybe they are real things but probably not." [Johnson, 2000, 243–4]

The Nobel Prize in Physics 1969 was awarded to Gell-Mann without a mention of quarks "for his contributions and discoveries concerning the classification of elementary particles and their interactions." At this time, quarks were dubious at best and the Swedish physicist Ivar Waller who introduced Gell-Mann in the award ceremony mentioned that "The quarks are peculiar in particular because their charges are fractions of the proton charge which according to all experience up to now is the indivisible elementary charge. It has not yet been possible to find individual quarks although they have been eagerly looked for. Gell-Mann's idea is none the less of great heuristic value."

Martinus Veltman [2003, 240] recalls that Gell-Man was calling quarks "symbolic" in the early seventies and in a lecture given in 1972, Gell-Mann still denied the existence of quarks:

> [T]he quarks presumably cannot be real particles. Nowhere have I said up to now that quarks have to be real particles. There might be real quarks, but nowhere in the theoretical ideas that we are going to discuss is there any insistence that they be real. The whole idea is that hadrons act as if they are made up of quarks, but the quarks do not have to be real. [Reproduced in Fritzsch and Gell-Mann, 2015, 7]

He made similar comments in a conference in 1972 as well [Johnson, 2000, 275–6].

Gell-Mann's attitude towards the quark model was shared by the physics community but mid-1970s transformed its status.

**Quarks as physical entities.** In late 1960s, the so called *deep inelastic scattering* experiments shattered protons to see the patterns of the remaining debris. The results were interpreted as showing that a proton is not a point particle, but rather has an inner structure. In some popular accounts of the quark model, this development is falsely said to be a vindication of the reality of the quark model. For example one source writes that:

> At the end of the 1960s deep-inelastic scattering experiments at the Stanford Linear Accelerator Center (SLAC) showed for the first time that these quarks were not just hypothetical mathematical entities, but indeed the true building blocks of hadrons. [Flegel and Söding, 2014]

This is not true because, as I told above, Gell-Mann and other physicists were still lukewarm about quarks in early 1970s. In fact, the inner structure of protons were attributed to particles called *partons* by Feynman and the relation of these partons to quarks at the time was anything but identity.

What made the quark model successful and established quarks as physical entities was not a particular group of experiments, but rather the theoretical success and improvement of the model in 1970s. New experimental results kept pouring in and the simplest theory to deal with these turned out to be the improvements of the original quark model. This is a very complex story and I will only give a few headlines here.

The deep-inelastic scattering experiments showed that quarks carried only half of a proton's momentum and this missing momentum was attributed to a new kind of particle, *gluons*, which are the the mediators of the strong force, keeping quarks together. The quark-gluon model matured more and more throughout the 1970s, adding one success after another under its belt.

In the meantime, the problematic aspects of the model were ironed out. One nagging problem was that quarks coming together to form some particles had to be in the same state, violating the well-accepted Pauli exclusion principle. The way out was to introduce a new property of quarks, *color*, which guaranteed that they were in different states, saving the exclusion principle.

There was also the problem of free quarks — none was ever found. Why? The solution came with the concept of *asymptotic freedom* put forward by Frank Wilczek, David Politzer, and David Gross in 1973 which says that the attraction between quarks increased with distance (contrary to most forces). The *bag model* is an analogy employed since mid-1970s[20] that helps to explain asymptotic freedom by thinking quarks to be contained in an elastic bag or tied together with a rubber band. To break apart a pair of quarks, a lot of energy has to be put in stretching the band, and this increases the more you stretch. But the required energy to separate the pair exceeds the pair production energy of a quark-antiquark pair, so what you get in the end if you break the band is not a couple of disjoint quarks, but rather a multitude of pairs.

When Gell-Mann introduced quarks, there were only three of them. In 1974, two teams (at SLAC and Brookhaven) independently found a new particle, now called J/ψ. This particle caused quite a stir because it was so massive, more than three times than a proton, and it also had an unexpectedly long life for such a massive particle. As Michael Riordan puts it,

> it lived a thousand times longer than normally expected. Why did it refuse to play along with its buddies, who disappeared far more quickly? Something had to be inhibiting its decay, and that something might well be a new property of matter never seen before. [Riordan, 1987, 291–292]

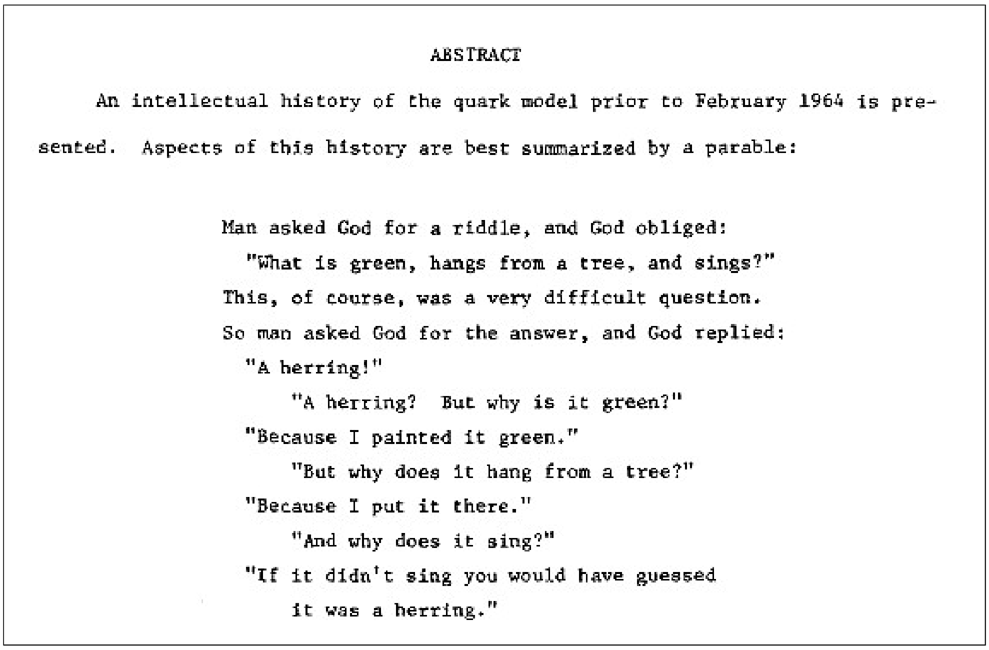
The discovery of J/ψ is known as the *November revolution.* The next two years saw discoveries of further massive particles that defied theory — except for the quark model. It was realized that the only available theoretical explanation of the new particles was by the quark model now enlarged by the addition of a fourth quark. Various alternative theories to account for the new particles were entertained, but eventually found to be inadequate. In the end, it was quarks that prospered:

> By the end of 1976, physicists the world over began to share a simple picture of the subatomic world, in which all matter was built of fundamental, pointlike quarks and leptons influencing one another by means of gauge forces. After all the confusion of the previous three decades, it was a great relief to have such a minimal, powerful theory. [Riordan, 1987, 321]

Quarks became one of the foundation stones of modern physics and no physicist called them fictions any more.

**Philosophy of quarks.** There are a number of philosophically interesting aspects of quarks.

[20] There are a number of such quark models, see for example [Thomas and Weise, 2001].



ABSTRACT

An intellectual history of the quark model prior to February 1964 is presented. Aspects of this history are best summarized by a parable:

Man asked God for a riddle, and God obliged:
"What is green, hangs from a tree, and sings?"
This, of course, was a very difficult question.
So man asked God for the answer, and God replied:
"A herring!"
"A herring? But why is it green?"
"Because I painted it green."
"But why does it hang from a tree?"
"Because I put it there."
"And why does it sing?"
"If it didn't sing you would have guessed it was a herring."

Figure 2.22: Zweig gave an invited talk titled "Origins of the Quark Model" at the Baryon 1980 Conference. The above image is the complete abstract of this talk. I take it that this parable shows how unlikely and unexpected the quark theory turned out to be compared to the past physics.

(1) The quark episode shows us how background or guiding assumptions (in this case about integer charges) can have strong influences. Past success of associated theories can entrench a guiding assumption, making new approaches quite hard to accept. In the case of quarks, it took physics community years of theoretical and experimental developments to finally move past beyond the notion of fractional charges. (See figure 2.22.)

(2) The status of mathematical entities is not a matter of theory but is a matter of acceptance. Whether an entity is considered to be mathematical or physical depends on the physicists' demeanor towards it. Even though two physicists disagree about an entity's status, they may both happily accept the associated theory. We saw that when Gell-Mann and most physicists considered quarks to be mathematical, Zweig, Nambu, and a handful of other physicists took them to be real. This situation of course reflect their confidence on the quark model and their philosophical views. In any case, physicists were happy to use the quark model. The difference between the two camps is not a theoretical one, but a pragmatic one.

(3) The more theoretical the physics gets, the harder it is to tell whether one *discovers* or *invents* new theories/models. Did Gell-Mann invent or discover quarks and their properties? The abstract parts of physics are so theoretical and the related experiments are so dependent on theoretical interpretation that the distinction between invention and discovery blurs. Moreover, a lot of metaphysical problems get in the door.[21] Metaphysical problems aside, I do not think that these questions can be answered in a meaningful way. Johnson sums up

[21]Here is a related metaphysical puzzle: Joe and Eve create (as far as they know) new molecules $M_1$ and $M_2$ respectively. It turns out that, unbeknownst to Joe, $M_1$ was already made in another lab a week ago. Did Joe invent or discover $M_1$? It also turns out that $M_2$ already exists in nature. Did Eve invent or discover $M_2$? Would your answers change if what they made were not molecules but some kind of blueprints?

the case about quarks as follows:

> As the particle physicists went on to elaborate their abstract, invisible world, it became less and less clear where to draw the line between something that was real and something that was mathematical. For many physicists, like Gell-Mann, this didn't seem to be a meaningful distinction. . . .
>
> For those who lived with this belief, it was harder than ever to tease apart the distinction between what was invented and what was discovered. Through the eyes of the willfully skeptical, the story of the Standard Model might be caricatured like this: According to the mathematics of group theory, the hadrons are built up from fractionally charged particles called quarks. Since quarks violate the Pauli exclusion principle, it was decreed that they come in three different "colors." Since no one could find particles with fractional charges, it was shown with some fancy mathematics that quarks are trapped inside the hadrons — that, going against all intuition, the strong force gets stronger, not weaker, with distance. And since electromagnetism and the weak force don't quite mesh, a new particle, the Higgs boson, was invented to break the symmetry between them. Finally, when new particles popped up that could not be accommodated with these contrivances, the theorists just kept adding quarks — charm, truth, beauty — until everything was accounted for. All this theorizing yielded testable predictions — particles that should show up in the accelerators. But when the experiments required so many layers of interpretation, how could the physicists know when they were reading too much into the lines and squiggles, seeing what their brains were primed to see, like pictures in the clouds? Were these really discoveries, or inventions? It was a question that Gell-Mann, among many others, refused to be distracted by, waving his prescription forbidding him from discussing philosophy. [Johnson, 2000, 296]

The issue of invention/discovery gets more complex as we move down the rabbit hole of physics, and at some point it stops to make sense. I will also touch on this issue in section 3.4.

(4) The more abstract the physics gets, the more theory-laden it becomes. There is so much interpretation involved in the experiments that one cannot talk about different realms of theory and experiment. Quarks are a case in point: In the monumental *Review of Particle Physics* by the Particle Data Group, we find the following paragraph in the section *Quark Masses:*

> Unlike the leptons, quarks are confined inside hadrons and are not observed as physical particles. Quark masses therefore cannot be measured directly, but must be determined indirectly through their influence on hadronic properties. Although one often speaks loosely of quark masses as one would of the mass of the electron or muon, any quantitative statement about the value of a quark mass must make careful reference to the particular theoretical framework that is used to define it. It is important to keep this *scheme dependence* in mind when using the quark mass values tabulated in the data listings. [Olive et al., 2014]

As Johnson notes above, "experiments required so many layers of interpretation" that we should not think that experimental results stand on their own, free from any theory.

(5) Physicists do not always believe their theories literally. They might only consider them as useful tools. This instrumental approach can be revised in time. Those theories that served well can be discarded if their usefulness ceases or more useful theories come along. In some cases, like the quarks, what started as a fiction can "become instead a boy" and no more considered a wooden puppet.

Of course, the instrumentalism discussion is a very old one [see Duhem, 1969]. But the quark case shows that scientists not only try to *save the phenomena*, but they also they also try to *save the theory* by postulating mathematical entities.

## 2.12 Neuron Doctrine

> But the ideas marched on, the ideas marched on, just as though men's brains were no more than stepping-stones, just as though some great brain in which we are all little cells and corpuscles was thinking them!...
>
> H. G. Wells, *The New Machiavelli*, 1910

In this section I will look at the dispute about the neuron doctrine which features the Italian scientist Camillo Golgi (1843–1926) and the Spanish scientist Santiago Ramón y Cajal (1852–1934) who jointly received the Nobel Prize in 1906 "in recognition of their work on the structure of the nervous system". The period leading up to this prize is told in *Encyclopedia of Medical History* as follows:

> The development of cellular theory in the first half of the nineteenth century owed to improved microscopy and particularly the achromatic lens. Jan Evangelista Purkinje, a Czech working at Breslau, gave the first detailed account of nerve cells in 1837, demonstrating that they had a protoplasm-filled main body which enclosed a central body and a fibrous "tail" extending from the main body. The nature of the tail, or "process," was not clear, though von Helmholtz thought that in many cases nerve fibers were composed of nerve cell processes, a view supported by more detailed work on nerve fibers by the Danish student A. H. Hannover. But the basic problem of the nerve cells' role in electrical action and the related problem of whether or how the nerve cells were connected remained unanswered.
>
> Progress on understanding nerve cell function began to be made when methods for fixing and sectioning tissue for microscopic study improved in the 1850s and 1860s, while Joseph von Gerlach's carmine stain (1854) greatly improved contrast. It was still very difficult to observe nerve connections through the microscope, however, though it had become clear that nerve cells had each one main process and that "nerves" were bundles of these processes. But there agreement stopped. One school of thought adhered to a theory set out by von Gerlach that the nerve cells were connected by a network of fibers. The network served to conduct impulses from cell to cell. This was known as the reticularist theory. Their opponents criticized the argument without having any clear alternative, though in 1887, Wilhelm His, a noted histologist, and Auguste Forel suggested that nerve cells not only were not connected but also had "free endings" in the central nervous system's gray matter. Available microscopy neither confirmed nor denied their views, and the reticular theory gained new support from findings by Camillo Golgi, who had discovered a new stain for central nervous system tissue. The Golgi stain, which required a week to prepare, revealed the nerve

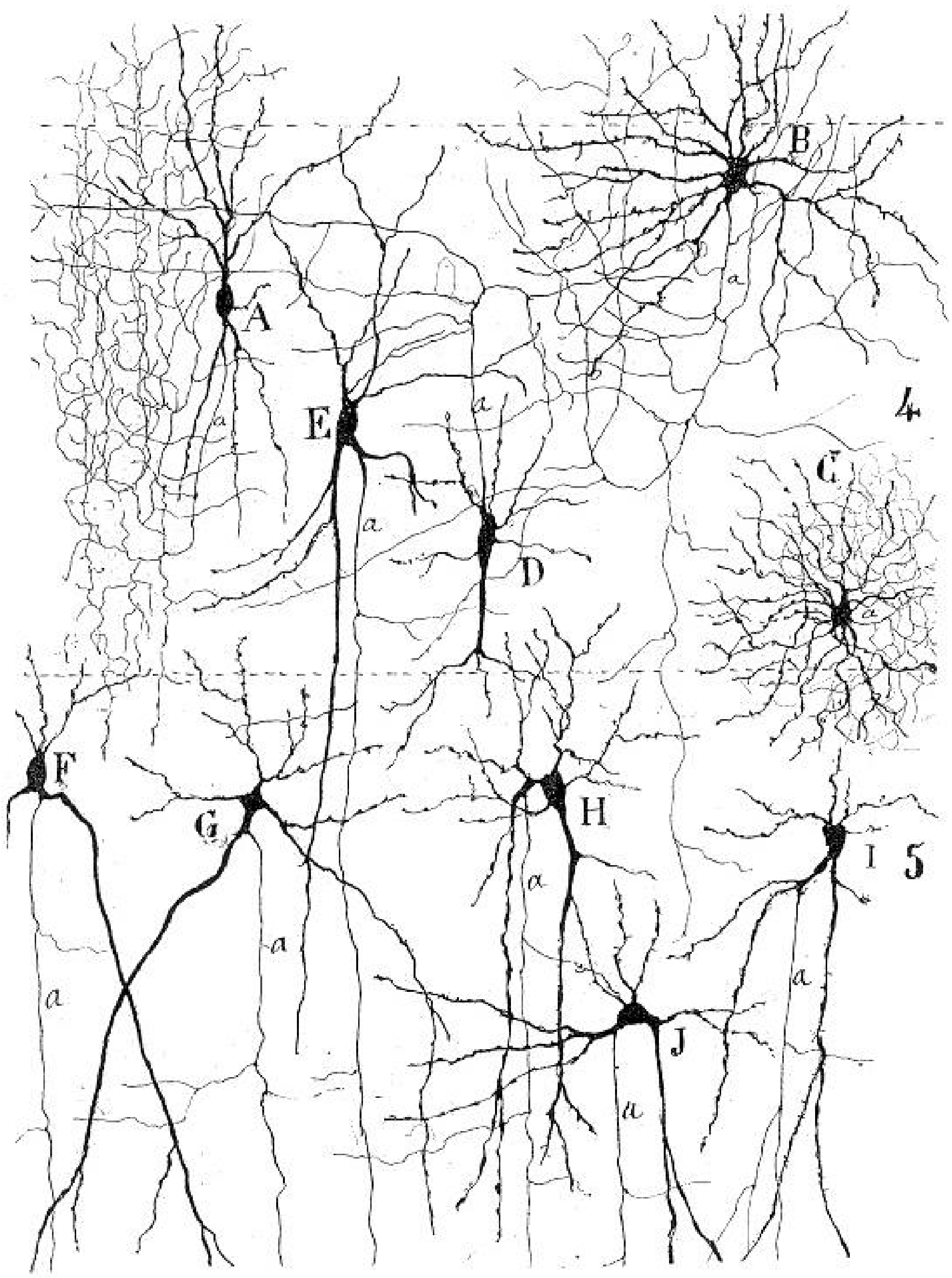



Figure 2.23: An illustration of nerve cells by Ramón y Cajal from *Texture of the Nervous System of Man and the Vertebrates*, 1898.

> cell clearly, showed its process, and for the first time revealed smaller processes (dendrites). But the junctions remained unclear. Golgi, nevertheless, supported a reticularist position which was challenged by an obscure Spanish histologist, Santiago Ramon y Cajal. Cajal became a master microscopist whose ambition was to unravel the problem of the brain cell, what he called "the aristocrat among the structures of the body" containing the promise of "knowing the material course of thought and will." Cajal, who shared a Nobel prize with Golgi in 1906, became the Italian's critic and competitor. He improved the Golgi stain and, working systematically with embryonic tissues of birds and small mammals, began to develop new images and descriptions which he published at his own expense. In 1889, he went to the Berlin conference of the German Anatomical Society, where his demonstrations, delivered in halting, clumsy French, astounded and converted some of Europe's leading biological scientists, including Rudolf von Koelliker. Cajal's theory became the basis for what Wilhelm Waldeyer named the neuron doctrine. It held that each nerve cell is a self-contained unit with an axon reaching toward but not connected with another cell. Nerve fibers are composed of these processes, but it is the cell which forms the communicating links in nerve tissues. The processes are insulated against each other except for the axon end, which makes contact with the next cell. Cajal's 1889 presentations supported the views outlined by His and Forel and swung the weight of evidence against the reticularist theory. Cajal continued to build on this early work, improving his stains and contributing to a fuller understanding of nerve pathways and the flow of electrical impulse. But his primary contribution, as Sir Charles Sherrington pointed out, was that he "solved at a stroke the great question of the direction of nerve currents" and swept away the reticularist theory. Cajal held that nerve circuits were valved, and he located the valves "where one nerve cell meets the next one." It was Sherrington who named that valvular connection, later found to be a gap, the synapse. [McGrew, 1985, 211-212]

At the end of the nineteenth century there were two competing paradigms: the reticular theory and the neuron doctrine. The first one, championed by Golgi, holds that the nerve fibers form a connected web forming a unified entity. This was a holistic approach that saw the nervous system as a single continuous network. The neuron doctrine, championed by Ramón y Cajal, on the other hand replaced the unitary approach by a reductionist one:

> The neuron doctrine recognized the nerve cell as the developmental, structural, functional and trophic unit of the nervous system and insisted that nerve cells communicate at sites of contiguity not continuity. Together with the law of dynamic polarization, which is often included as a part of the doctrine and which recognized the dendrites and the cell body as the receptive surface of a nerve cell with the axon serving as the (single/unitary) effector portion of the cell, the neuron doctrine formed a powerful tool for analyzing the nervous system. [Guillery, 2007, 412]

The neuron doctrine contained more than "cell theory applied to nervous systems". The law of dynamic polarization[22]

[22] Giovanni Berlucchi [1999] investigates the history of the the law of dynamic polarization including the contribu-


Figure 2.24: An 2001 issue Angolan stamp showing a portrait of Ramón y Cajal with the erroneous "Camillo Golgi" caption. A mistake fitting for joint Nobel Prize winners. See the article by Triarhou and del Cerro [2012] which gives detailed information about this stamp.

> recognizes dendrites and cell bodies as the receptive surface of nerve cells and the axon as the effector surface. . . . and added very significantly to the analytical power of the neuron doctrine because it provided crucial clues regarding the direction in which messages pass through the nervous system from one cell to another. [Guillery, 2007, 414]

The doctrine also requires that neurons are not connected to each other. The gap between them was eventually observed in mid-1950s using electron microscopes. But the neuron doctrine was well accepted by that time and the observation was a final confirmation of it.

Even though Golgi and Cajal were the first scientists ever to ever share a Nobel Prize, they did not share the same view. Golgi was an adamant supporter of the reticular theory. His staining method considerably improved the visibility of the neuron fibres in microscopes. But the images were not enough to settle the matter and what different scientists took those images to be were very different. Whereas Golgi was seeing a reticulum, Ramón y Cajal was seeing nerve cells. They received the Nobel Prize jointly for "their work on the structure of the nervous system", though only one believed in the neuron doctrine. "Golgi resented that by using his method Cajal had arrived at a view of the nervous sytem that was opposite to his own, and surely could not be pleased that physiologists seemed to be more supportive of Cajal than of himself." [Berlucchi, 1999, 200] (See figure 2.24.) Essentially, one of them was getting a Nobel Prize for his stain, the other was getting it for his *interpretation* of the stained samples.

But what was it that made one interpretation better than the other? *Not* because there was strong evidential ground for any one of them. The real reason is that the neuron doctrine

tions of William James and Charles Scott Sherrington which I neglect here.

turned out to be a successful theory to analyse structures of the neural system which are classified by R. W. Guillery [2007, 416] into the following headings: *The synapse, The motor unit, Neural degeneration and axonal transport, Developmental studies, Molecular markers.* Guillery succinctly shows the presupposition and success of the neuron doctrine in these areas and argues that "a reticularist view would have provided an inadequate guide for understanding" the structures in question. I will not repeat the success stories of the doctrine here and I refer you to Guillery's text for details. What is important for us here is that the neuron doctrine was the more fruitful and explanatory interpretation, and the reductionist programme of understanding the nerve system in terms of simple neurons have worked wonders.

The neuron doctrine has changed throughout the twentieth century by addition or removal of different elements. For example, *Dale's law* of 1935 "which stated that a single nerve cell produces the same transmitter at all of its axon terminals" was "modified when it was discovered that a single nerve cell could produce and release more than one transmitter" [Guillery, 2007, 418]. Neurons became the basic information processing unit sometime in mid-century. Even some of the main tenets of the doctrine were challenged by findings such as fused neurons, gap junctions, serial synapses, back-propagation of action potentials and so on. Metabolic subunits within the neuron were discovered. The neuron doctrine at the end of the century was much different then the one at the start. Slowly and regularly, one version of it gave way to a newer version.

But what is the neuron theory used today? Guillery [2007] argues that the neuron doctrine is today reduced to only cell theory and its parts that surpasses cell theory have fizzled out: "where it goes beyond the cell theory, it can no longer be defended on the basis of contemporary evidence." The reductionist approach of the neuron doctrine has been hugely successful, but its limits are acknowledged today.

> The neuron doctrine has to be seen as a tool, and should not to be regarded merely as an accurate and single view of what all neurons are 'really' like. The doctrine has proved particularly useful in the analysis of long pathways, and where it seems to be weakest today is in the study of local circuits. We have to recognize that the neuron doctrine has been extremely useful in the past, that it continues to serve us as a practical conceptual tool today especially for the study of long pathways and their development, and that its role in providing a general abstract view of some ideal that would fit all nerve cells has never been an important practical part of its function. [Guillery, 2007, 416–417].

There are a number of philosophically interesting aspects of this case:

(1) The neuron doctrine shows how a theory can change steadily for over a considerable time. In section 3.4, I introduce the name "extended building" for this kind of theory change. We might never see the complete theory or at least for a long time. By influencing each other, experiment, theory, interpretation, and inter-theoretic relations all contribute to the gradual change.

(2) The reductionist program of understanding a complex structure by simple units is an important part of science; and, in the neuron doctrine case, a very successful one. There is never a guarantee that the committal to a reductionist program will succeed since it may turn out that the complex structure defies the simplistic attack. But, then again, there are no guarantees in science that *any* approach will be successful. Reductionism is an experiment

in seeing how much way we can make in explaining the structure in question using simple things. Hopefully, all the way. If not, we will still learn a lot along the way.

Can we see reductionism as a case of parsimony? Strictly speaking, no. Parsimony tells us that nature *is* simple. Reductionism tells us to analyze nature (which can be complex) in simple terms.

(3) Scientists are mindful of the limits of their theories. The usefulness of the neuron doctrine has reduced and there are areas where it is more useful than others. It is kept as a "practical conceptual tool" but exceptions are well-known and expected.

(4) Experiments and observations are influenced by theory. At the end of the end of the nineteenth century and early twentieth century, supporters of the reticular theory and the neuron doctrine were interpreting what can be seen in microscopes quite differently: reticulum vs. neurons. This is an example that shows that evidence is *theory-laden* (see section 3.4).

(5) It is quiet telling how Ramón y Cajal philosophizes about the luxury of evidence:

> It is a rule of wisdom, and of nice scientific prudence as well, not to theorize before completing the observation of facts. But who is so master of himself as to be able to wait calmly in the midst of darkness until the break of dawn? Who can tarry prudently until the epoch of the perfection of truth (unhappily as yet very far off) shall come? Such impatience may find its justification in the shortness of human life and also in the supreme necessity of dominating, as soon as possible, the phenomena of the external and internal worlds. But reality is infinite and our intelligence finite. Nature and especially the phenomena of life show us everywhere complications, which we pretend to remove by the false mirage of our simple formulae, heedless of the fact that the simplicity is not in nature but in ourselves.
>
> It is this limitation of our faculties that impels us continually to forge simple hypotheses made to fit, by mutilating it, the infinite universe into the narrow mould of the human skull, — and this, despite the warnings of experience, which daily calls to our minds the weakness, the childishness, and the extreme mutability of our theories. But this is a matter of fate, unavoidable because the brain is only a savings-bank machine for picking and choosing among external realities. It cannot preserve impressions of the external world except by continually simplifying them, by interrupting their serial and continuous flow, and by ignoring all those whose intensities are too great or too small. [Ramón y Cajal, 1899]

The context of this quotation is the neuron doctrine and Ramón y Cajal admits that he did not accept the neuron doctrine solely on evidential grounds. Scientists do not restrict themselves to evidential values. As this case highlights, usefulness, fruitfulness, explanatory power, etc. are all essential values in science. Moreover, discovery can be a very complex process, depending on interplay of all kinds of values.

(6) I want to end this section with a couple of quick pointers. Consider the view that scientists always consider all the theories related to their research and make well-informed decisions/choices. This is not always true:

Figure 2.25: Camillo Golgi in 1889. [Image in the public domain.]

> Golgi and Cajal alike were mostly oblivious to the work and ideas of contemporary physiologists and psychologists, possibly because they felt that the nervous system could be fully understood by means of morphological investigations alone. As a result, they held wrong functional views inspired by morphology that could have been corrected by paying due attention to existing physiological evidence or common-sense behavioral considerations. [Berlucchi, 1999, 191]

Giovanni Berlucchi [1999, 191] gives an analysis of this case. For a recent philosophical discussion of the "unconceived alternatives" see the book *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives* by P. Kyle Stanford [2010].

(7) Scientists cannot always give a complete explanation or detailed mechanism of something. In those cases an incomplete or partial explanation/mechanism might be the aim of the research. As we have seen in section 2.3, there are cases that scientists might not even be trying to give an *actual* mechanism, but only a *possible* mechanism.

In the last decade, the topic of mechanisms in science has been a hotly debated issue in philosophy of science. Interestingly, examples from neurology are frequently met in this literature. This is not surprising because the history of this disciple is saturated with attempts to explain various neural structures and all kinds of mechanisms are offered. See, for example, [Colombo et al., 2014].

## 2.13 Pharmaceutics

> Proper treatment will cure a cold in seven days, but left to itself, a cold will hang on for a week.
>
> Darrell Huff

If there is one section of this text that I am not happy about then it is this one. The reason is not that it presents an important or relevant case study, but rather it is a topic which cannot


Figure 2.26: Stamp on drug-making issued by Japan in 1986 features what can be considered as the main symbol of pharmacology: a pill.

be done justice to in a short section or even a book.

The dark side of the medical establishment was publicly known to a degree but in the last ten years different aspects of it were scrutinized in numerous books which made the situation shockingly grave.[23] It was Marcia Angell's *The Truth About the Drug Companies: How They Deceive Us and What to Do About It* [Angell, 2004] which started the avalanche. She is an American physician who served as an executive editor and then as the editor-in-chief of the prestigious journal *The New England Journal of Medicine*. When she published her book, she opened the floodgates. Numerous other researchers joined her detailing the awful state of medicine.

The last but not the least in this chain of writing is British physician and writer Ben Goldacre's well-researched 2012 book *Bad Pharma* which shows the severity of the problems.[24] He does a great job of improving and putting together all investigations on this topic together and the picture he meticulously comes up with is shocking. In the introduction of *Bad Pharma*, Goldacre puts forward his main thesis as follows:

> Drugs are tested by the people who manufacture them, in poorly designed trials, on hopelessly small numbers of weird, unrepresentative patients, and analysed using techniques which are flawed by design, in such a way that they exaggerate the benefits of treatments. Unsurprisingly, these trials tend to produce results that favour the manufacturer. When trials throw up results that companies don't like, they are perfectly entitled to hide them from doctors and patients, so we only ever see a distorted picture of any drug's true effects. Regulators see most of the trial data, but only from early on in a drug's life, and even then they don't give this data to doctors or patients, or even to other parts of government. This distorted evidence is then communicated and applied in a distorted fashion. In their forty years of practice after leaving medical school, doctors hear about what works through ad hoc oral traditions, from sales reps, colleagues or journals. But those colleagues can be in the pay of drug companies — often undisclosed — and the journals are too. And so are the patient groups. And finally, academic

[23] Some of the books on this topic are [Abramson, 2013; Angell, 2004; Davies, 2013; Goldacre, 2013; Goozner, 2005; Kassirer, 2005; Rost, 2006; McGarity and Wagner, 2008] I have especially made use of Angell's and Goldacre's book.

[24] Note that different printings of *Bad Pharma* have different subtitles. My page references are to the revised 2013 edition with the subtitle *How Medicine is Broken, and How We Can Fix It.*

> papers, which everyone thinks of as objective, are often covertly planned and written by people who work directly for the companies, without disclosure. Sometimes whole academic journals are even owned outright by one drug company. Aside from all this, for several of the most important and enduring problems in medicine, we have no idea what the best treatment is, because it's not in anyone's financial interest to conduct any trials at all. These are ongoing problems, and although people have claimed to fix many of them, for the most part they have failed; so all these problems persist, but worse than ever, because now people can pretend that everything is fine after all. [Goldacre, 2013, x–xi]

Goldacre goes on to successfully and thoroughly document and defend these assertions in the 400+ page book. The book paints a very grim but complex picture that is impossible to to capture here. Rather, I will look at a single problem that Goldacre mentions: *publication bias*. But this is just the tip of the iceberg and I refer you to Goldacre's book for the whole story.

Publication bias can inflict both physical and social sciences.[25] This type of bias can be seen when there is a discrepancy between available and published data:

> Publication bias is the term for what occurs whenever the research that appears in the published literature is systematically unrepresentative of the population of completed studies. Simply put, when the research that is readily available differs in its results from the results of all the research that has been done in an area, readers and reviewers of that research are in danger of drawing the wrong conclusion about what that body of research shows. In some cases this can have dramatic consequences, as when an ineffective or dangerous treatment is falsely viewed as safe and effective. [Rothstein et al., 2005, 1]

Publication bias is especially important in medicine as doctors can only make judgements based on the evidence that is available to them. If they receive incomplete or distorted evidence then their decisions can threaten the health and even the lives of their patients. And unfortunately there is ample evidence that shows that publication bias is widespread.

To begin with, drug companies carry out or sponsor numerous trials but only publish those with favourable results. If, for example, there are seven unfavourable and three favourable trials, doctors only get to hear about those three. Not only that, the fact that there are seven unpublished trials is kept secret. That is, as far as doctors know, there are only three trials and they are favourable. The situation is not helped by the fact that there is more variance in the results if the trials are carried on only a small number of patients. Just pick and choose the suitable ones and throw away the rest.

Another issue is the difficulty of publishing negative or repeat trials. Medical journals tend to accept new and exciting results more than old and boring ones. If a research group publishes a favourable result about a new drug, then their research must be worth checking and publishing to make sure that they got their results right. But if another group carries out that trial again then they will have a hard time finding a journal that will publish the repeated result. This gets in the way of correcting the published results. Why would anyone carry out a trial if they are unlikely to get published? This is one of the reasons of the reproducibility problem which I will look at in section 4.3.

[25] See http://www.nature.com/news/social-sciences-suffer-from-severe-publication-bias-1.15787 for a discussion of publication bias in social sciences. For an example from marine science see http://www.eurekalert.org/pub_releases/2016-02/oup-ijo022916.php. Retrieved on February 29, 2016.

A third issue is related to the changes that can be made to trials to make them look more effective. Frequently, a sponsoring company of research group gets to see the data as its collected and has the right to stop the trial at any time. As Goldacre [2013, 39–40] puts it "if you stop a trial early because you have been peeking at the preliminary results, then you can either exaggerate a modest benefit, or bury a worsening result." Moreover, the fact that the sponsoring company had the right to stop the trials at any time is not mentioned in the published papers.

There is another problem related to sponsoring companies: confidentiality and control clauses in agreements that prevent researchers and academicians from disclosing various aspects of the research and blocking the full access of the researcher to the data. Both can get in the way of open availability of unfavourable results.

Finally: "At the end of your trial, if your result is unimpressive, you can exaggerate it in the way that you present the numbers; and if you haven't got a positive result at all, you can just spin harder." [Goldacre, 2013, 218]

All these issues result in a publication bias in medicine that make negative or unfavourable results never see the light of the day. Here I have only hastily touched a few of these ways without giving any concrete examples. I just opened the Pandora's box and the rest is in the books I mentioned.

There are a few points to consider. To begin with, there are biases in sciences, but, as far as we know, a lot more in some then others. Medicine, in particular, is inflicted by many, starting with the publication bias I examined here. But there is a difference between knowing a bias and its dissolution. Further in this text we will see a science with biases waiting to be overcome, that is, researchers are actively looking for a way to abate biases. The difference in the medicine case is that the way to deal with the biases is known but *not* implemented (yet). Goldacre makes an impressive list of suggestions which would improve the situation drastically. Even the drug companies acknowledge some of these possible improvements and make promises to implement them in the future.

We can therefore talk about five types of bias:

(1) There is an unknown bias in a science.

(2) The bias is known but (at least for the time being) ignored.

(3) The way to diminish the bias is unknown but actively searched for.

(4) The way to diminish or overcome the bias is known but not implemented.

(5) The way to diminish or overcome the bias is known and it is implemented.

Historical and current science show examples of biases from each of the five categories. Of course, we can know the first type only in hindsight. Provided that it is a known bias, philosophers, sociologists, and researchers can determine the place and importance it has in the research tradition and also which type it is.

Unfortunately, biases in pharmaceutics stagnated at the penultimate level. But there is hope as more and more articles and books are drawing attention to the current failure of action. Two of the positive developments is the establishment of a journal for repeat and negative trials and also the announcement by some of the editors of leading journals on


Figure 2.27: The Berlin stamp issued in 1980 to commemorate Alfred Wegener and shows continents fitting each other.

improving their acceptance criteria. But "big pharma" is yet to respond conclusively to these problems.

Topics discusses in this section naturally lead to the reproducibility problem which I will discuss in section 4.3. Also appendix B contains a list of ways double-blind clinical trials can go wrong in medicine.

## 2.14 Continental Drift

> "Impossible" is usually defined by our theories, not given by nature.
>
> Stephen Jay Gould [1979, 165]

It is very likely that school children all around the world make the observation that continents fit each other like jigsaw puzzle pieces; at least I did when I was a kid. (See figure 2.27.) This phenomenon was written by numerous scholars as well. But no one made much out of it until Alfred Wegener (1880–1930) published his seminal book on the subject in 1915: *Die Entstehung der Kontinente und Ozeane*. There is an English translation [Wegener, 1966] of the fourth German edition of 1929 and figure 2.28 lists the chapter headings of it.

Wegener's book is a brilliant compilation of evidence for continental drift. As you can see from the table of contents, he uses evidence from different branches of geoscience to backup his argument. His tour de force turned continental drift from children's play to serious science. But there is something very surprising about its reception: Whereas European scientists were more than eager to jump on the bandwagon, American scientists brushed off the idea until late 1960s. Initially, all American contemporaries of Wegener bar a few exceptions rejected his theory. The British, on the other hand, were "cautiously receptive" [Oreskes, 1999, 125].

(1) Historical Introduction

(2) The Nature of the Drift Theory and Its Relationship to Hitherto Prevalent Accounts of Changes in the Earth's Surface Configuration in Geological Times

(3) Geodetic Arguments

(4) Geophysical Arguments

(5) Geological Arguments

(6) Palæontological and Biological Arguments

(7) Palæoclimatic Arguments

(8) Fundamentals of Continental Drift and Polar Wandering

(9) The Displacement Forces

(10) Supplementary Observations on the Sialsphere

(11) Supplementary Observations on the Ocean Floor

Figure 2.28: The chapter headings of *The Origin of Continents and Oceans* [Wegener, 1966].

More and more Europeans started to support the theory in the next three decades but the Americans continued to keep their distance and ridiculed Wegener's ideas.

We have seen national divisions before in this text. In section 2.1, we saw how Israeli convention is forcing differs from the world. I also told the well-known case of the Piltdown man in section 2.10. The difference is that forcing is benign, that is, one approach is just a linguistic variant of the other and does not lead to controversy in any way. On the other hand, Piltdown man was a real difference in scientific opinion and we saw what made the British stand out. The theory of continental drift caused a divide of this kind: Whereas Europeans believed in moving continents, Americans took it to be a joke. This is how Stephen Jay Gould [1979] recalls the standing of continental drift in American science in early 1960s:

> Kenneth Caster, the only major American paleontologist who dared to support it openly, came to lecture at my alma mater, Antioch College. We were scarcely known as a bastion of entrenched conservatism, but most of us dismissed his thoughts as just this side of sane. ...A few years later, as a graduate student at Columbia University, I remember the a priori derision of my distinguished stratigraphy professor toward a visiting Australian drifter. He nearly orchestrated the chorus of Bronx cheers from a sycophantic crowd of loyal students. [Gould, 1979, 160]

Naomi Oreskes is a a geologist turned historian of science whose book *The Rejection of Continental Drift: Theory and Method in American Earth Science* [Oreskes, 1999] is my main source in this section. She agrees with Gould that

> [In America] continental drift was widely discussed and almost uniformly rejected, not merely as unproved, but as wrong, incorrect, physically impossible, even pernicious. American scientists were much more hostile to the idea than their European counterparts; some even labeled the theory *unscientific.* [Oreskes, 1999, 5]

But Oreskes and Gould have a difference of opinion about *why* the scientific communities on the two sides of the Atlantic did not see eye to eye.

According to Gould, continental drift "was dismissed because no one had devised a physical mechanism that would permit continents to plow through an apparently solid oceanic floor. In the absence of a plausible mechanism, the idea of continental drift was rejected as absurd." [Gould, 1979, 161] According to this reasoning, the cause of drift was unknown, so it was rejected. There were different attempts at giving the driving mechanism of drift by various scientists (including Wegener himself) but none was found to be satisfactory until late 1960s. It was this lack of mechanism that turned scientists away from the drift theory.

Gould's *lack of mechanism* explanation for the American rejection of continental drift is completely wrong for a number of reasons. First, geologists were quite open to theories that did not have explanatory cause. Oreskes gives the following examples:

> The ice ages (whose causes are still being debated today), the Alpine overthrusts (whose existence helped stimulate the drift debate), and geomagnetic polarity reversals (which were critical to the establishment of plate tectonics) are three important examples; there are many others. The late Marshall Kay, Professor of Geology at Columbia University, went so far as to describe the operating premise of geologists as "anything that has happened, can."
>
> There was a long tradition in geology of accepting the reality of phenomena without requiring causal accounts of them. [Oreskes, 1999, 63]

So even if geologists had not had a mechanism, they would not necessarily have rejected drift. But this brings us to the second point: Geologists did indeed have various theories for the cause of drift:

> By 1929, three powerful theories — Daly's gravity sliding, Joly's periodic fusion, and Holmes's subcrustal convection currents — had been developed to explain the kinematics of drift. All were consistent with the known physical properties of the earth. Moreover, Daly and Holmes offered dynamic explanations as well. None of the theories came as an ad hoc adjustment to Wegener's proposals[.] . . . All three were developed by prominent and eminent scientists; all were published in readily accessible form; and all were widely known and discussed at the time.
>
> Together, the three theories incorporated the fundamental aspects of modern theory: a rigid moving surface riding on convection currents in a weak zone beneath, with portions recycled into the substrate. Not one of these theories required the continents to plow through the rigid ocean floor. . . . The point is that *the same driving force that is generally accepted today for plate tectonics was proposed in the 1920s.* The theory of continental drift did not fail for lack of a mechanism. [Oreskes, 1999, 119–120]

Just as Einstein could not have come up with relativity theory 50 years earlier, Wegener could not have done so either. The reason is that Wegener owed much to his scientific milieu just as Einstein did. There were already theories around that allowed some movement of the world, though not to the same extent. There were lots of evidence of local movement. There were numerous evidence of continuity between continents. Most importantly, *isostasy* was known and accepted to be true:

> By the time Wegener's Origin of Continents and Oceans was translated into English in 1924, a rough consensus had emerged among geologists in the United States and Europe that the earth contained a mobile layer beneath the crust that provided both the seat of isostatic adjustment and the kinematic explanation for surficial horizontal dislocations. By imagining the mobile zone as either a plastic solid or a highly viscous fluid, geologists reconciled their empirical observations with the theoretical constraints provided by physics and astronomy. Continents were not viewed as rigid, fixed bodies, but more like large rafts. Their vertical motions were accepted as fact, and from these vertical motions, Hayford and others had argued, horizontal dislocations arose as a consequence.
>
> Wegener built his theory on this consensus but challenged its last aspect. Rather than viewing horizontal displacement as a side effect of vertical oscillation, he viewed it as a fundamental process in its own right. If continents could move vertically through the substrate, he argued, then at least in principle they could move horizontally as well. Thus the novel point in Wegener's argument was not the idea of horizontal mobility per se but the scale and extent of that mobility. Given a mobile substrate, large-scale horizontal motions were at least plausible. [Oreskes, 1999, 77–78]

Far from "plowing through a solid oceanic floor", Wegener employed and extended isostasy theory accepted at the time to argue that continents were plowing through a *fluid* oceanic floor. Without considering the historical context of Wegener's theory, it is easy to think that Wegener came up with a crazy theory out of the blue. Nothing can be far from the truth.

Finally, Gould's account fails to explain the asymmetric views on the two sides of the Atlantic. Even if we accept the counterfactual that there was no mechanism given for drift, this still does not answer the fundamental puzzle: Why did only Americans reject drift theory? Having a mechanism or not is something symmetrical, that is, both Americans and Europeans shared this knowledge. Something symmetrical cannot explain the asymmetric views.

Having dismissed Gould's *lack of mechanism* explanation for the American rejection of drift, let us move on to Oreskes' explanation for it. Oreskes crucially draws attention to the fact that the difference of opinion regarding drift is not a one isolated instance. Indeed, Americans and Europeans had different views regarding various theories in geology. Just to give one example:

> At the end of the nineteenth century, geology was an international science, and geologists frequently traveled abroad to visit colleagues and to see sites of interest. Nevertheless, European and American geologists found themselves subscribing to incompatible views of earth evolution. From the same starting point — the secular cooling of the earth — two different pictures emerged. In the European view, the earth was in a state of continual flux with complete interchangeability of its parts. Ocean basins could be elevated into continents, continents could collapse to form ocean basins, and change occurred across the globe. In the American view, the basic outlines of the earth had been set at the beginning of geological time and had not changed fundamentally since then. Continents were always continents, oceans were always oceans, and change was confined to discrete zones at the interface between them. The two theories also differentially weighed the available facts. The American perspective emphasized the physical properties of minerals, the contrasting compositions of continental rocks and the ocean

> floor, and the asymmetry of folding in the Appalachians. The European view emphasized the biogeographical patterns, the stratigraphic evidence of interchangeability of land and sea, and the diverse patterns of folding in European and African mountain belts. [Oreskes, 1999, 19]

Oreskes details the differences between the two research traditions extensively which I will not repeat here. This backdrop is crucial in understanding the American rejection of drift. It turns out that there were already considerable differences between the American and the European geology before 1910s. Europeans already entertained theories that were much closer in spirit to Wegener's. Moreover, the "biogeographical patterns and the stratigraphic evidence" preferred in Europe was fulfilled in Wegener's writings.

There is one point repeatedly brought up in American reviews of Wegener's work: that it is *unscientific.* Why did Americans insist that the drift theory is unscientific? This makes sense only in the light of another point Oreskes makes: American geologists has much different methodological commitments and explanatory frameworks than their European counterparts to degree that they found the drift theory unacceptable on such grounds. What were these commitments?

(1) A salient problem for American geologists was to develop an independent science standing on its own, that is, independent from its European roots. For example, influential American geologist T. C. Chamberlin required the next generation "be individual and independent, not the mere following of previous lines of thought ending in predetermined result." [Quoted in Oreskes, 1999, 138] This attitude already shows a tendency to avoid foreign approaches, but more importantly, it led to a development of an American way of doing geology.

(2) Being independent meant to avoid "excessive respect for European authority" and "to develop their own science on their own ground". [Oreskes, 1999, 135]

(3) American geology was funded by "governmental agencies with practical mandates" which "had an impact on the work that Americans did and the way they did it." [Oreskes, 1999, 133]. This tended to put practice ahead of theory.

(4) American geologists entertained an inductive logic of justification that put observation before theory. They "believed that reliable knowledge is grounded in observation" [Oreskes, 1999, 144].

(5) There was a suspicion towards all-encompassing theoretical systems amongst American geologists. These types of universal theoretical systems were known to fail and they were "no better than propaganda". "Such behavior was considered inimical to scientific discovery." [Oreskes, 1999, 134]

(6) Americans followed Chamberlin's theoretical pluralism which required there to be a number of "working hypothesis" in order to get best science possible. Only when observations are considered in the light of alternative hypotheses, impartial progress can be achieved. Debate of different hypotheses nurtures discovery, develops potentials to a fuller extent, and eliminates weak spots.

Oreskes elaborates these points to considerable detail and gives numerous examples from the works and writings of American geologists. She concludes that "to accept these [Wegener's] ideas in the 1920s or early 1930s would have forced American geologists to abandon many fundamental aspects of the way they did science. This they were not willing to do." Wegener's theory was deemed *unscientific* by Americans because that is not how they did science.

I have only glanced over Oreskes' reasoning above but I do not need to get in the nitty-gritty of her account to show the difference of her account to Gould's. The latter sees the reasons of rejection of the drift theory in the theory itself: it required "plowing through a solid oceanic floor" and an account of the mechanism was lacking. Not only Gould was historically wrong, he was also *logically* wrong: a reason internal to the drift theory cannot explain the differences of the two geology communities. Even though were are talking about times well-before the fast internet access to knowledge era, the theory and evidence for it was out there and known by the international geology community. If one side accepts and the other side rejects the *same* theory, then there must be something external to theory that causes this. Contrary to Gould, Oreskes finds reasons for the difference in the historical background and the cultural context of the theory as well as the practices and traditions of the community.

# CHAPTER 3

# PHILOSOPHY OF VALUES

In the previous chapter, we saw values and practices in action in many different cases from science. Scientists appeal to different values when they work and this observation must be taken as a *reductio ad absurdum* of any philosophical view of science that artificially restricts values. Philosophy of science should be based on an investigation of values and practices of science similar to those carried out in the preceding chapter. It is time to see the picture that emerges from this investigation.

## 3.1 Properties of Values

> When scientists must choose between competing theories, two men fully committed to the same list of criteria for choice may nevertheless reach different conclusions. Perhaps they interpret simplicity differently or have different convictions about the range of fields within which the consistency criterion must be met. Or perhaps they agree about these matters but differ about the relative weights to be accorded to these or to other criteria when several are deployed together. With respect to divergences of this sort, no set of choice criteria yet proposed is of any use. One can explain, as the historian characteristically does, why particular men made particular choices at particular times.
>
> Thomas S. Kuhn [1977a, 324]

There are a number of philosophically important aspects and properties of values and how they work.

**Richness.** There is one thing that should be clear from the case studies: There is no one and uniform science to talk about. Values (and practices) of science differ so much from one research tradition to another that it would be misleading to bundle all science under one banner when discussing various philosophical problems about science. Values and practices are so rich and varied. This statement may sound platitudinous or obvious to you, but it nevertheless needs to be stated as its neglect is the root of a number of philosophical problems about science. (This topic will be a running theme in the rest of this text.)

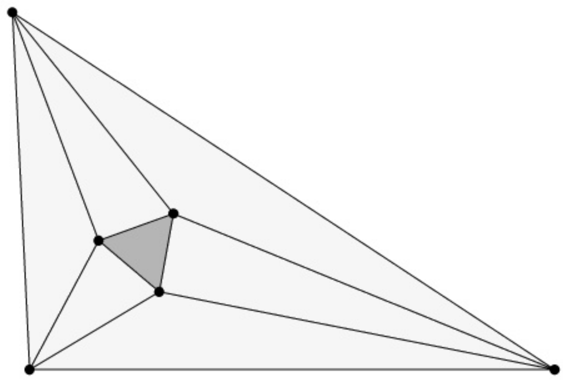


Figure 3.1: Morley's trisector theorem states that in *any* triangle the intersection points of adjacent angle trisectors form an equilateral triangle. [Image adapted from http://commons.wikimedia.org/wiki/File:Morley_triangle.png by Dbenbenn licensed under Creative Commons Attribution-Share Alike 3.0 Unported.]

**Concrete and general values** The most essential thing to note about values is the difference between the general version of values put forward *without* a particular case in mind, and the concrete or specific values as it applies to a case.

When we talk of values, we can talk in the ideal sense, without thinking of a specific case. When one says "scientists prefer simple theories" they are only expressing a rough idea, an abstract ideal. The way an ideal is realized in particular theories can be quite different from each other. For example, what is found to be innovative or fruitful in different fields can have very little common with each other.

The process of recognizing various values in a scientific activity can be quite complex. The relationship between a general value and a particular one can be tenuous or not easily recognizable. It might be even controversial whether that particular case is an instance of a general value. This is why it is important to carry out historical and sociological studies to complement the philosophy of science. These studies can serve as an input to determine values and their relationship. They can also lead to different insights about general values.

The relation of a general value to a specific one can be clear as the case studies I presented in chapter 2 exemplify. For example, there is no puzzle as to why Dirac used the delta function: the simplicity of using $\delta$ is clear to everyone. But which values pertain to a case is not always that clear. I once listened to a talk by a logician in which he mentioned Morley's trisector theorem (see figure 3.1) as an example of surprising but not beautiful piece of geometry. This claim itself was a surprise to me since I always thought Morley's theorem to be beautiful (as well as surprising).

Here is a list of main points about general/concrete values of different theories:

(1) Concrete values of theories can be vastly different than one another, and they generally are.

(2) Different theories can have concrete values that are quite different which nevertheless belong to same general value. For example, what is considered parsimonious in life on

Mars issue has nothing to do with parsimony in ToM issue though both involve the same general value of parsimony.

(3) What is important in theory appraisal is not general values but rather concrete values related to that theory.

(4) The palette of concrete values is so rich and colorful that it is impossible to neatly group them under general values.

Let me elaborate the third point with an analogy. If a hard object hits my face and hurts me, the cause of my pain is not a general concept of hardness, but rather the hardness of that particular object hitting me. Similarly, the concrete values of a particular theory are what is important in its evaluation, not some abstract general value. For example a theory that is thrown away because it has not been useful is not rejected because there is a general value of usefulness, but rather, that theory in that context is found to be not useful. *Efficacious values are concrete ones.* There is a tendency in philosophy of science to blur the lines between two types of values, and even I am guilty of that sometimes. But we should not forget that general values lack punch and what does the heavy lifting are concrete values.

I once read an article by a philosopher[26] which claimed something to the effect that the value-centric approach in philosophy of science is useless because it is highly unlikely that anyone will make a complete list of general values and no one seems to agree on any partial list out there. But this confuses the role general values with concrete values. General values are of no consequence in science — they are only interesting for philosophers of science for some cataloging purposes. What matters are concrete values. And there is no point in cataloging them as they are unique to a theory. Confusing the potency of the general and the concrete values results in bad philosophy such as confirmation theory (see section 5.5).

**Degrees and types.** As Bas C. van Fraassen [1980, 9] writes, "If belief comes in degrees, so does acceptance." The relation of scientists to theories is not in binary. Furthermore, a scientist need not accept a theory to work on it. He might be simply curious about it — trying to see the value of it. It can also be the case that he believes that the theory has a promising future — it does not have a lot of value now, but it has prospects. "There is a broad spectrum of cognitive stances which scientists take towards theories including accepting, rejecting, pursuing, entertaining, etc." [Hacking, 1983, 15] These stances can involve values as well.

There is a spectrum of how much value scientists attribute to a theory. All values have comparative forms: less/more fruitful, has wider/narrower scope, better/worse experimentally supported, has more/less (important) novel predictions, and so on. We have seen these comparisons at work over and over again.

Another component is about scientists' beliefs about the future values of a theory. For example, one need not think that a theory is fruitful (applicable, has wide scope, etc.) to use it. It is possible to use a theory because of a belief that it will eventually turn out to be fruitful (applicable, has wide scope). Scientists consider not only values present, but also those expected.

A couple of quick examples:

[26]Unfortunately, I cannot find the article and do not remember the reference.

(1) We have seen in section 2.11 that two inventors of the quark model, Gell-Mann and Zweig, had different types and degrees of commitment to it.

(2) Different physicists judge the prospects of string theory quite differently — some believe that the only theory with the promise of unifying physics is string theory; while others do not see this possible virtue of it.

The role of a value in science is complicated by not only by different shades of a value, but also by the present and future assessments of that value.

**Factors.** There are a number of factors which are effective in shaping the values:

- Social and cultural factors. These include the social and cultural environment one develops and makes science in, as well as the political, ethical, and religious factors.
- Biological factors. I classify our sensory capabilities and cognitive abilities in this heading as well as all other biological factors you can think of.
- Environmental factors. These include both the role of the environment that shapes us as well as the nature under study.
- Historical factors. Historical factors can be history at large, history of a (group of) scientist(s), or the history of a scientific discipline.

I only put forward this classification as a rough guide. It does not mean that different class of factors are independent. For example, environmental factors have a say in our cognitive abilities. It is also important to note that any of these factors could come apart in different circumstances. For example, different cultural practices can favour different outcomes. Nevertheless, I see more similarities in a class of factors than with others, so I will make use of this classification.

I claim that all these four class of factors are important in shaping values. To begin with, the role of history is as clear as day. The past practice of science influences the current practice through scientific and pedagogical works. If an upcoming scientist is continuously exposed to a particular value in such works, no doubt he will be disposed to recognize and favour that value in his own work.

Let me explain the other three factors by giving a few examples.

- **Simplicity** To begin with simplicity is species dependent. For example, we should not expect that an alien creature with a different ontology, conceptualization powers, etc. share our views of what is simple. Our cognitive capabilities are the main factor in what we find simple. For example, we use mnemonics to help with information retention. But if we had different memory, we might find some of these mnemonics to be more cumbersome then the information itself. Simplicity has a social side as well. Language is the main social ingredient of (common-sense and scientific) simplicity. Scientists learn scientific language and common practices socially.
- **Empirical Adequacy** We observe, measure, and experiment with nature. So clearly environment is a factor. No doubt our sensory capabilities have a strong say on what we observe. But it takes years of education and being a part of scientific community to

learn how to carry out and report these experiments. So the social comes into play as well.

- **Parsimony** According to this value, as we perceive it, environment presents itself to us as simple. The role of environment is clear. Again, the simplicity we attribute has biological and social factors as above.

- **Beauty** Social influences are clear in our non-scientific conception of beauty and scientific communities also have their idiosyncrasies about beauty. Lately, there is considerable evidence that points to a biological basis of beauty, see [Heinric, 2013] for example. What we find beautiful results from a combination of all types of factors.

Although this list is selective and each item features just a cursory examination, we can conclude that our values are collectively determined by the socio-cultural, biological, environmental, and historical factors.

It is important to note that these factors pertain to concrete values, not general. Or rather, these factors act on the community of scientists associated with that concrete value.

It is possible that two communities of scientists have a similar value but the reason why one community has that value can be different than the other. For example, "avoiding Darwinism" is a value which can be seen in very different communities in the last century, for a variety of reasons.

Since there is a many to many relation between the set of factors and the values they associate with, only through a detailed investigation of a token value one can determine the exact involvement of these factors.

**Categories.** It is possible to group values related to theories and models into four categories:

- Pragmatic values
- Theoretical values
- Epistemic values
- Societal values

For me, *epistemic values*[27] are those that connect a theory to the world such as predictive accuracy, empirical adequacy, parsimony, and so on. *Theoretical values* on the other are those related to the formulation and structure of the theory: internal coherence, inter-theoretical consistency, theoretical simplicity, and so on. *Pragmatic values* include fruitfulness, usefulness, pedagogical/psychological factors, etc. *Societal values*[28] are those values which are the most influenced by social factors (like politics).

Having made this categorization of values, I am immediately backtracking from it: It does not work. There are three reasons why:

[27] I am not following the classifications given by previous generations of philosophers who used "epistemic values" to cover what I call "theoretical values" as well.

[28] I reserve the word "social" for the factor as used above and "societal" for the value category.

(1) As explained above, no value is determined by a single type of factor. Far from it, all factors have an influence on all values. This somewhat implies that the candidate categorization I gave is just for show. Deep down, they are all made from the same cloth.

(2) Consider the list of values given in section 1.2. If you go through the list one by one, you will see that a lot of them are impossible to neatly put into one of the four categories.

(3) There are good reasons why different categories melt into each other. For example, theory-ladenness (see section 3.4) implies that we cannot separate epistemic and theoretical values. There are numerous reasons why pragmatic overlaps with epistemic only one of which is mentioned at the end of section 2.6. There is an a chicken and egg relationship between pragmatic and societal.

In the end, all categories collapse. Nevertheless, I find it useful to name the values highly influenced by social factors as *societal values*, especially since their role in science is controversial as I will discuss in section 3.5. Other than that, I will stay away from this fake categorization.

## 3.2 Value Analysis

> I don't think Einstein, or anyone else in 1905, realized how simple and elegant the new theory of relativity was.
>
> Stephen Hawking [1993, 37]

In most scientific endeavors, different values come into play. I have given a long list of possible values in section 1.2. But, of course, that list is only partial; and more importantly, in real cases of scientific judgements, a smaller set of values comes into play. Moreover, concrete values unique to work at hand can appear that has nothing to do with any general values mentioned. I have seen two approaches to investigate the values pertinent to a case.

The first approach is to *assume* what the values are and how they work and then try to find these preconceived ideas in case studies. This "analysis" tries to transform case studies into confirming cases of the prejudged ideas about values. I think that this is a misleading approach and I have much to say about it in chapter 5. At best, this attitude leads to a very impoverished picture of science.

**Value analysis.** The better approach and the one I endorse is to make as few assumptions about values as possible and try to find them out from science itself. This forms the basis of what I call *value analysis*. Instead of stamping assumed ideas about values on science, you rather read them from science itself. Only after a value analysis of science you decide the values in play.

This is not a perfect approach either because three problems arise. (1) It is impossible to get rid of all preconceived/theoretical ideas about values. (2) It is prone to sampling bias: you will only learn about the values dominant in the case studies you have chosen. (3) There is a theoretical bias connecting the first two problems as I explain below. These problems are all relevant to the case studies I presented in chapter 2.

The way out of the first problem is naturally through discussion and (self-)criticism. The ideal is to reach a minimal common ground. My own view about the cases I presented is that I have avoided any unreasonable assumptions about values (but see the caveat below).

The solution to the second problem is to choose diverse and numerous case studies. As Thomas Nickles [2001, 88] puts it, "Just as one can appeal to the Bible to 'prove' almost anything, so one can find a historical case to 'establish' or 'refute' practically any methodological claim." That is why I am not satisfied by one or two case studies in this text. But of course, this is not a problem that can be remedied in a single text.

There are (to my knowledge) three kinds of sampling bias in my case studies of chapter 2: (a) Clearly these studies reflect my own interests and familiarities. In particular, they are all examples from twentieth or current century science. (b) I tried to choose examples which can be tackled in short sections. No doubt, not all science is suitable for such a bite-sized analysis. (c) I also tried to come up with examples in which evidential values are not a big factor. But as computer programmers like to say, *this is not a bug, it is a feature*, that is, it is deliberate on my part. One of my aims was to show the wide range of values, not repetitively showcasing evidential values.

The third problem of theoretical bias is the relation of generalizations and individual case studies. Assume that (for example) fruitfulness is an important value in most case studies at hand. One may naturally conclude from this situation that fruitfulness is an important value in science. This would dispose further analyses of case studies to attribute more importance to fruitfulness than that value deserves. So sampling bias leads to theoretical bias and theoretical bias leads to sampling bias. I do not think that this is a vicious circle. The more diverse examples we investigate and the more we criticize the cases at hand, the better our theoretical understanding will get. In turn, this theoretical improvement will lead to a better handling of case studies. So on and so forth.

For example, I believe that one of the most important values is theoretical simplicity. It might be true that this leads me to see more of it in the case studies I look at, and this enforces my views about simplicity. Am I doomed? Do I have to live my days left on this world in an inescapable simplicity prison? Not really. I try to be critical of my opinions about simplicity and look forward to other people's criticisms about it. Our understanding of simplicity improves and I might revise my views in the future. Even if I do not, others can; and if this simplicity-importance is a disease, it does not have to turn into an unrecoverable epidemic.

To sum up, even though there are some limitations involved, it is possible to carry out value analysis of science. I will argue in section 5.2 that science should aim to be unbiased and can eliminate biases in time. The same goes the value analysis of science as well. I acknowledged different types of bias above. It is possible that there are more but we can find them out and be more careful about them in the future. Nevertheless, value analysis is much better than the alternatives, that is, simply assuming the values.

There are some parts of science that are more susceptible to value analysis. As I have shown in my collection of case studies there are examples in which values are apparent. Scientific controversies are especially suited for value analysis since a clash of competing theories tends to make the different values of the theories stand out. But not all values are as conspicuous as those in controversies.

I have talked about the *narrow sense of value analysis* — to find about about particular values of a theory (ideally) without any biases and especially without assuming the values to be found. We can build on this basis by comparing the values of theories, comparing

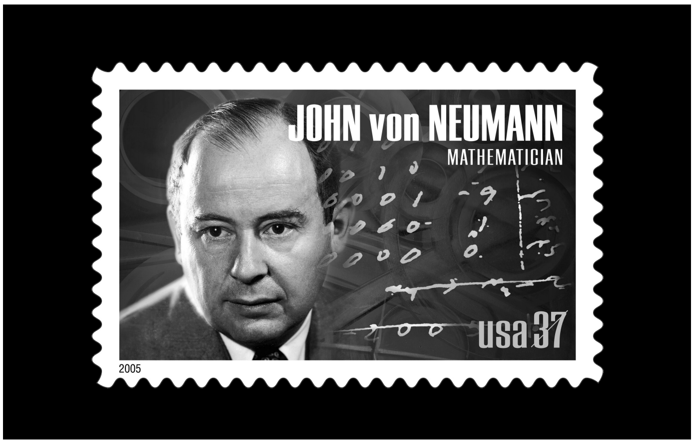



Figure 3.2: A stamp by U.S. Postal Service issued in 2005 featuring John von Neumann (1903–1957).

the particular values of a theory to general values, finding similarities and dissimilarities in how values are employed, finding out new general values, seeing the relationship of values to scientific practices, how changes in theories affect values, and so on. This project which strives to understand values and puts the intuition gained to good use on philosophy of science problems can be called the *wide sense of value analysis*. From now on, I will use "value analysis" ambiguously, referring both the narrow and wide senses.

**Background Values.** A theory may have some values which are/were not considered to be primary in its appraisal. These present but ineffective values are traps for value analysis. Just because one observes a value of a theory does not mean that that value is/was important in the (non-)adoption of the theory. This can happen in five ways:

(1) It is possible that a theory has a particular value that plays no or limited role in the appraisal of the theory. For example, the life on Mars issue (section 2.6) has various religious or cultural consequences that do not factor in the scientific discussion.

(2) Two competing theories $T_1$ and $T_2$ can be well-matched about a relevant value $V$ so that $V$ cannot be used to judge between them. For example, if $T_1$ and $T_2$ are equally simple, then simplicity cannot be used to choose one of the theories over the other. This value $V$ can be any value but the one philosophers have found the most interesting is evidence. This situation is called *the underdetermination of theory by evidence* which I will look at in section 5.4.

(3) It can be the case that the initial formulation of the theory which did not exhibit a value can be changed to a formulation which brings the value to the light. For example, the lack of mathematical rigour in Dirac's formulation is not seen in von Neumann's formulation of quantum mechanics.

(4) Further scientific, environmental, or social developments can bring about new values of a theory. For example, a not very useful theory can be turn out to be quite useful in time. In these cases it is important not to mistake these later values for those at the time time of theory adaptation or development. These further values can strengthen or weaken a theory.

Scientists sometimes express[29] the sentiment that science should be done for its own sake,

[29] See for example [Shockley, 1950, vii] or David Hilbert's 1930 radio address available from http://www.maa.org/

not necessarily for its applications and technology since it is impossible to predict what these would be. There are successful applications of various theories which had no such practical uses initially.

(5) A value can tag along other values, that is, a value may be dependent on others and it might be indirectly involved. To put in another way, some values somewhat supervene on others. For example, (mathematical) simplicity most of the time results in dynamism[30] as well. [Frank, 1962, 352] writes that "It seems that mathematically simple theories are also dynamic, are fit to be generalized into theories that cover a wide range of facts."

It might be the case that the values are so tangled up that it is hard to decide which values are primary and which tag along. Jim Holt writes in The New Yorker magazine[31] that

> The gold standard for beauty in physics is Albert Einstein's theory of general relativity. What makes it beautiful? First, there is its simplicity. In a single equation, it explains the force of gravity as a curving in the geometry of space-time caused by the presence of mass: mass tells space-time how to curve, space-time tells mass how to move. Then, there is its surprise: who would have imagined that this whole theory would flow from the natural assumption that all frames of reference are equal, that the laws of physics should not change when you hop on a merry-go-round? Finally, there is its aura of inevitability. Nothing about it can be modified without destroying its logical structure.

In this analysis, Holt reduces the beauty of general relativity to its simplicity, novelty, fruitfulness, and internal coherence. But one might as well have taken beauty as an independent value or found even more values which it depends. There is a lot of wiggle room here. The acceptance of general relativity is such a complex case that there might be different plausible analyses.

Kuhn [1970, 185] claims that the values are often imprecise so that different scientists might make different conclusions about a particular virtue when assessing a theory: "judgements of simplicity, consistency, plausibility, and so on often vary greatly from individual to individual." He later extended his views in *Objectivity, Value Judgment, and Theory Choice* where he writes:

> I am suggesting, of course, that the criteria of [theory choice] function not as rules, which determine choice, but as values, which influence it. Two men deeply committed to the same values may nevertheless, in particular situations, make different choices as, in fact, they do. But that difference in outcome ought not to suggest that the values scientists share are less than critically important either to their decisions or to the development of the enterprise in which they participate. Values like accuracy, consistency, and scope may prove ambiguous in application, both individually and collectively; they may, that is, be an insufficient basis for a shared algorithm of choice. [Kuhn, 1977a, 331]

See also McMullin [1982, 16–17].

The fact that values do not function as rules is only one part of the story. The other part is about the whole set of values involved and how complicated their relations can be. It might

publications/periodicals/convergence/david-hilberts-radio-address. Retrieved on December 15th, 2016

[30] Recall that a theory is dynamic if it is "more fit to expand into unknown territory." [Frank, 1962, 352] Such a theory can accommodate future developments of science or itself can be included in a future theory which is more general.

[31] October 2, 2006 issue available from http://www.newyorker.com/magazine/2006/10/02/unstrung-2. Accessed on January 17, 2016.

not always possible to navigate the labyrinth of values and decide which values are primary and which depend on others.

Even though values are not rules and even though we can not always be sure of the relevant concrete values or their relationship, there is good reason to analyse values. As I have shown in chapter 2, most of the time, the values can be untangled. Even if we cannot negotiate all the intricacies of the values of a theory, we would still learn a lot in the process.

**Changing Values** Value analysis should take into account that the set of values associated with a theory can change in time. I will give a few examples how.

(1) We have seen in section 2.6 that two values — simplicity and parsimony — are confronting each other in the life on Mars issue. For now, parsimony is winning and the majority of scientists deny life on Mars. But it is possible that further evidence or analysis of currently available evidence make the life hypothesis less parsimonious. In that case simplicity might overwhelm and the tide can change to the life hypothesis.

(2) Recall the Piltdown Man fraud discussed in section 2.10. Between the years 1912 when the bones were found and the year 1953 when the fluorine tests on the bones were carried out, the prevalent values were simplicity, theoretical consistency, nationalism, and aligning with cultural influences. But when the fluorine test results came in, the evidence weighed itself so strongly that there was not any more discussion about the authenticity of the skull. See figure 3.3.

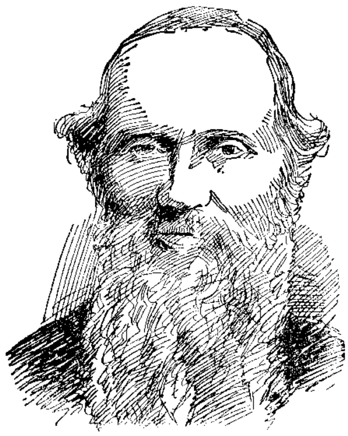
(3) John Worrall discusses the wave theory of light in a couple of papers [Worrall, 1976, 1989]. Young put forward his theory in the early 1800s though the scientific community only began to accept the wave theory of light after Fresnel's contributions twenty years after. Why is this delay? What did Fresnel's theory had that Young's lacked? Worrall convincingly dismisses the widely cited view that the novel prediction of the Fresnel's theory, namely the Arago-Poisson spot, is the main or contributing reason for the acceptance of the new theory. Rather, what made Fresnel's work valuable is its mathematical novelty and the improved method of measuring diffraction fringes. These two virtues together with the ad-hocness of Young's theory made the difference.

(4) In the late nineteenth century, geologists and biologists had suggested that the Earth should be at least 300 million years old. But Lord Kelvin (William Thompson, figure 3.4 ) used Fourier's heat conduction theory to calculate the age of the Earth to be at most 100 million years. This caused a controversy for some years until the discrepancy was resolved by the discovery of radioactivity in the early twentieth century. Interestingly, "Fourier himself had obtained a similar result, but it seemed to have no significance at a time when geological periods were measured in thousands rather than millions of years; only after Lyell encouraged geologists to 'think old' did 100 millions years seem like a short time." [Brush, 1996, 5]. This shows that the exact same theory or argument can either be completely neglected or become the center of attention depending on the other theories of the time.

The examples can be multiplied, but the above ones are enough to show that values associated with a theory can change in time. Values can put on or lose weight, disappear, do time and get out of prison, go in hiding, become famous, and so on. Value analysis should be sensitive to the possibility of such change.

Figure 3.3: Group portrait painted by John Cooke in 1915 of the Piltdown skull being examined by some of the leading British archaeologists and anatomists. Piltdown Man, who had an ape-like jaw and and a human-like upper skull, was defended by the British for theoretical and cultural reasons until it was established by fluorine content of the bones that the jaw and skull fragments were fraudulently brought together. Back row (from left): F. O. Barlow, Grafton Elliot Smith, Charles Dawson, Arthur Smith Woodward. Front row: A. S. Underwood, Arthur Keith, William Plane Pycraft, and Ray Lankester. [Image in the public domain from http://commons.wikimedia.org.]

**Kel'vin, Lord,** formerly Sir William **Thomson,** world-famed physicist and inventor, was born at Belfast, Ireland, in 1824. At the age of eight he removed to Glasgow with his father, a professor in the university. Kelvin graduated at St. Peter's, Cambridge, in 1845, being second wrangler and Smith's prizeman. Beginning in 1846, he filled the chair of physics at Glasgow University for 50 years. In 1866 he was knighted for his signal achievement in solving the scientific problems connected with the first Atlantic cable. In 1892 he was raised to the peerage. His contributions to physics are so numerous and cover so wide a range that it will be impossible to do more than allude to a few of the more remarkable ones. His contributions to science are contained in *Papers on Electrostatics and Magnetism*, *Mathematical and Physical Papers*, *Popular Lectures*, *Molecular Dynamics* (a course of lectures on *Light*) and his masterly treatise on dynamics, written in conjunction with Professor Tait of Edinburgh. His inventions and improvements of electrical measuring instruments, navigator's compass and deep-sea sounding apparatus have been of exceeding value to commerce as well as to science. Few men have ever lived who combine the mathematical ability and the experimental acumen which united in Lord Kelvin. Honors of all kinds from the presidency of the Royal Society down have been conferred upon him, including burial in Westminster Abbey. His death occurred on Dec. 1, 1907.



LORD KELVIN

Figure 3.4: William Thomson (1824–1907) was a distinguished British physicist and engineer who in 1860s used cooling rate of the Earth to estimate its age. Since the heating effect of radioactivity was unknown at the time, his numbers were very modest and this caused a controversy with evolutionary theorists. See [Hallam, 1989, chapter 5] for details. [Entry from *The New Student's Reference Work*, 1914, reformatted to one column.]

## 3.3 Inheriting Values

— *I worry that all of my wisdom is derived from bad analogies.*
— Ratbert, sometimes a good wine has to age before it is perfect.
— *So … I'll get smarter over time?*
— To the extent that you are like a grape.

Scott Adams, *Dilbert*, http://dilbert.com/strip/2006-10-26

**Analogies** In appendix D, I mention how mathematical tables can be interpreted differently. They may be used to refer to different things as seen fit. But of course, this is true of all mathematics. When you write down mathematics, it is just mathematics. Scientists may use that mathematics to represent different phenomena or interpret it as a part of some science. These representations or interpretations are dynamic in the sense that whether/how/what they represent can change in time. This shows the pragmatic approach of scientists to mathematical things, putting usefulness and simplicity at the forefront.

This flexibility is not unique to mathematics. Sometimes it is possible to use analogies to apply concepts or knowledge of one domain to another. Here is an example from chess:

> There are various ways of creating a plan [in chess], but one of the most common is by analogy. A knowledge of the plans available in similar positions may suggest one which can be transferred to situation on the board. Often the key factor is the pawn-structure; if this is similar to or the same as a known position, it may well be possible to adapt a plan. [Nunn, 2011, 51]

More interestingly for us, analogies are ubiquitous in the history of science. I will only investigate the role of values in analogies. For a general account see [Bailer-Jones, 2002]. Analogies are used in science relate a model/theory/system unknown to a model/theory/system more well-known. Bailer-Jones [2002, 113] writes that "an analogy can be analyzed in terms of similarity, similarities of relationships (e.g. encountering interference in water waves *and* in light) and similarities of object attributes (e.g. oxygen and helium being gaseous at room temperature)." An analogy can have different purposes:

- It is a tool of discovery. To explore an unknown terrain, one uses tools successfully employed in the past.
- It can save time and energy by using readily available methods; it facilitates intellectual economy.
- An analogy may explain by improving understanding.
- An analogy makes unfamiliar familiar by connecting an alien landscape to a known one.
- It can serve as a pedagogical tool.

See [Bailer-Jones, 2002] for an elucidation of all but the last item of the above list.

The interesting point about analogies related to our topic is that they are carriers of values. If a target model X is analogous to a source model Y and Y has a value V, then that value can rub on to X. Just a quick sample of examples:

- Simplicity. Most of the time analogies lead to simplicity as it involves a simpler model.
- Visualisibility. A geometrical or solid model can aid in visualisibility.
- Incentives. If you connect your model to a popular one you might get published easier.

Analogies can be carriers of any value you can think of.

What about familiarity? Since analogies in science are always between a well-known source model and an unfamiliar target model, familiarity is always a benefit the target model gets. But what other values are carried over depends on the particular analogy and we cannot know without examining the analogy in question. Like most other things in philosophy of science, there is no quick formula that could a priori give a list of transferred values.

Most of the time, the values transmit from the source to target model. But the transmission of virtues can work in the opposite direction as well. For example, one might find a class of new interesting problems or mathematical ideas when working with the target and see if they make sense in the source as well. In this way, for example, fruitfulness may spread over to Y. But this direction of virtue transmission is less common.

An analogy can become so engraved in our minds that we can forget that there is an analogy. For example, we do not think of "electric current" as an analogy any more, though it started its life as one.

Analogies are a salient way of transferring values and they deservedly need all the attention they can get from philosophers of science. Now, let us turn our attention to a more dubious way of value transmission.

**Inclusion** The following paragraph is from an introductory text on string theory available online:

> String theory, if a true unified theory, shouldn't just predict new phenomena, but also phenomena we are already aware of. The Standard Model of particle physics contains all of the forces except for gravity, and an entire zoo of particles, most of which have been experimentally verified. In effect, String theory should be able to be reduced to the Standard model in the low energy limit. In other words, the Standard model should be, in a sense, predicted by String theory. In other words, *the consistency of string Theory with other theories with provided experimental evidence is another route in which string theory might be tested, though indirectly.* [Svesko, 2013, 27, my emphasis]

The author is not alone in voicing such sentiments. Lately, there is a debate about the usefulness of string theory (see figure 3.5) which has spilt over to internet forums and blogs and the above ideas are met frequently in the blogosphere as well. Actually, there are two different ideas voiced here:

(1) If the standard model can be derived from the string theory, then this somewhat gives credence to the string theory.

(2) Assuming that the string theory implies the standard model, the well-confirmed status of the standard model is transferred to the string theory.

We can also consider these sentences in a more general and abstract way:

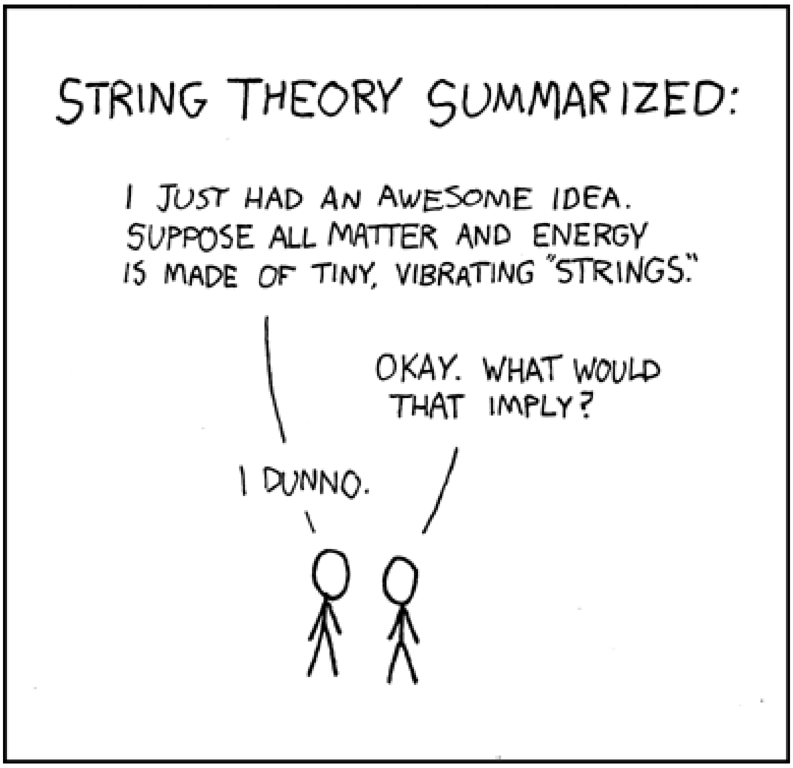



Figure 3.5: A comic by Randall Munroe on string theory. [Image from http://xkcd.com/171/ licensed under Creative Commons Attribution-NonCommercial 2.5 License.]

(3) If a theory/model A implies B where B is a well-known, successful, etc. theory/model, then this gives some credence to A.

(4) If a theory/model A implies B where B is a well-known, successful, etc. theory/model with a value X, then A inherits some X as well.

The items (1) and (3) are not exactly the topic I want to deal here but let me get them out of the way. When Einstein's general relativity implied the precession of Mercury (in a natural way), this was considered a big plus for general relativity. On the other hand, when Hendrik Lorentz's (see figure 3.6) ether theory (circa 1910) implied constancy of the speed of light (in an ad-hoc way), nobody cared. This goes on to show that (3) fails as a general rule. The credence transfer can only be investigated case by case.

So if the string theorists get the standard model, how important would that be for string theory? I can only guess: If the standard model arises from string theory in a natural and non-ad-hoc way, then it would be a huge success. But if it gives the feeling that you are getting the standard model by tweaking some parameters the right way just to get the standard model, it will not be very impressive. Let us move on to (2) and (4).

In section 5.5, I have a few things to say about confirmation theory, and all bad. But this does not mean that we should ignore lessons learned from it. One such lesson is: If a theory A includes a sub-theory B and a piece of evidence E confirms B then it is *not* necessarily true that E confirms A as well. Just for a trivial example, let A be B&¬E. This generalizes to all values:

Figure 3.6: Albert Einstein and Hendrik Lorentz photographed by Paul Ehrenfest in 1921. [Image in the public domain from https://en.wikipedia.org/wiki/File:Einstein_en_Lorentz.jpg]

(5) If A includes B and B has some value X then it is *not* necessarily true that A has X as well.

For example: a very complex theory can have a simple sub-theory; an unfamiliar theory can have a familiar part; a confusing theory can have clear parts; a non-commonsensical theory can be an extension of a commonsensical one; a theory that contradicts other theories can have consistent sub-theories; and so on.

We can conclude that if A is a new theory that includes/generalizes/extends an old theory B then there may be no transference of a particular X from B to A. Of course, A can have that value independently, but that is another matter. The point is that (4) fails as a general principle.

In particular, for string theory the claim goes as follows: (a) String theory gives the standard model. (b) There is so much evidential support for the standard model. Hence (c) string theory gets some (indirect/partial) evidential support transitively. There is no substance to this claim because inclusion by itself is not a sufficient reason for value transference.

Moreover, the argument (a), (b), (c) in the above paragraph is most of the time offered as an answer to those who challenge string theory on evidential grounds. This would not work because string theory is yet to give the standard model. The challengers are not saying that string theory would not have any evidential merit in the *future* event of a successful unification, but rather, they are saying that string theory has no any evidential merit *now*.

To sum up, while analogies transfer some values, inclusion (by itself) does not. In both cases, only a philosophical investigation can show us which values are shared or transferred.

## 3.4 Theory, Evidence, and Discovery

> One Ring to rule them all, One Ring to find them,
> One Ring to bring them all, and in the darkness bind them,
> In the Land of Mordor where the Shadows lie.
>
> J. R. R. Tolkien, *The Lord of the Rings*

In section 1.2 I introduced the *value myth of positivism*. The first assumption of the myth is that evidence is the one and only value to rule them all. Logical positivists thought that there is a clear and complete distinction between theory and evidence. The claim is that observation is separate from theory in the sense that evidence bears on theory but not the other way around: evidence stands on its own and weighs on theory. Similarly, there is a separation between the contexts of discovery and justification. Scientists could come up with theories in any way but regardless of how they do that, theories are autonomously tested or confirmed. According to this view, discovery and validation of theories are distinct processes. (See figure 3.7.) This one-dimensional approach was replaced in 1950s and 1960s by philosophers such as Norwood Russell Hanson, Paul Feyerabend, and Thomas Kuhn building on the philosophy of Pierre Duhem.

> [D]iscovering a new sort of phenomenon is necessarily a complex process which involves recognizing both *that* something is and *what* it is. Observation and conceptualization, fact and assimilation of fact to theory, are inseparably linked

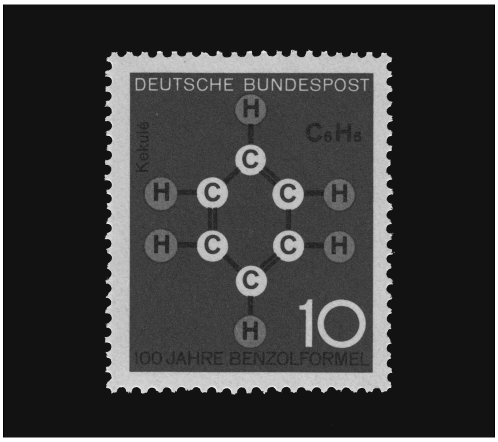



Figure 3.7: A German stamp issued in 1964 for the hundredth anniversary of August Kekulé's discovery of the structure of benzene. According to the story, Kekulé found the structure after dreaming of a snake biting its own tail. Logical positivists recite this story to show that one can discover new hypothesis in the most creative ways. But justification, on the other hand, is an autonomous and logical process that separates itself from discovery.

> in the discovery of scientific novelty. Inevitably, that process extends over time and may often involve a number of people. [Kuhn, 1977b, 171]

Today a much highly relationship between theory and evidence is acknowledged and the notion of pure observation is rejected by most philosophers of science. The aim of this section is to demonstrate some of this complex relationship.

**The selection role of discovery.** Thomas Nickles [2001] considers an under-appreciated role of discovery in science:

> Consider the following general argument from economy of research that coupling [of discovery and justification] is necessary to achieve the goals of science, and hence that discovery is an essential topic for epistemology. The central problem of methodology is to show that the methods advocated have a reasonable chance of achieving the stated goals of the enterprise, that they are better than blind luck. The problem is to show how to achieve infinite aspiration ("Find the one true theory in an infinite domain of possibilities!") by finite means. Suppose that the goal of research is to find true laws, theories, models, and/or explanations. Of the infinite number of possible laws, theories, etc. possible for any scientific domain, scientists will, over time, actually formulate and consider only a finite set of candidates; or, at least, infinite subsets of the points in search space will go unnoticed. (For one thing, as the histories of deep conceptual change have taught us, we cannot now canvass possibilities that will only become available in future eras.) And these are precisely the law or theory candidates furnished by discovery procedures of whatever kind. Whatever its character, the discovery process filters out these few from the limitless set of potential laws, theories, or explanations. These are in turn poured through a second filter consisting of empirical testing and checks for compatibility with theoretical and methodological constraints. Thus research amounts to a two-stage filter or selections process. Unless a true (or sufficiently reliable) candidate is selected at the discovery stage (or developed during the transformation to "final discovery"), it has no chance at all of being selected at the second stage. In this sense at least, the discovery process is epistemically relevant. There must be some degree of coupling between

> the modes of generating theories and criteria of epistemic appraisal. Furthermore, since consequential testing obviously cannot govern the first-stage filter (since the candidates must already be selected before testing can commence), the economy argument establishes the necessity of a generative component of methodology, one that epistemically informs the initial selection/construction process. [Nickles, 2001, 91–92]

Since scientists always have a finite number of theories at hand in a sea of infinite possibilities, how they come up with those have a strong bearing on what theory they will end up with. The history and methods of the field, interests of its practitioners, inter-theoretical relations all affect what Nickles calls the first stage.

But the role of discovery does not end with the first stage and discovery influences justification in a stronger sense as we shall see.

**Theory-ladenness.** *Oxford Dictionary of Philosophy* gives the following definition:

> A statement is theory-laden if its terms only make sense in the light of a set of theoretical principles. The judgement that an electron has been emitted, for example, would be more heavily theory-laden then the judgement that there was a white flash on the screen. [Blackburn, 2008, 362]

I use "theory-laden" not only with this meaning but also in a more general sense which is that observation/data/evidence are influenced by theoretical presuppositions. One consequence is that theories can render some group of observations more salient than others if the former is theoretically more interesting than the latter.

At the end of section 2.3, there is a long quotation on some geochemical work carried out in 1982–1983 which shows the amount of interpretation that goes into grasping of evidence. The ratio of the various isotopes were not simply measured and accepted, but they were rather interpreted or "corrected" with theoretical considerations in mind. Let us recall the corrections mentioned about the nitrogen isotopes:

> Becker and Pepin's measurement of the nitrogen isotopes trapped in the same meteorite was received a bit more tentatively. The ratio of extracted nitrogen-15 to nitrogen-14, even after a *correction* for nitrogen in the unshocked part of the meteorite, was only +130 per mil in the standard notation of isotopic ratios. The value for nitrogen in the Martian atmosphere is $620 \pm 160$ per mil. To test the Martian origin hypothesis, Becker and Pepin made *another correction* based on the assumption that there was more nitrogen in the minerals that formed the gas-containing glass than found elsewhere. They took as a measure of that excess the relative sizes of the nitrogen-argon ratios in the glass and in the Martian atmosphere. This *second correction* raised the value to +500 per mil, which is within the Martian range. Most listeners viewed the necessity of a second correction as regrettable but took some reassurance from the high nitrogen-argon ratio of the meteorite, which lies between that of Earth and Mars. [Kerr, 1983, 289, emphases added]

The first correction gave only 130 per mil, not enough to be in the Martian range. "Regrettablly," a second correction was needed to raise the ratio within the Martian range. This was as good as any time to stop correcting the data since it was now consistent with the Martian hypothesis. But wait, was not the original idea to test the Martian hypothesis?

Theoretical presuppositions do matter in the interpretation of data. Sometimes even the theory that is to be tested can end up influencing the data supposed to test it. When scientists who support that theory are the ones to test it, the tests and their interpretations naturally reflect their interests. This is one reason why separation between the contexts of discovery and judgement fizzles out. As a case in point, let me briefly mention spontaneous generation.

**Spontaneous generation.** Spontaneous generation says that some living things can originate suddenly by chance from non-living matter independently of any parent [Farley, 1977, 1]. If we consider the modern history of the spontaneous generation controversy to start with Francesco Redi's (see figure 3.8) celebrated experiments, it took two centuries to settle it. But why did the controversy took so long to settle? This is an extremely complex issue with a number of reasons, only two of which I will discuss.

First, the experiments which the opponents of spontaneous generation carried out were accommodated by the proponents as well. For example, demonstrating that no life occurred in boiled and sealed containers that contained broth was not enough to establish the failure of spontaneous generation because the proponents claimed that such a process destroyed the "vegetative force" needed for spontaneous generation. Other experiments that purportedly showed failure were similarly explained away by the proponents for destroying the "elasticity of air". Whatever clever experiments the opponents came up with to show that no spontaneous generation occurs, the proponents came up with with equally clever explanations to assimilate those results with their theory. On the other hand, any experiment that allegedly established spontaneous generation was attributed to experimental error or some other inadequacy by the opponents. The theory-ladenness of these experiments kept the controversy alive for two centuries.

Second, the spontaneous generation theory did *not* claim that life is generated from given suitable materials *all* of the time, but rather, *some* of the time. That means that there is actually no logically conclusive way to show that spontaneous generation theory is false. Any experiment that fails to show spontaneous generation can be chalked up to one of those times that it did not occur. Moreover, complete sterility was very hard to achieve and even the best experimenters had an occasional slip up and saw signs of life in their experimental set-ups.

All these experiments were interpreted very differently by the supporters/detractors and assimilated to their theories. Naturally, each camp interpreted the experiments to fit their theories.

**Invisible evidence.** Sometimes observational data is theory-laden to the extreme, that is, the evidence or data is invisible until the right theory is in place.

Parity violation of weak interactions was found only after theoreticians put the question in front of the experimenters. But it turns out experimenters were observing it all along without recognising it: "Looking back, physicists realized they had seen parity violation before but dismissed it as an anomaly or explained it away. What they had thought was noise — a dirt


Figure 3.8: Statue of Francesco Redi (1626–1697) at The Uffizi Gallery in Florence, Italy. Redi was an Italian biologist and poet who carried out experiments to test spontaneous generation by comparing the occurrence of maggots on meat placed in open and gauze covered jars. [Photo from https://commons.wikimedia.org/wiki/File:Florence_Italy_Statues-in-the-Uffizi-outside-Gallery-01.jpg by Uwe Aranas licensed under Creative Commons Attribution-ShareAlike 3.0 Unported license.]

effect — was the true signal." [Johnson, 2000, 149]

Writing about the discovery of the $\Omega^-$ particle that I mentioned on page 55, Johnson [2000, 218] writes that "Obviously, ... a great deal of interpretation was involved. The omega-minus was not something anyone would likely notice if not already primed to see it." And this "interpretation" and "priming" makes theory and evidence intertwined, far complex than the simple positivist picture.

As I mention on page 57, Norwood Russell Hanson [1963] investigates the discovery of the positron in early 1930s. The story of the discovery of the positron (1932–1933) is complex and interesting but I want to touch upon just a portion of it. It turns out that physicist were seeing tracks of positively charged electrons well before 1932. As Hanson puts it, physicist "certainly encountered electrons, before Anderson's discovery [in 1932], which they described as 'falling back into the source', 'curving the wrong way', 'coming up from the floor', or 'moving backwards'." [Hanson, 1963, 138] But these electron tracks were either neglected or "Whenever seen, such tracks were discounted as 'spurious', or as 'dirt effects'." [Hanson, 1963, 139] Once physicists accepted the new particle, they returned to some of these neglected or puzzling effects and "many 'spurious' phenomena now seemed amenable to possible re-interpretation." [Hanson, 1963, 164]

These examples completely demolish the claim that theory and evidence are separate categories. It might be the case the evidence is just in front of the scientists who do not recognize it until the theory points them at the right direction.

**Conflicting evidence.** The positivist picture of evidence introduced at the beginning of this section also neglects the fact that different evidence can conflict each other or theories.

(1) A collection of data can have accidental properties (see the discussion about *data dredging* on page 123). That is why different studies can have conflicting evidence: one might say that a drug is good for a particular condition, the other might contradict it. That is why systematic reviews that compare all such studies are of paramount importance. But, of course, there has to be sufficiently many studies for a systematic review to make sense.

(2) Since evidence is theory-laden and it can be mediated by theories in different and complex ways, it should not be surprising that we may end up with conflicting evidence in science. For example, consider the age of Earth issue of the late nineteenth century: Geological evidence supported a very old Earth but the physical calculations based on the heat dissipation of Earth supported much younger Earth. [See Brush, 1996, 5–7]

(3) An issue can be so complex that it might not be clear what counts as an evidence for or against unless you bound or simplify the issue. As there is no unique way to achieve this, what is evidence for/against changes depending on the outlook. The cycling helmet issue discussed in section 2.8 is such an example.

(4) Evidence can conflict with other theories. As we have seen in section 2.3, the geochemical evidence for the Martian origins of certain meteorites was shelved for a considerable time because the geophysical theory of the time precluded the possibility of interplanetary travel of such rocks. In such conflicts, the theory or evidence can give way, and in the case of Martian origins, it was geophysical theory that caved in.

When there are conflicting relevant evidence (and theories), one cannot simply gather these under a banner and call it a day. If a theory was to subsume all relevant evidence without

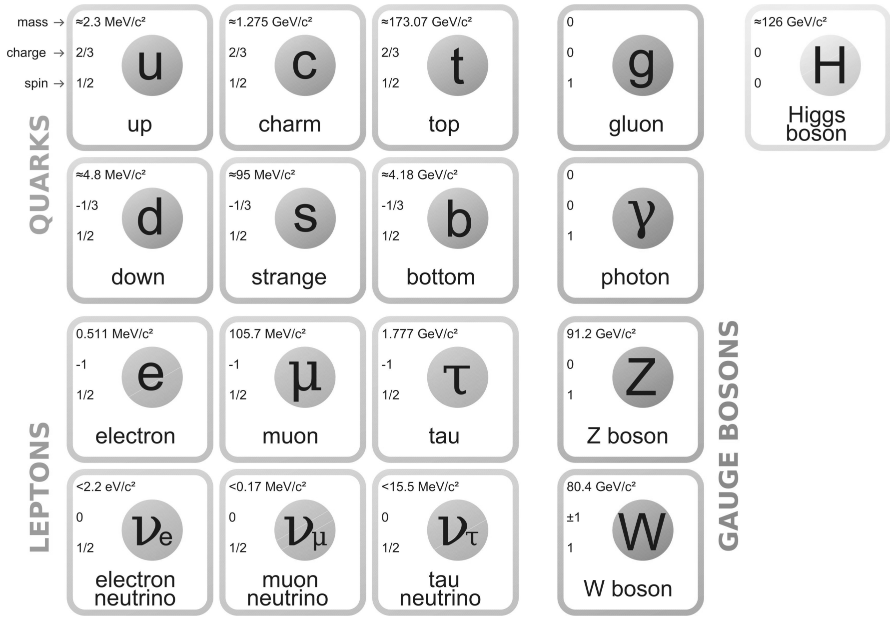



Figure 3.9: The quark model eventually led the way to the standard model of elementary particles shown here. [Image adapted from https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg by MissMJ licensed under Creative Commons Attribution 3.0 Unported.]

interpreting or rejecting some of them, it would end up as a jumbled mess of contradictory bits. The theory *has to* mediate through evidence to resolve the conflicts, though there is no guarantee that this can be achieved. Scientists might shelve some of them as *anomalies*, hoping for eventual resolution.

**Extended building.** "Scientists come up with a theory and then it is tested independently" view also does not take into account what I call *extended building* which is the long meandering path of discovery through evidence and theory.

Consider the case of quarks. The story starts with only three quarks in 1964. Then for the next decade there is a rich interplay between experiments, their interpretation, and theory. This is a very complex interaction and my presentation of it in section 2.11 gives only gives a few headlines as I indicated there. The theory is modified to account for new experiments and other theoretical concerns. On the other hand, the theory suggests new experiments, new interpretations, and so on. In the end you end up with a much richer quark model with six quarks each of which comes in three colors, and also a bunch other additions to the model (see figure 3.9).

Nickles [2001] writes about something similar to my notion of extended building:

> While AI experts fail to appreciate the degree of conceptual reconstruction and refinement of skill typical of scientific work, the same is true even of historical philosophers and sociologists! The discovery process is far more drawn out and structured than most methodologists appreciate. Theory construction, as exhibited in the original papers, is only the first stage of discovery, only the first "round" of innovation. To stop even at "final justification" (empirical confirmation of the new claims) is to leave the task of describing/explaining discovery only half

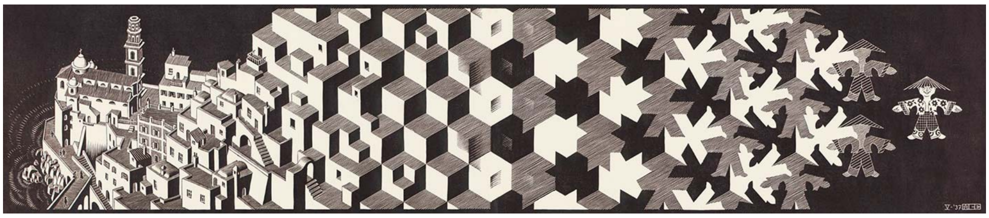


Figure 3.10: *Metamorphosis I* by M. C. Escher. [All M.C. Escher works © 2015 The M.C. Escher Company - the Netherlands. All rights reserved. Used by permission. www.mcescher.com]

> finished. For much or most of the innovation of every major discovery occurs in the successive technical refinements that occur in the decades after acceptance of the initial work into the literature — years after "the discovery" was supposed to have been made. [Nickles, 2001, 90–91]

Nickles considers "technical refinements" that can happen after initial "theory construction". I would subsume such episodes under the heading of "extended building", but there can also be cases of extended building consisting of a long process with extensive or important changes at any time, not only technical refinements. There can be an association or unification of justification, discovery, and construction. More accurately, I think that these separate categories fade away, and what we have is a new process, namely, extended building. As I remarked in section 2.11 the discovery/invention distinction fails to cover episodes such as quarks. Both "discovery" and "invention" suggest a precise time of discovery/invention which does not apply to extended building. There are many steps taken during extended building any of which might vaguely or strongly resemble discovery, invention, or justification; but the whole process cannot be pigeon-holed into one of these categories.

When I think of extended building, I cannot help but recall (analogically) Dutch graphic artist M. C. Escher's (1898–1972) three woodcut prints[32] *Metamorphosis I* (1937), *Metamorphosis II* (1940), and *Metamorphosis III* (1967–1968). These are long narrow prints of different lengths featuring transitions between patterns, tilings, and various objects. These patterns morph and change shape as you move along the print and the interaction of shapes eventually result in an image of a small town. Let us think of the town as the complete theory at the end of the process. The shapes represent the elements of extended building. They somehow interact and influence each other, but you never know how they will change or where they will end up. The shorter of the three, *Metamorphosis I*, has more "technical refinements" along the way. In the longer prints *Metamorphosis II* and *Metamorphosis III*, there are some patterns that look like intermediary theories or discoveries but they change as well. The way I think of the analogy, there are no distinctive "justification" or "discovery" shapes that span all the image — you cannot isolate "justification" or "discovery" blocks in the image. Thus the longer images *II* and *III* represent a case like quarks better than the shorter *I*.

Extended building is not something that happens in isolation. It can be influenced by or influence different theories, technical advancements, and so on. It is possible that extended building can lead to many important results or methods along the way and some of them can function independently and even become important in their own right.

[32] Here I reproduce *Metamorphosis I* as figure 3.10. I strongly suggest that you look up the other two images on the internet yourself.

The gradual change of the neuron doctrine in the twentieth is an example of extended building. The doctrine has lost its power to a considerable degree today (see section 2.12) but the huge array of results obtained by using it stand tall. The doctrine has been accompanied by numerous technical and theoretical improvements that have led to a better understanding of the nervous system. These have affected numerous branches of science, from anatomy to zoology.

The spontaneous generation episode also exemplifies how extended building can join or separate different elements of a theory. For most of the duration of the controversy, the issues of spontaneous generation of life and the appearance of first life on earth were joined at the hip. The latter issue was seen as a related problem and featured prominently in the spontaneous generation discussions. In the twentieth century the first creation of simple organisms continued to be an active interest for biologists and geneticists but the other problem was completely dismissed.[33]

**Rationality of theory selection.** Sometimes discovery and theory-ladenness is related to the issue of rationality of science. The claim is that if one has a method of justification in science completely independent from discovery, especially if the method depends only on neutral evidence, then science is rational. I do not share with such views the same view of rationality (see section 5.3). More importantly, as we have seen in this section, discovery is *not* independent of justification and evidence is *not* theory neutral.

## 3.5 Societal Values

> There are three schools of magic. One: State a tautology, then ring the changes on its corollaries; that's philosophy. Two: Record many facts. Try to find a pattern. Then make a wrong guess at the next fact; that's science. Three: Be aware that you live in a malevolent Universe controlled by Murphy's Law, sometimes offset by Brewster's Factor; that's engineering.
>
> Anonymous

It is easy to understand why some philosophers and scientists are/were afraid of societal values in science. If scientific assumptions/results depend heavily on or result from various social structures or forces, what would happen when these change? It is possible that the scientific assumptions/results change along as well. This presents a serious problem. Everyone wants science to be robust and reproducible from one society to another. To put it another way, science must be immune to social change which can come in two ways:

- *Synchronic change* is the difference between societies existing at a single time.
- *Diachronic change* is the change of societies over time.

The ideal is that science should be able to stay valid through synchronic and diachronic change. One can think of science-fiction ideas which could tell creatively what could go wrong: (1) Our

[33] For the history of the relationship of the spontaneous generation and the first life problem see [Farley, 1977] and for a modern view see [Harris, 2002, 157–159].

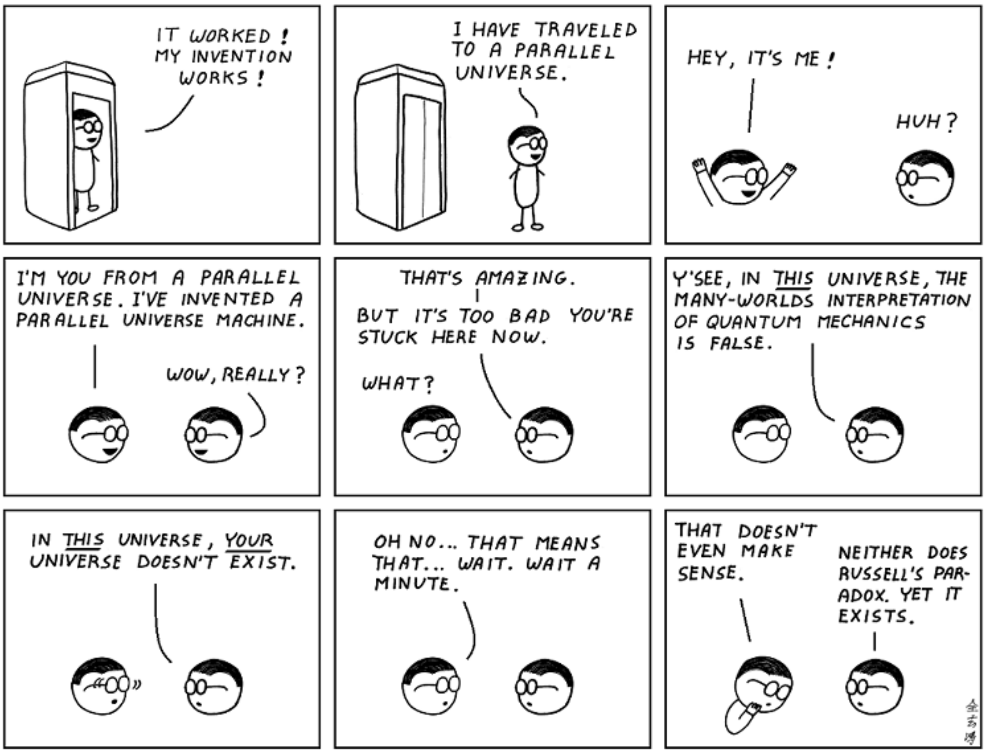



Figure 3.11: A comic from www.abstrusegoose.com/457. [Image licensed under Creative Commons Attribution-Noncommercial 3.0 United States License.]

hero travels to another place and dies there because his medicine is ineffective in that society. (2) Our hero travels to another time and gets stranded there because his time machine fails to work in that society. (See also figure 3.11.) Those that are afraid of societal values in science do not envision such drastic cases but nevertheless there is a real potential for societal values to undermine science in times of synchronic or diachronic change. That is why the ideal of *value-free* science appealed to many philosophers in the past.

Even though the myth of value-free science is shattered in our times, I believe that there is a serious philosophical problem left intact: understanding the types and roles of societal values in science and especially investigating whether any type of societal value breaks the stability of science under synchronic or diachronic change. I believe this to be one of the more salient issues in philosophy of science today. This section is my modest attempt to differentiate the types of societal values in science and see whether or not each type is problematic in the sense that it gets in the way of robust and stable science. I will call this problem the *stability problem* in the face of societal values.

**Conventions.** There is one group of assumptions in science that are accepted almost solely on social grounds: conventions. To put it in another way, the community you belong to determines the conventions you use. Sometimes a group of scientists choose one set of alternatives over the other either by fiat or accident (but, as we shall see below, there might be subtle reasons such as uniformity with other chosen conventions that favour one convention over alternatives) and once you are part of a community that has a particular convention,


Figure 3.12: Australian "metric conversion" stamps of 1973.

then you learn to live with that convention.

At the first sight, conventions seem to be as innocuous as it can get: synchronic or diachronic changes can bring about a new convention but scientists can easily move between conventions and such a change, though possibly inconvenient, is not a threat to stability of science. A cursory look at some conventions seems to enforce this idea.

In section 2.1 I gave an example from mathematics: there are two notational conventions ("boolean" and "Israeli") in set theory. These are linguistic variants and certainly do not cause set theory to be divided into two distinct branches. To adapt the medicine story mentioned above, a set theory "doctor" educated abroad has no problems of curing set theory "illnesses" in Israel, or vice-versa.

Let us move onto *units* which are one of the fundamental conventions. Here is an example of an amusing one:

> A *Thaum* is the basic unit of magical strength. It has been universally established as the amount of magic needed to create one small white pigeon or three normal sized billiard balls. (from *The Light Fantastic* by Terry Pratchett)

Consider the stamps in figure 3.12. Each stamp features a man reporting a quantity both in imperial and metric system. The following paragraph is from an information card accompanying these stamps:

> In 1970, Australia began a ten-year changeover from "imperial" weights and measures to the metric system used by most of the world. In 1973, when important phases of the programme occur, the Australian Post Office issued four stamps aimed at producing interesting visual messages about metric conversion. They feature the four aspects which will most affect the Australian public — length, mass, volume and temperature. The cartoon style was used to "humanise" the subject, and to emphasize that the man-in-the-street is vitally concerned, and will share metrication's benefits — simplified calculations, fewer errors, and efficiency increases.

The mentioned advantages of the metric system over the imperial one show that the choice between two conventions need not be purely accidental. We also saw in the forcing case that there are some reasons offered by the each camp for their convention. This does not change the fact that a convention, after all, is a convention. A group of people agree to use a convention and as the above two examples show, conventions do not seem to cause any stability problem. Or do they?

The above two examples show cases where it is trivial to translate between two conventions. But not all cases of convention choices are benign.

Hasok Chang [2004] tells the complex story of temperature measurement. The picture he presents is much different than my simplistic one above. It is a trivial matter to move from one convention to another *today* because what we have today is the complete and polished set of conventions related to a well-established theory of heat. Before that, when scientists struggled to develop a theory of heat and a universal scale of temperature, there were numerous scales that were either not compatible or not known how they could be made compatible. There can be growing pains in the development stages of a branch of science and measurement systems related to it need not be compatible.

There is one related way of how different conventions can lead to a stability problem. Suppose two conventions are tied respectively to two competing irreconcilable theories. This irreconcilability might transfer to the conventions themselves though of course it does not have to. What I have in mind is something like the following:

Wolfgang Pietsch [2014] draws attention to the recently proposed redefinition of physical base units. This "overhaul of the metric system" is generally referred to as an "explicit-constant formulation" because it involves "new definitions of the kilogram, the ampere, the mole, and the kelvin in terms of fixing a number of fundamental constants of nature." The motivation for these changes are "the accuracy and stability of definitions as well as the more theoretical concern of universality" and "The pragmatic and contextual nature of these criteria is obvious."[34] [Pietsch, 2014, 86] The odd thing about the base unit change is that some experimental statements of the old (new) convention are definitions in the new (old) one! Therefore there is a very fundamental conceptual incompatibility between the two systems.

This is a fascinating topic that ties in various topics from philosophy of science such as underdetermination, scientific revolutions, and incommensurability. For more on this topic,

[34]There are other values involved as well: precision, easiness of realizability, availability, applicability, simplicity and so on. See [Pietsch, 2014, 87–88].

I refer you to Pietsch's article.[35]

To sum up, conventions most of the time do not cause stability problems. But this is not a rule and there is a possibility that one might not be able to move from one convention to another seamlessly.

**Underlying assumptions.** Without a doubt, the most important cause of stability problems in science is the adoption of different set of underlying or guiding assumptions by different research groups or traditions. The problem especially becomes conspicuous if these assumptions conflict with each other.

When one research tradition highly values an underlying assumption that contradicts another tradition's, this can also be seen as a clash of values. A theory is deemed more *valuable* if it caters for the underlying assumption.

The history of medicine shows how much the practices of various traditions differed until a unified theory emerged in the late nineteenth century. Patients were treated quite differently in different schools; for example, blood-letting was performed in some but not others. One of the reasons for such incompatible practices was the different underlying principles.[36]

History of particle physics is full of different guiding assumptions giving direction to research. I touched upon this topic in section 2.11 where I wrote that "The quark episode shows us how guiding assumptions (in this case about integer charges) can have strong influences." There are numerous other such assumptions in particle physics, some with more colorful names, for example, Geoffrey Chew's "nuclear democracy" and Shoichi Sakata's "dialectics of nature" (see the history of particle physics books mentioned in section 2.11).

Today here is a deep division between mainstream and independent economists. The first group takes average income to be a good measure of prosperity; some of the latter try to come up with alternatives such as human development index. Mainstream economists see no end to growth and believe that technological advances will squash all problems in its way. Opponents question this optimism and assert that such growth is not only economically but also physically impossible. One takes nature to be an infinite resource at our service, the others caution us against an environmental crisis. These clashing views have political consequences that cannot be overstated. From global warming to human rights problems, numerous important problems are relevant to this conflict.[37] When some of the most basic assumptions made about economy change, not surprisingly all economics and politics built on it change as well.[38]

When one tradition chooses to work with a different underlying assumption than another, there will be at least difference in the science produced as a result. But if synchronic or diachronic change brings about contradictory underlying assumptions, then for sure there

---

[35] Especially Pietsch's paragraph on page 92 that starts with "A related difference ..." is quite important. It suggests that the three-part separation of theory, experiment, and instrumentation of science can be supplemented by a fourth one: measurement.

[36] For short introductions to the history of medicine see different entries in *Encyclopedia of Medical History* [McGrew, 1985], especially the entries *Pathology*, *Physiology*, and *Bloodletting*.

[37] See [Sabin, 2013; Oreskes and Conway, 2011; Foster et al., 2010; Easterly, 2013; Newell, 2012] for starters.

[38] But of course, the influence between economics and politics is a two-way street.


Figure 3.13: A Dutch stamp issued in 1993 showing cyclists without helmets.

will be stability problems. Indeed, I believe that the main reason for stability problems in science are different underlying assumptions.

**Philosophical differences.** As we have seen in the case of quarks, scientists can have different philosophical inclinations. We could take these to be a special subset of underlying assumptions. Philosophical assumptions of scientists can influence their beliefs in theories.

Surprisingly, there were times even mathematics witnessed philosophical cacophony, for example at the time of the foundational crises after Cantor's work on set theory.

Philosophical differences can lead to stability problem just as underlying assumptions.

**Different ways of making science.** The difference between two research traditions can be in the underlying assumptions, as I have explained above. But sometimes it is not an assumption underlying science that makes the difference, but rather the way science is made. Different methodological and explanatory frameworks can favour one hypothesis over another.

In section 2.14, I praised Naomi Oreskes' historical account of the theory of continental drift. It turns out that there were serious differences of opinion between American geologists and their European counterparts regarding how geology should be made that caused the rift about the status of the continental drift theory. Some of the reasons why American geologists resisted the continental drift theory were also philosophical in nature.

Another example is the string theory controversy that I discussed in section 3.3. A number of physicists have been attacking the string theory for a number of reasons [See Smolin, 2006; Woit, 2007; Baggott, 2013] one of which is its being *unscientific* because it makes no observable predictions. This is a controversy very much about the nature of physics.

Different ways of making science can lead to stability problems if these different ways favour different theories.

**Complexity and importance.** An issue can be so complex that some kind of selection might be necessary to get it out of the quagmire. Recall the controversy about cycling helmets (section 2.8) which is such a complex issue that ties to numerous factors. Another example is climate modeling which turns out to be a very complex undertaking full of uncertainties [See Biddle and Winsberg, 2010]. In both cases, it is absolutely a must to cutdown on uncertainties or variables involved to make headway. This means to make judgement calls on what factors to give importance to. And in these cases, they include political judgements.

This brings me to a more general point: importance is a very essential part of scientific appraisal. First, as we have seen above, scientists need to make judgements based on the importance they assign to various factors in order to negotiate complexity and uncertainty.

Second, the importance of an hypothesis affects the work that goes into checking it. If it is a matter of life or death, we'd better make sure that we get it right. But another hypothesis could be easily brushed off if it is deemed unimportant. One of the main arguments against genetically modified organisms was their early use without proper risk assessment of their effects on health and the environment. The objection did *not* say they were harmful, rather there should be more investigation of their effects since it is such an important issue with possible dire consequences.

Third, importance can change the relation of values as I will mention in section 4.1. For example, in some situations we might give more weight to parsimony compared to simplicity if the theory in question is a risky or controversial one.

As I note at the end of section 2.6, scientists did not dwell too much on the biological origin of various microfossils until they faced the Martian one. The importance of the Martian life hypothesis made scientists studiously investigate the Martian fossils. It became clear that scientists had not shown thorough and careful attention to detail in analysing terrestrial microfossils — some of these were accepted hastily to be of biological origin. The importance of the Martian theory changed the norms of evaluating signs of life and caused scientists to use the new norms to re-evaluate the past attributions of life.

The gist of these points is that importance and risk are factors that can influence values, practices, and content of sciences. Even though importance is a societal value, it might not cause a stability problem if the entire scientific community gives the same importance to an issue. Only when different research traditions assess the importance and risk quite differently from each other might we get instability.

**Essential involvement.** Societal values can be a part of a science as we have seen. It is even possible for a science to be completely immersed in societal values in a very essential way. *Conservation biology* is a case in point. It involves normative questions such as: Why should we protect this species? Why is biodiversity good for us? How does (say) economics relate to biodiversity? These and similar normative questions are part of conservation biology and it looks like they are there to stay. There have been some attempts at developing a more "objective" look at biodiversity but all failed. [See Sarkar, 2010]

Such essential involvement of societal values can lead to instability. There is no guarantee that another community will agree with us on protecting biodiversity at all.

**Early adoption.** Sometimes, to proceed with research, scientists have to accept some assumption as valid, though uncertain at the time. The idea of spontaneous generation (see section 3.4) was left aside by many well before there was conclusive vindication of the germ theory. We also saw the premature acceptance of the neuron theory in section 2.12. Please read again Santiago Ramón y Cajal's wise words I quoted there that start with

> It is a rule of wisdom, and of nice scientific prudence as well, not to theorize before completing the observation of facts. But who is so master of himself as to be able to wait calmly in the midst of darkness until the break of dawn? ... [Ramón y Cajal, 1899]

I think that early adoption can be seen as a special case of underlying assumptions and it can lead to stability problems.

**Incentives** Political, economical, or other incentives can be a serious threat to stability of science in the hands of special interest groups whose aim is solely to protect and improve their agendas. One example is the big pharma bending science for their profit. Another example is "mercenary scientists" representing various special interest groups. Naomi Oreskes and Erik M. Conway's book *Merchants of Doubt* [Oreskes and Conway, 2011] has a subtitle which gives an idea of what I am referring to: *How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*.[39]

Political motivations[40] can have a role as well. The rejection of the "Jewish science" in the Third Reich immediately comes to mind.

*Most of the time*, incentives lead to diachronic instability, that is, at one point the influence of the special interest groups are in some way foiled and "the truth comes out". But at best, we have "lost years" at our hand; at worst, much more.

**Popularity.** Popularity can turn a theory into self-fulfilling prophecy. Lee Smolin [2006] says that string theory is so popular in theoretical physics that it is a must to know it to get a position in a theoretical physics department. So graduate students specialize in string theory and at the fast pace of graduate programs in our day, they hardly have a chance to learn anything else. So you get a monoculture leading to even a narrower monoculture, and it spirals out of control and string theorists eventually grab their pitchforks and chase all survivors out of town.

I am deliberately exaggerating here. Actually, I have no idea how much self-fulfilling prophecy popularity can lead to. It is a possibility I just wanted to draw your attention to.

**Conclusion.** Societal values are part of science. But they are not harmful per se. Problems can arise if they lead to stability problems. This is why it is important to (1) find more about societal values in science; (2) find out which ones cause instability; (3) find out which ones *can* cause stability; and, (4) determine if any stability can be remedied. It is not enough to acknowledge the presence of societal values; we also need to understand them.

---

[39] See also [Michaels, 2008; McGarity and Wagner, 2008] as well as the books mentioned in section 2.13.

[40] Of course I am not claiming that political interests are independent from economical ones.

# CHAPTER 4

# CLASH OF VALUES

Ideally, all values are compatible. A theory can be simple, parsimonious, beautiful, empirically adequate, politically correct, and so on. But as we have seen, in practice, theories cannot have all. Values collide. Why this is so depends on the particulars of the case. But there are some pairs of value types that are more conspicuous at opposing each other. This opposition appears at the level of concrete values.

## 4.1 Parsimony versus Simplicity

> Land bridges are something like Internet company start-ups:
> once they were everywhere but today there are relatively few.
>
> Andrew Alden, geologist

Parsimony tells us that nature is simple and favours theories that finds its domain uncomplicated, uncluttered, unfussy, and unembellished. On the other hand, (theoretical) simplicity tells us that theories (not nature) themselves must be simple and favours theories that simply and clearly cohere with all accepted observations, evidence and background assumptions. These values (or rather their tokens) are frequently seen clashing. At the first sight this is quite surprising. What stops us from having simple theories of simple things?

Let me explain by an abstract example. Suppose that the objects $A_1, A_2, \ldots, A_n$ have respectively the properties $P_1, P_2, \ldots, P_n$ and we are looking for a theory to explain this situation. Most of the time, parsimony requires that we look at each $A_i$ separately and give an individual explanation why the corresponding property $P_i$ applies to it. When you are looking for a basic explanation of why $A_i$ has $P_i$, you do not want to confuse matters by looking at the whole picture. What you end up with is a theory $R_1 \& R_2 \& \cdots \& R_n$ that finds a separate reason $R_i$ for each $A_i$ and $P_i$. On the other hand, simplicity does not care if we give a parsimonious explanation of why each $A_i$ has $P_i$. The important point is the simplicity of the end theory. So simplicity would rather look at all of the $A_i$ and $P_i$ and come up with a single explanation. But this process can come up with a different theory than the above one.

What I have been trying to tell can be summed up by saying that parsimony favours low level theories and simplicity favours high level theories.

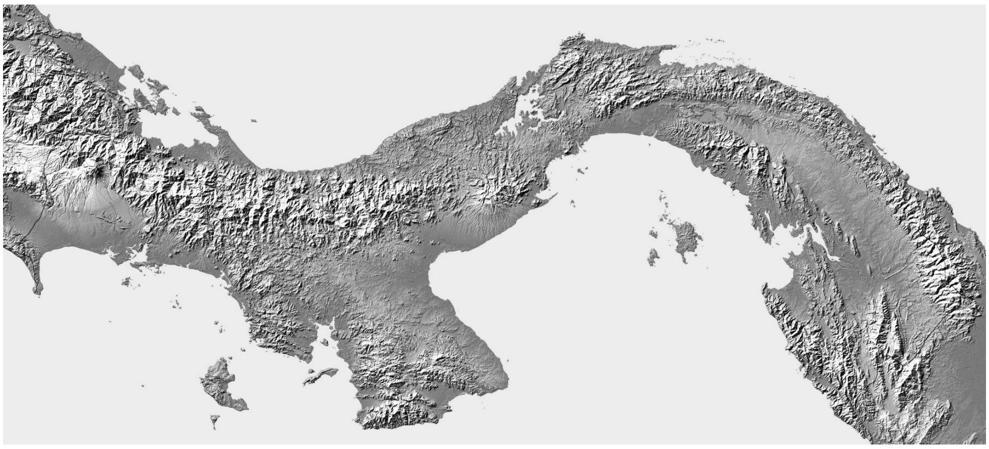


Figure 4.1: When the Isthmus of Panama connecting North and South America was formed about 10 million years ago, it had a huge impact on climate and its environment, and in particular, made it easier for animals and plants to migrate between the two continents. [Image in the public domain from NASA Earth Observatory.]

This clash can be even seen in political history. Consider one country with a history of atrocity crimes targeted against minorities. The high level explanation of this list of crimes is ethnic cleansing. But the ruling class of that country might not like this conclusion and rather look at each atrocity on its own and try to give a justification for it that does not involve ethnic cleansing.

We have seen this clash of values even in philosophy. Consider the now discredited view of operationalism in philosophy of science (and similarly behaviourism in psychology). This view entails dealing with observations and measurements individually and a refusal of a higher level explanation of these using theoretical entities. See [Chang, 2009].

We have already seen an example of this clash in section 2.9 about ToM. Some scientists prefer to explain the behaviour of apes by attributing theory of mind to them, whereas others prefer to stay away from this assumption and provide low level explanations.

How is this conflict between parsimony and simplicity resolved? The number of individual cases that can be provided only *ad hoc* low level explanations seems to be important. If you have to come up with one contrived low level explanation after another, then the high level explanation looks more appealing.

This situation can be seen in the efforts of a country accused of ethnic cleansing in trying to conceal as many atrocities as they can. It is easy to explain away these when they are low in number.

Before plate tectonics brought geology together under a banner in 1960s, the only possible course available for opponents of continental drift theory to account for different geological processes/observations was to give isolated, fragmented, and *ad-hoc* geophysical theories, two of which I will touch upon now.

(1) In 1930s, patterns in fossil record and biodiversity were explained by using a large number of *land bridges* which, like the Isthmus of Panama (see figure 4.1), connect separate areas allowing plants and animals to cross over. As an alternative to continental drift, numerous land bridges up to 5,000 kilometres in length were conjectured to exist

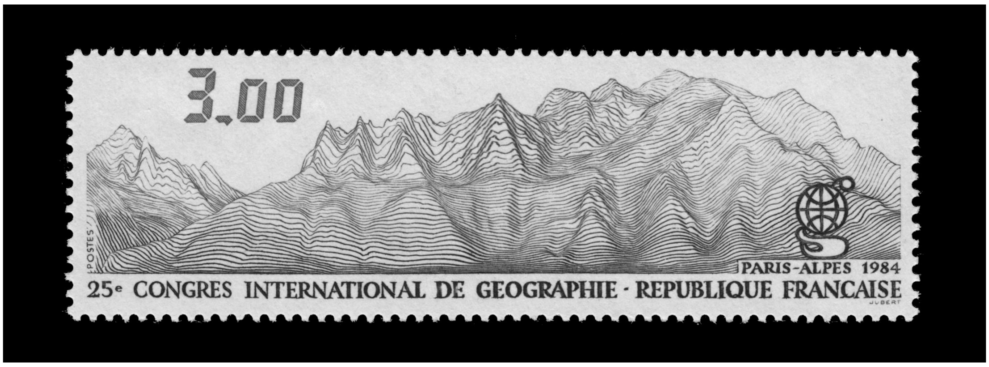



Figure 4.2: 1984 France stamp showing Mount Blanc. Orogenesis or mountain formation was one of the outstanding problems in geology in late nineteenth and early twentieth century.

between continents in the past. Stephen Jay Gould [1979, 165] contends that "The only common property shared by all of these land bridges was their utterly hypothetical status; not an iota of direct evidence supported any one of them." These inter-continental patterns are now explained by the movement of continents and only a few land bridges are acknowledged.

(2) Without theories permitting lateral movement of continents, the problem of *orogenesis* (see figure 4.2) was tackled with by using only vertical movements. The so called *geosyncline* theory was used to explain the formation of all mountain ranges, although some of its problems were recognized at the time (see [Holmes, 1944, chapter XVIII] and [Oreskes, 1999, 18–19] for details).

Plate tectonics replaced all these and similarly contrived theories with a powerful framework able to explain features of Earth in a unified manner. The last opposition to continental drift (or plate tectonics) that I know of came from Harold Jeffreys' *The Earth: Its Origin, History and Physical Constitution* [1976, 481–497] which is a collection of ingenuous alternative explanations, but it amounts to beating a dead horse considering the support plate tectonics had at that time (see figure 4.3).

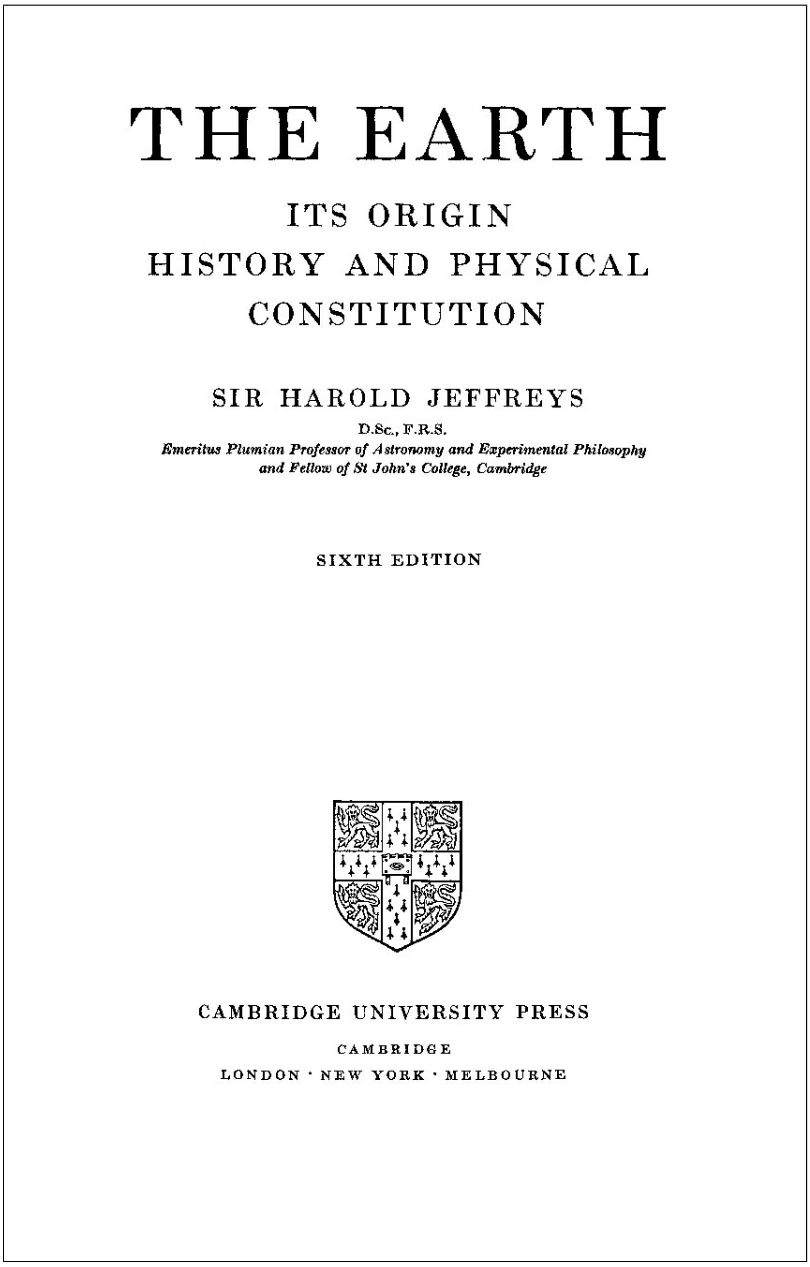
Another case in point is the life on Mars example of section 2.6. Even though the initial announcement was "life found on Mars" the scientific community was hesitant to accept this high level assumption. Rather scientists looked (and are still looking) for ways to explain all the "evidence" of life in the Mars rock using low level physical and chemical mechanisms. But if the day comes when we have a huge number of Mars data that makes low level explanations suspect, only then scientists will except (past) life on Mars hypothesis.

The number of evidence in the theory of mind debate mentioned above seems to have reached a level that makes it more likely to attribute some kind of theory of mind to apes.

There is also a science-stopping aspect of this value clash which I will turn to now.

**Science-Stopping.** We have seen that parsimony can stand up against simplicity and there are three good reasons why.

First, parsimony increases in importance when the simpler theory is the controversial one. For example, alien life is such a theory. Admitting to Martian life (see section 2.6) might



THE EARTH

ITS ORIGIN
HISTORY AND PHYSICAL
CONSTITUTION

SIR HAROLD JEFFREYS

D.Sc., F.R.S.
*Emeritus Plumian Professor of Astronomy and Experimental Philosophy*
*and Fellow of St John's College, Cambridge*

SIXTH EDITION

CAMBRIDGE UNIVERSITY PRESS
CAMBRIDGE
LONDON · NEW YORK · MELBOURNE

Figure 4.3: The title page of the sixth edition of Harold Jeffreys' *The Earth: Its Origin, History and Physical Constitution* [1976]. Jeffreys (1891–1989) was an eminent British polymath who made major contributions to mathematics, statistics, geophysics, and astronomy. He was a strong opponent of continental drift for all of his carrier. In a section of *The Earth*, he gives alternative explanations for various arguments for continental drift. [Image scanned by the author.]

be simpler but it is obviously more sensational.

Second, choosing a parsimonious theory over a simple one does not mean that we have completely forsaken simplicity. In practice, parsimonious theories are also simple, though less than the simplest one.

The third reason is more important and deeper than the first two. Let me explain it by using the concept of *science-stopping* which I am borrowing from Alvin Plantinga:

> But there will be little answer ... if, in answer to the question, Why does so and so work the way it does? or What is the explanation of so and so? we regularly and often reply "Because God did it that way" or "Because it pleased God that it should be like that." This will often be true, but it is not the sort of answer we want at that juncture. ... Claims to the effect that God has done this or that (created life, or created human life) *directly* are in a sense science stoppers. If this claim is true, then presumably we cannot go on to learn something further about how it was done or how the phenomenon in question works; if God did it directly, there will be nothing further to find out. [Plantinga, 2001, 356]

Although Plantinga's context[41] is unrelated to mine here, I will adopt his notion of science-stopping.

In a very real sense, high level assumptions are science stoppers. Once you accept a high level hypothesis, you no longer look for a low level hypothesis.[42] A high level assumption gets in the way of finding low level mechanisms or explanations. High and low level hypotheses generally contradict each other, so choosing a high level one is tantamount to rejecting low level ones. Since parsimony tells us to go with low level hypotheses and simplicity tells us to favour high level ones, there is the danger that simplicity can stop science.

One of the things I like in Louise Barrett's book *Beyond the Brain: How Body and Environment Shape Animal and Human Minds* [2011] is that she gives a lot of (natural and artificial) examples of very complex behaviour caused by very simple mechanisms. I cannot do justice to her examples here. The lesson of those examples is that you should not jump to high level conclusions about complex behaviours. One really needs to lay one's intuitions aside and look diligently to find these simple mechanisms. It is prudent to keep searching for low level explanations, not because that such explanations are inherently better, but because it is too easy to accept a high level explanation and be done with it.

We may think of this "do not let science stop" approach as a tacit heuristic. But sometimes it is explicitly stated: The Morgan's Canon that I introduced on page 44 and the astropaleontological principle on page 32 are examples of this kind.

When do we cease our search for low level theories and opt for higher level ones? There is no rule and it really depends on the circumstances of the situation. What makes the difference is the context and importance of the hypothesis in question. While scientists might determinedly search for low level hypothesis in a context, they might easily settle for a high level hypothesis in another. This is an example of societal values at work as I mentioned in section 3.5.

---

[41] Plantinga argues that a Christian science that does not stop science is possible.

[42] Note that I do not claim that all high level assumptions are similar or on par with "because God did it that way" assumptions.

## 4.2 Precision

> Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise.
>
> John Tukey [1962]

Precision is a desirable value in different aspects of science: experimentation, prediction, etc. But it sometimes runs into other values.

**Precision and clarity.** In his autobiographical book *Unended Quest* [2002], Popper compares precision and certainty with clarity:

> What I do suggest is that *it is always undesirable to make an effort to increase precision for its own sake — especially linguistic precision — since this usually leads to loss of clarity,* and to a waste of time and effort on preliminaries which often turn out to be useless, because they are bypassed by the real advance of the subject: *one should never try to be more precise than the problem situation demands.*
>
> I might perhaps state my position as follows. *Every increase in clarity is of intellectual value in itself; an increase in precision or exactness has only a pragmatic value as a means to some definite end* — where the end is usually an increase in testability or criticizability demanded by the problem situation (which for example may demand that we distinguish between two competing theories which lead to predictions that can be distinguished only if we increase the precision of our measurements). [Popper, 2002, 22]

Suppose you want to do a back of the envelope calculation to roughly determine the area of a circular garden using its radius in order to buy enough seeds. Would you take $\pi$ to be equal to

$$3.14159265358979323846264338327950288419716939?$$

I bet not. Taking such a $\pi$ value will only make the calculation more cumbersome and tiresome. This is unnecessarily precise for the job at hand.

Consider *taxonomy* in biology which aims to identify, classify, and name organism groups. One needs to decide on an appropriate methodology to organize the classification. This is already problematic (see figure 4.4). Assuming you have successfully chosen a method of classification, there is a second problem: how fine grained do you make your classification? There is a trade-off between clarity and precision. Clarity demands that you find similarities between organisms. On the other hand, precision demands that you are sensitive to differences between organisms. Wearing magical clarity glasses you would get a tree of life with few branches and with magical precision classes you would get a tree with enormous number of branches. A sweet spot needs to be found between them.

There is no benefit of being precise more than the job at hand requires. It is a balancing act between certainty and clarity.

**Precision and resources.** Precision is most of the time not free to get. It costs resources which could be time, money, computing power, infrastructure, organization, so on. Anyone who worked on computer modelling in any branch of science appreciates the value of computer

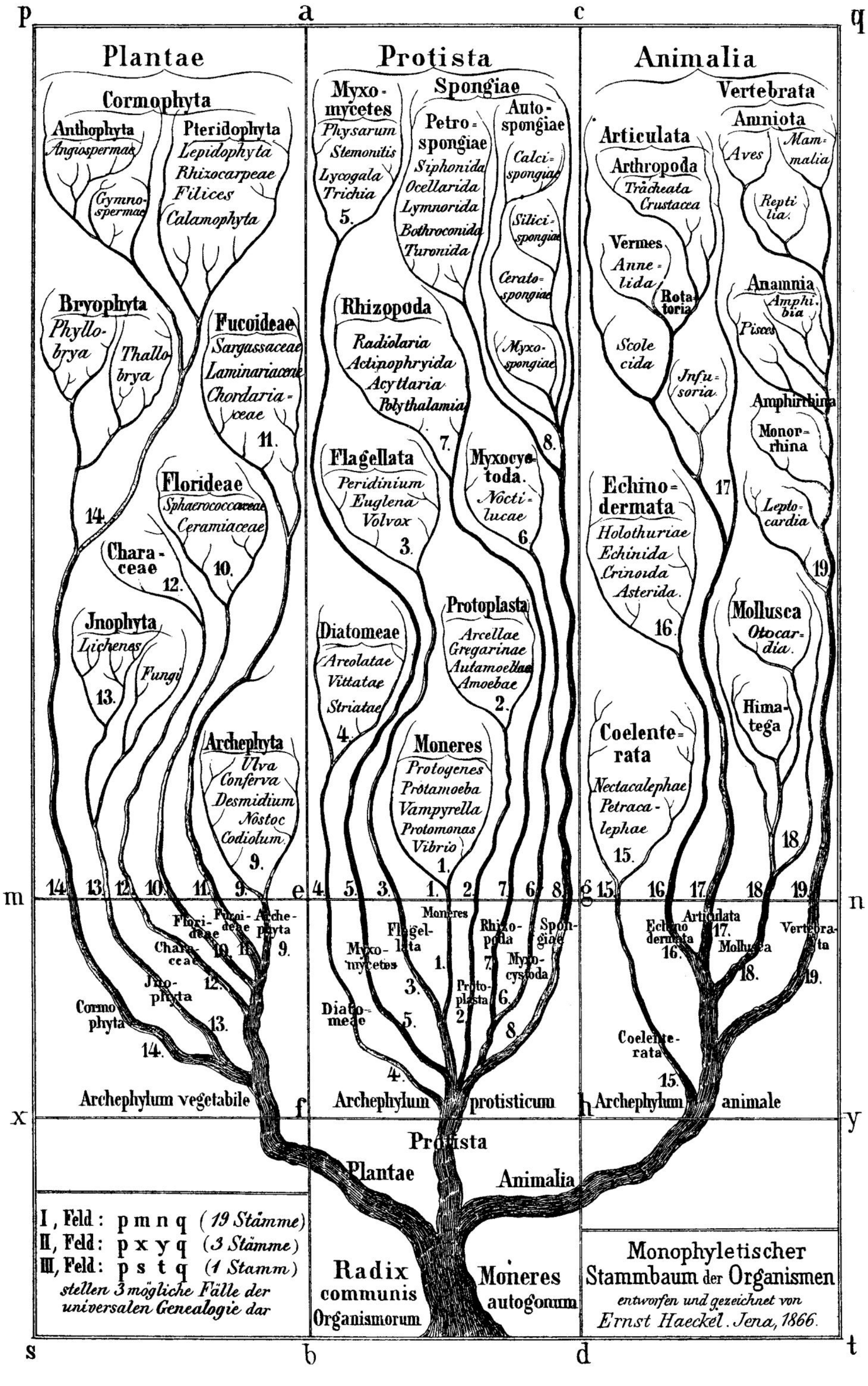



Figure 4.4: Until Darwin, biological classifications were built on traits shared between organisms. The evolutionary theory made a second scheme possible: classifying organisms by their common ancestry. There are some in-between approaches as well. See [Garvey, 2007, chapter 8] for an introduction. Darwin used the *tree of life* metaphor to describe the relationships between organisms in *On the Origin of Species*, though he only gave a hypothetical drawing. The above tree is the one drawn by Ernst Haeckel in *Generelle Morphologie der Organismen* (1866). [Image in the public domain from www.wikimedia.org.]

time. But sometimes precision comes at a much larger price. Writing about how "[b]etween 1945 and 1965, digital computers revolutionized weather forecasting," Paul N. Edwards highlights the changes this move to precision required:

> Computer models for weather forecasting rapidly came to require hemispheric data, and later global data. Acquiring these data with sufficient speed demanded automatic techniques for data input, quality control, interpolation, and "bogusing" of missing data points in sparsely covered regions. The computer itself helped solve these problems, but their full resolution required *substantial changes to the global data network. Like all infrastructure projects, these changes involved not only scientific and technological innovation, but also institutional transformation.* [Edwards, 2010, 111, my emphasis]

Edwards goes onto discuss more recent work on climate models and how limited resources necessitates resource allocation. Such an allocation involves a decision regarding the importance of different parameters: If you do not have resources to make everything precise, then you better make those that matter more precise. But what *matters* has a strong social element as I discuss in section 3.5.

**Accuracy and convenience.** Surrogates in medicine is a topic I will look at in section 4.5. There is a forerunner of the issue discussed in Claude Bernard's 1865 book *An Introduction to the Study of Experimental Medicine*:

> Let me further point out that the reduction of physiological phenomena to an expression in kilograms of body weight is vitiated by many sources of errors. For a certain number of years this method has been used by physiologists studying the phenomena of digestion. We observe, for instance, how much oxygen or how much food an animal consumes in a day; we then divide by the animal's weight and get the intake of food or of oxygen per kilogram. This method may also be applied to measure the action of toxic or medicinal materials. We poison an animal with a maximum dose of strychnine or curare, and divide the amount by the weight of the body, to get the amount of poison per kilogram. *For greater accuracy in the experiments just cited, we should have to calculate, not per kilogram of the animal's body taken as a whole, but per kilogram of blood and of the unit on which the poison acts;* otherwise we could not deduce any direct law from the reductions. But other conditions would still remain to be established similarly by experiment, conditions varying with age, height, state of digestion, etc.; in these measures, physiological conditions should always hold first rank. [Bernard, 1865, 135, emphasis added]

Physicians of the era looking for a "reduction of physiological phenomena" used *amount by the weight of the body* as a surrogate for that phenomena. Barnard suggests as an improvement to the old one to use *amount by the weight of the unit* where the *unit* is the related physiological system. For example, a poison injected into blood should be measured *per kilogram of blood*, not *per kilogram of body weight.*

Here we see a clash of values between accuracy and easiness of measurability. As Bernard writes, using the amount by the weight of the related unit is more accurate than by using the whole bodyweight but this comes at a cost of convenience. Weighing the bodyweight of an animal is easy. But how can you weigh the blood? This can be done or inferred somehow, but nevertheless, it is at least inconvenient or harder to figure out.

In current medicine both Bernardian and non-Bernardian approaches have their place, that is, we see *per bodyweight* used frequently along with more refined units. It is a matter of balancing accuracy and feasibility.

## 4.3 Reproducibility and Its Discontents

> Replication is central to the progress of science: if others cannot reproduce the evidence backing a scientific claim, then the claim loses status as scientific knowledge. This process differentiates science from other ways of knowing for which the power, authority, ideology, or persuasiveness of the person making the claim determines its truth.
>
> Errington et al. [2014]

One of the important tenets of scientific experiments is their reproducibility. The view that reproducibility is one of the main principles of the scientific method and even a hallmark of science is prevalent. But the problem with reproducibility is that it is not always easy to get. There is a crisis of reproducibility in biomedicine and psychology widely acknowledged today.

In 2011, three scientists working at the drug giant Bayer reported their company's attempts at reproducing experiments from 67 biomedical publications. "This analysis revealed that only in 20–25% of the projects were the relevant published data completely in line with our inhouse findings." [Prinz et al., 2011] The biopharmaceutical company Amgen trying to reproduce the findings of 53 "landmark" studies in cancer research published in top journals managed to reproduce only 6 of the 53 studies [Begley and Ellis, 2012]. As I am writing this paragraph in late 2015, the project, *PsychFileDrawer*[43] dedicated to replication of published articles in experimental psychology shows a 14 out of 50, or 28% reproduction success rate.

This crisis in reproducibility has led to creation of a number of projects to systematically check published experiments in different fields, changed the way journals handle submissions, caused scientists to campaign for openness of data and methods, and suggest numerous possible solutions to overcome different biases in their fields. For details of some of these developments see the books mentioned in section 2.13, the *Nature* special on reproducibility at http://www.nature.com/nature/focus/reproducibility/index.html, as well as the Bayer/Amgen papers mentioned above. I will be able to mention only a small subset of the suggested solutions in this section and my interest lies more with the relations with other values.

**Irreproducibility can be good.** Generally, irreproducibility points to a slipped cog in the machine called science, but this need not always be the case. Debugging why an experiment or result is not reproduced can lead to new insights or discoveries.

The reason for irreproducibility can be an unknown factor affecting the experiments. For example, when the spontaneous generation debate was in full swing in mid-nineteenth century, there was a recalcitrant problem about microbial life after extensive boiling of organic solutions — some scientists consistently had it to the chagrin of those who were able to avoid.

[43] http://www.psychfiledrawer.org/view_article_list.php

Figure 4.5: Ferdinand Julius Cohn (1828–1898) was a founder of modern microbiology and bacteriology. [Image adapted from https://commons.wikimedia.org/wiki/File:Portrait_of_Ferdinand_Julius_Cohn_Wellcome_M0009956.jpg by Wellcome Images licensed under Creative Commons Attribution 4.0 International.]

It was the German biologist Ferdinand Cohn (figure 4.5 ) who solved the puzzle in 1876 by showing that some bacteria have heat resistant forms. It turns out that those that had microbes after boiling were using infusions of hay that contained *Bacillus subtilis* which can produce endospores when subjected to a deleterious environment. See [Strick, 2000, 28] for more on this episode.

The following quotation is from biologist Mina Bissell, known for her research on cancer:

> A third example comes from a non-malignant human breast cell line that is now used by many for three-dimensional experiments. A collaborator noticed that her group could not reproduce its own data convincingly when using cells from a cell bank. She had obtained the original cells from another investigator. And they had been cultured under conditions in which they had drifted. Rather than despairing, the group analysed the reasons behind the differences and identified crucial changes in cell-cycle regulation in the drifted cells. This finding led to an exciting, new interpretation of the data that were subsequently published. [Bissell, 2013, 334]

As this example shows, irreproducibility can lead to new appreciation of the results. Sometimes lack of reproducibility does not lead to new discoveries but rather methodological insights. One measure to abate the above mentioned reproducibility crisis in biomedicine is the step taken by a number of journals (including *Nature*) to remove the length limit on the methodology section of the submitted papers. The hope is that the authors will be more careful in detailing their methodology and this will in turn improve the robustness of their experiments. Noorden [2014] tells how one detail omitted in the methodology section of a biology paper has led to failed replication attempts.

Hence irreproducibility may point to new understanding, but this does not mean that we should welcome it with open arms. If it is systematic, it can make accepted results suspect and even make a whole field of science questionable.

**Fraud and Reproducibility.** It is only natural to think that fraudulent practices lead to irreproducible research. But this is not always the case. Actually, there is good reason to think that fraudulent practices frequently lead to reproducible results, though not for the right reasons. For example, when one cooks up data confirming a hypothesis, he is more likely to choose a hypothesis that he thinks will stand the test of time. If he is to avoid detection, he cannot just put forward any hypothesis, but rather one that is consistent with background assumptions and predicted by theory as we have seen in section 2.10. Such a hypothesis is more likely to be backed up by future experiments.

German physicist Jan Hendrik Schön working at the renowned Bell Labs was a rising star in the field of nano-electronics when it came to light in 2002 that he had committed extensive misconduct in his work. He had made up his results and not seen any of the effects he claimed to have observed. But in the following years other teams of scientists did real experiments and went on to observe the effects that Schön claimed to have observed. In her book about the Schön scandal, Eugenie Samuel Reich writes that:

> The investigators of Schön had considered this possibility, writing in their report that the finding of scientific misconduct against Schön would remain valid even if the science of Schön's claims was validated in the future. Even so, it was interesting to think about what might have happened if this research, or other work similar to Schön's, had been completed sooner. Schön might have earned the credit for being one of the first to jump into a novel area, while the scientists who did the work to test his claims appeared to come in second. . . .
>
> Hooked by this idea, I couldn't help but wonder whether it had occurred to Schön too. Had Schön been banking on the possibility that his false but plausible scientific claims might one day be validated through the honest work of others? At the same time, the appearance of similarity between Schön's work and other genuine results helped to explain why other scientists had been so willing to believe Schön in the first place. Schön had apparently imitated the outline of real scientific breakthroughs well enough that his data seemed both groundbreaking and plausible at the same time. [Reich, 2009, 3–4]

Just because a research is fraudulent does not mean that it is irreproducible and just because a research is reproducible does not mean that it is not fraudulent.

**Incentives versus Reproducibility.** One might think that drug companies determine the effectiveness and safety of treatments by sound and unbiased methods including double-blind clinical trials. But if these results are systematically irreproducible, then there must be something wrong with the methods. There are huge profits in pharmacology and when one aims to make profit in this competitive field, robustness of clinical trials and experiments are sacrificed for positive results. I have already mentioned a very limited set of these problems in section 2.13. But as I have said in the beginning of that section, the shocking nature of biomedicine is too vast to tell here and I refer you to the books mentioned in that section. Initially, one might think that double-blind trial is the cure for all bias in medicine, but nothing can be further than truth. Although they eliminate some biases, they do not eliminate all (see appendix B).

The incentives that lead to irreproducibility are not necessarily associated with big companies. Individual scientists can be lured by incentives as well. Job security, money, success, and "publish or perish" policies can motivate scientists to massage data or commit outright fraud. For example, Diederik Stapel (discussed in section 2.10) explains the motives for his fraud in an interview with *The New York Times* as follows:

> What the public didn't realize, he [Diederik Stapel] said, was that academic science, too, was becoming a business. "There are scarce resources, you need grants, you need money, there is competition," he said. "Normal people go to the edge to get that money. Science is of course about discovery, about digging to discover the truth. But it is also communication, persuasion, marketing. I am a salesman. I am on the road. People are on the road with their talk. With the same talk. It's like a circus." [Bhattacharjee, 2013]

When incentives take precedence over robustness, unsurprisingly, irreproducibility follows.

**Innovativeness versus Reproducibility.** Innovation of new experimental techniques and technology can lead to reproducibility problems.

One obvious way this can happen is if there are faults at the new technology. In September 2011, scientists working at the OPERA experiment in CERN announced that they detected neutrinos traveling faster than light. In March 2012 they retracted the result due to a loose fiber-optic cable.

New technology can be unreliable and can lead to reproducibility problems. Shapin and Schaffer [1985] (cf. [Principe, 1998, 111–117]) consider the seventeenth century pneumatic experiments which featured scarce and undependable air-pumps. The way out of this irreproducibility conundrum was carrying out experiments in front of an audience as depicted in Derby's celebrated painting *An Experiment on a Bird in the Air Pump* (see figure 5.8 on page 151). The role of the audience was to witness the experiment and remove the possible doubts that can arise about experimenter's claims.

One reason why the spontaneous generation controversy mentioned above lasted so long was the intricacies of the experiments. Contamination could be avoided only by the very skilled experimenters and even they had occasional difficulties. Even the slightest contamination resulted in microbial life in the experimental set-up confirming the spontaneous generation theory.

New experimental set-ups continue to cause problems today. Bissell draws attention to the complexity and sensivity of biochemical experiments:

> Many scientists use epithelial cell lines that are exquisitely sensitive. The slightest shift in their microenvironment can alter the results — something a newcomer might not spot. It is common for even a seasoned scientist to struggle with cell lines and culture conditions, and unknowingly introduce changes that will make it seem that a study cannot be reproduced. [Bissell, 2013, 334]

She attributes some of the reproducibility problems in biomedicine to the sophistication and difficulty of the techniques.

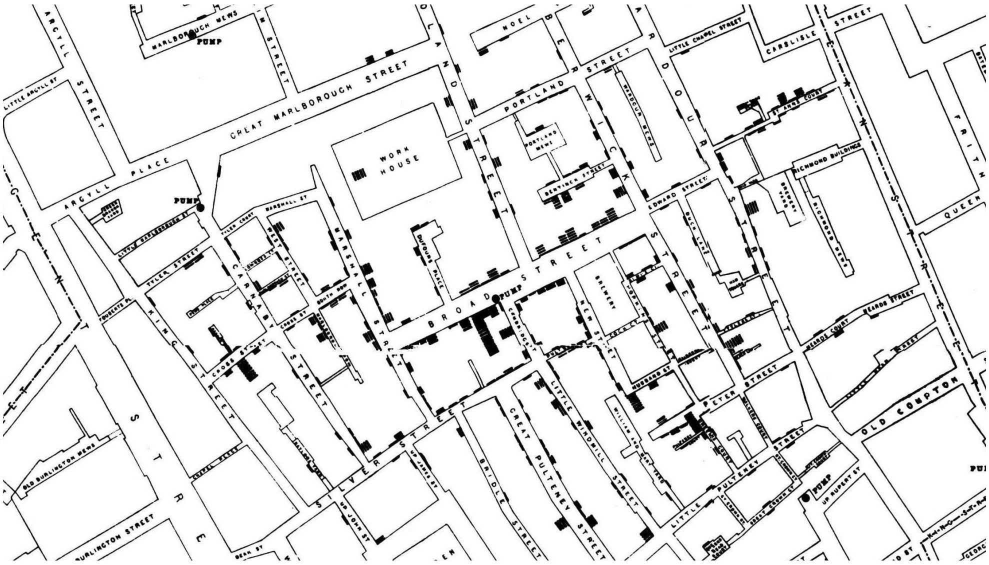



Figure 4.6: The English physician John Snow (1813–1858) was successful in tracing the source of an outbreak of cholera in 1854 in London to contaminated water in a well. Snow mapped out the cases of cholera and recognized that they were spread around the Broad Street pump. The image above is the map made by Snow in 1854 in which cholera cases are highlighted in black. Snow's method of identification of the focus of infection set a precedent which led to formation of epidemiology. [Image in the public domain from commons.wikimedia.org.]

Pushing the limits of technique, technology, and methodology brings in an element of unknown and uncertainty which can lead to irreproducibility, at least until the new becomes mainstream.

**Fruitfulness and Reproducibility.** Data analysis is an important part of different fields of science which can result in important discoveries (see figure 4.6). But not all data analysis ends up to be fruitful. Sometimes a scientist gathers data to test a particular hypothesis which turns out not to be supported by it. Instead of throwing away the useless dataset, he might try to put it to some use by coming up with a hypothesis related to the data. Without any particular objectives in mind, he can try a few models until one fits the data. This *data dredging* is a double-edged process. Searching the data to find all correlations between the variables in the dataset might turn out to be fruitful and teach us new relations hidden in the data. On the other hand, every data set has its random peculiar relations that do not capture real relations. Search enough, you will eventually hit one of these accidental relations. Clearly, such relations would not be reproducible. In an editorial in *The BMJ* (formerly the *British Medical Journal*) authors state that

> Data dredging is thought by some to be the major problem: epidemiologists have studies with a huge number of variables and can relate them to a large number of outcomes, with one in 20 of the associations examined being "statistically significant" and thus acceptable for publication in medical journals. The misinterpretation of a $p < 0.05$ significance test as meaning that such findings will be spurious on only 1 in 20 occasions unfortunately continues. When a large number of associations can be looked at in a dataset where only a few real associations exist, a p value of $0.05$ is compatible with the large majority of findings

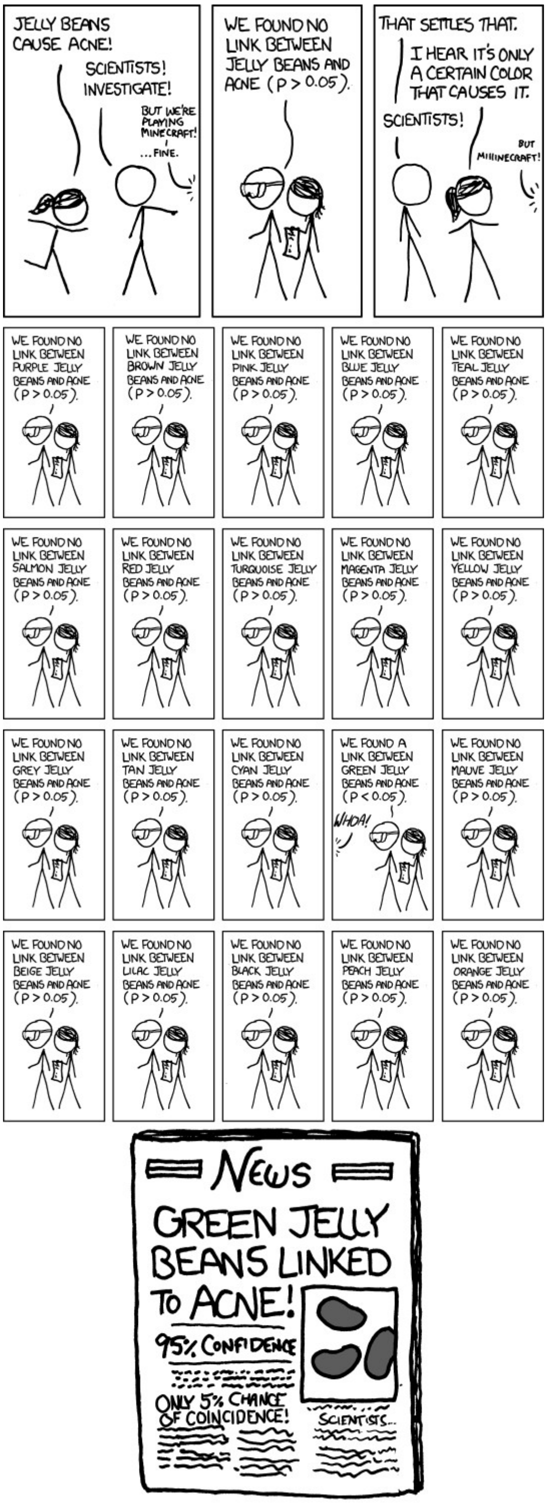



Figure 4.7: A comic by Randall Munroe about data dredging. [Image from http://xkcd.com/882/ licensed under Creative Commons Attribution-NonCommercial 2.5 License.]

still being false positives. These false positive findings are the true products of data dredging, resulting from simply looking at too many possible associations. [Smith and Ebrahim, 2002]

This aspect of data-dredging is the topic of a brilliant comic by Randall Munroe reproduced here as figure 4.7. The comic is hardly readable in printed text so you might want to read it from http://xkcd.com/882/ and a transcription of it is given in appendix C. There is a site dedicated to explaining xkcd comics and this is from their explanation[44] of this comic:

> At first the scientists do not want to stop playing the addictive game Minecraft …but they do eventually start.
>
> The scientists find no link between jelly beans and acne (the probability that the result is by chance is more than 5% i.e. $p > 0.05$) but then Megan and Cueball ask them to see if only one colour of jelly beans is responsible. They test 20 different colors each at a significance level of 5%. If the probability that each trial gives a false positive result is 1 in 20, then by testing 20 different colors it is now likely that at least one jelly bean test will give a false positive. To be precise, the probability of having no false positive in 20 tests is $(0.95)^{20} = 35.85\%$. Probability of having no false positive in 21 tests (counting the test without color discrimination) is $(0.95)^{21} = 34.06\%$.

So it is more likely that the correlation between green jelly beans and acne is a fluke. But if it is a fluke then it is irreproducible.

Data dredging can be fruitful but it can also lead to false positives and spurious correlations. These irreproducible results are not only seen in comics but in real science, see [Smith and Ebrahim, 2002] for examples from biomedicine.

**Generality and Reproducibility.** As I argued above, in some cases, lack of reproducibility can teach us about factors not known and lead to useful information. It is also possible that irreproducibility make us revise the generality of a theory. For example, if a particular biochemistry experiment fails to be reproduced, this may be due to a different cell line used as the above quotation from Mina Bissell shows. Instead of throwing out the experiment as irreproducible, scientist can restrict its scope.

An illustration of this aspect of reproducibility is an incisive critique of a particular trend in behavioural sciences by Henrich et al. [2010]. The authors wittily call people from Western, Educated, Industrialized, Rich and Democratic societies *WEIRD*. The authors mention that "A recent analysis of the top journals in six sub-disciplines of psychology from 2003 to 2007 revealed that 68% of subjects came from the United States, and a full 96% of subjects were from Western industrialized countries, specifically those in North America and Europe, as well as Australia and Israel [Arnett, 2008]. The make-up of these samples appears to largely reflect the country of residence of the authors, as 73% of first authors were at American universities, and 99% were at universities in Western countries. This means that 96% of psychological samples come from countries with only 12% of the world's population." They present the problem as follows:

> Behavioral scientists routinely publish broad claims about human psychology and behavior in the world's top journals based on samples drawn entirely from

[44] http://www.explainxkcd.com/wiki/index.php/882:_Significant

. . . WEIRD societies. Researchers — often implicitly — assume that either there is little variation across human populations, or that these "standard subjects" are as representative of the species as any other population. Are these assumptions justified? [Henrich et al., 2010]

They go on to convincingly show that these assumptions are *not* justified. The universal claims made with regards to human psychology by some behavioural scientists turn out to be not reproducible in non-WEIRD societies. The conclusion they draw is that the scope of the claims should be restricted to WEIRD cultures.

Reproducibility is a very complex value that is currently a focus of attention in scientific literature, particularly in behavioural and medical sciences. This section could only serve as a starting point and there needs to more more discussion of this value in philosophy of science.

## 4.4 Essential Tension

The interplay of old tradition and new necessities becomes at times very curious.

H. G. Wells, *Mankind in the Making*, 1903

According to Kuhn [1977c, 226], scientific revolutions

> are episodes — exemplified in their most extreme and readily recognized form by the advent of Copernicanism, Darwinism, or Einsteinianism — in which a scientific community abandons one time-honored way of regarding the world and pursuing science in favor of some other, usually incompatible, approach to its discipline.

Kuhn's account of scientific revolutions changed in time [see Nickles, 2013] and there is a debate in philosophy of science circles about what scientific revolutions are and whether they occur in science.

> It is controversial whether or not there have been any revolutions in the strictly Kuhnian sense. It is also controversial what exactly a Kuhnian revolution is, or would be. Although talk of revolution is often exaggerated, most analysts agree that there have been transformative scientific developments of various kinds, whether Kuhnian or not. However, there is considerable disagreement about their import. The existence and nature of scientific revolutions is a topic that raises a host of fundamental questions about the sciences and how to interpret them, a topic that intersects most of the major issues that have concerned philosophers of science and their colleagues in neighboring science studies disciplines such as history and sociology of science. [Nickles, 2013]

This debate about revolutions is not something I could settle here but I want to look at one aspect of revolutions, namely the role of innovative thinking. Kuhn [1977c] claims that long stretches of normal science which follow *convergent thinking* are disturbed by the *divergent thinking* of scientific revolutions. What he means by convergent thinking is following an existing consensus and trying to use the methods and techniques available at hand. But when research breaks with the old and follows a new direction that involves innovation, flexibility, and open-mindedness, you have divergent thinking. There is an *essential tension* between these two types of thinking since they are exclusive by nature:

> [N]ormal research, even the best of it, is a highly convergent activity based firmly upon settled consensus acquired from scientific education and reinforced by subsequent life in the profession. ... But revolutionary shifts of a scientific tradition are relatively rare, and extended periods of convergent research are the necessary preliminary to them. As I shall indicate below, only investigations firmly rooted in the contemporary scientific tradition are likely to break that tradition and give rise to a new one. That is why I speak of an "essential tension" implicit in scientific research. To do his job the scientist must undertake a complex set of intellectual and manipulative commitments. Yet his claim to fame, if he has the talent and good luck to gain one, may finally rest upon his ability to abandon this net of commitments in favor of another of his own invention. Very often the successful scientist must simultaneously display the characteristics of the traditionalist and of the iconoclast. [Kuhn, 1977c, 227]

This paragraph also has an important footnote:

> Strictly speaking, it is the professional group rather than the individual scientists that must display both these characteristics simultaneously. ... Within the group some individuals may be more traditionalistic, others more iconoclastic, and their contributions may differ accordingly. Yet education, institutional norms, and the nature of the job to be done will inevitably combine to insure that all group members will, to a greater and lesser extent, be pulled in both directions. [Kuhn, 1977c, 227–228]

Although the essential tension is most of the time discussed in the context of scientific revolutions, I think it deserves to be thought independently especially since it is not exclusive to scientific revolutions.

The essential tension between convergent and divergent thinking is also about clash of values. Competing theories, methods, etc. can engender different types of thinking. A theory that confirms with the traditions has convergent thinking on its side and an out of the box one has divergent thinking on its side. A choice between two types of thinking can result in a choice between theories. That is why essential tension is tied to clash of values and it is important to understand it for both on its own sake and for its relation to values.

The upside of convergent thinking is clear as day. To begin with, using readily available methods can save time and energy — it facilitates intellectual economy. Moreover, new ideas may not immediately fit in with the existing science and it might be necessary to rethink or reinterpret what is old in terms of what is new to make them compatible; and there is no guarantee of achieving such a feat. It is a risk to leave aside ideas that worked successfully in the past and sail uncharted waters. It is not surprising to see scientists that cling to convergent thinking. (See also the Planck quote on page 156.)

There is a further important point about convergent thinking: it is a very powerful tool for discovery/construction. Scientists employ methods that worked in similar situations in the past to solve new problems. They try to expand their old theories or come up with similar theories that explain the new.

When I was reading about history of particle physics [see Johnson, 2000; Riordan, 1987; Crease and Mann, 1986; Oerter, 2006] it surprised me how much physicists tried using old theoretical approaches and tricks in novel situations. Similar field theories, symmetry arguments, tricks such as postulating new quantum numbers are used again and again to tackle new particles or forces. They keep rehashing old ideas until it fails to work — that is only when the innovative genius comes into play.

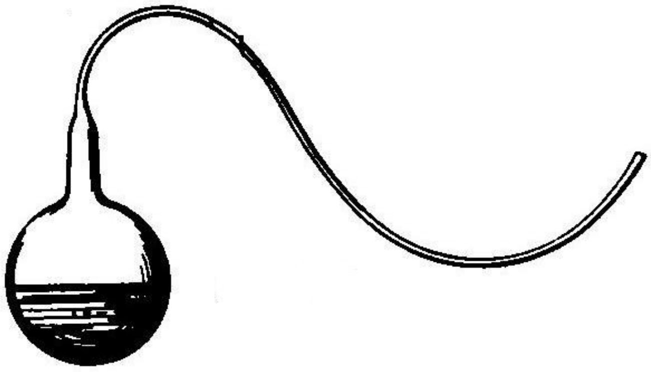


Figure 4.8: Part of a 1861 drawing by Pasteur showing a swan-necked flask used in his experiments. This flask trapped air-born particles in its neck and caused delayed microbial formation in the flask.

When convergent thinking is such an important and successful aspect of science, it is no wonder that scientists are resistant to divergent thinking. We have seen how various new theories were adopted slowly and with resistance in chapter 2. This is especially so if the new goes against something of the old. Martian origin of the SNC meteorites  was problematic in view of the geophysics of the time. Quarks had the unfortunate property of having fractional charge.

As Kuhn's footnote given above acknowledges, in each episode we see some individuals more in favour of divergent thinking than others. If the new idea can gather enough support it can overthrow the incumbent. But it is possible that it might never achieve this deed and left in the sidelines or completely forgotten. There is no guarantee that a new idea will be fruitful or win over the scientific community.

The turn to divergent thinking is more forcefully seen at times when the accepted theory runs into troubles. New ideas to tackle the issues can surge. At these times the likelihood of challenging the consensus of the field is dramatically increased. Scientists become more involved in the foundational problems of the field.

Divergent thinking may even cause us to question our various prior commitments. As I mention at the end of section 2.6, the norms of evaluating signs of life has changed after the Martian life debate and these new norms are used to re-evaluate the past attributions of life.

It is a mistake to think that divergent thinking is exclusive to scientific revolutions. I think Kuhn [1977c, 232] overplays his hand when he claims that "Except under quite special conditions, the practitioner of a mature science does not pause to examine divergent modes of explanation or experimentation." In science (and in our daily lives) we may explore novel approaches if the tested ones fail. This may be to solve a minor problem without any revolutionary outcomes.

For example, the experimental set-ups to test for spontaneous generation changed gradually during the debate that I mention in section 3.4. There were a lot of small but novel improvements made like Pasteur's swan-necked flasks (see figure 4.8). These were out of the box solutions to deal with concrete problems. But I would not call any one such improvement revolutionary.

Novel thinking and following tradition are two important aspects of science. But these two should be regarded in their own right and not necessarily tied to the revolution issue especially because the essential tension is reflected at the level of values.

## 4.5 Quality

> Data quality is an essential characteristic that determines the reliability of data for making decisions.
>
> from *IBM Analytics* website

Every group of scientists would want their data to be high-quality, wide-ranging, detailed, and easily processed. But there can be some obstacles in obtaining such data.

Chess engines are programs that play chess. This type of programming clearly shows the conflict between quantity and quality. Ideally, you would want your engine to contain as much (positional) rules about chess as possible and be as fast as possible. But the the more positionally aware it becomes, the slower it must get as it has to apply all these rules to its calculations, slowing it down. There is a trade off between how much an engine knows and how fast it is at analyzing new positions — the more rules to check, the slower it becomes.

Economics is one of the obstacle in the way of obtaining both quantity and quality in science. A medical researcher cannot conduct elaborate trials on a big population without having necessary funds. A climate modeler cannot process raw data without having the necessary computer resources. In such cases, the available resources would restrict the scope or quality of the research in question.

There are other interesting issues related to quality of data, two of which I will look at in this section: surrogates and anecdotes.

**Surrogates.** Suppose you want to measure a quantity A, but doing so is not feasible for some reason. What you can do then is to find another quantity B somehow correlated to A and measure B instead. In such cases, B is a called a *surrogate* of A. Measuring B would give indirect information about A. If the correlation between them is quite strong, measuring B can be almost as good as measuring A.

There are two aspects of having a suitable surrogate: (1) the theory that gives the correlation between A and B, (2) measurability of B. Both of them must be good to make that surrogate successful.

Using surrogates is a very widespread practice in our daily lives. For example, one can check whether there is any light in the room to determine if someone is sleep or not. Using surrogates is a very widespread practice in science as well. Let me mention a couple of examples from hydrogeology: (1) "Environmental tracers are natural and anthropogenic (manmade) chemical and isotopic substances that can be measured in ground water and used to understand hydrologic properties of aquifers. ... Different types of environmental tracers can provide different types of information about an aquifer."[45] (2) Groundwater temperature

[45] From http://water.usgs.gov/nrp/proj.bib/Publications/plummer.circ1222.pdf Retrieved on January 28, 2016.

surveys are valuable tools for identifying various interbedded clay layers.[46]

There may be more than one useful surrogate of a variable. In astronomy, there are various methods[47] to find exoplanets (extrasolar planets). Most of the time, planets around stars cannot be seen directly with telescopes. The *transit method* looks at the drop in the brightness of a star as a sign of a planet passing in front of the star. The *wobble method* on the other hand looks at the slight deviations of the radial velocity of the star as a sign of a planet orbiting it. Each method has its own weak and strong points with different rates of false positives for different types of star-planet pairs.

Above I mentioned the importance of the theory that gives the surrogate correlation. Naturally, the success of the surrogate method depends on the success of this theory and the development of it can change the fortune of the surrogate.

**Low-level Aerial Survey.** Today aerial and satellite imagery (as well as remote sensing) are methods used in numerous sciences from ecology to archaeology. I came across an interesting publication *Low-level Aerial Survey Techniques: Report of an International Workshop Held 6–11 November 1979 Nairobi, Kenya* [ILCA, 1981][48] (*LAST* from now on) on the topic that shows the growing pains of this field. Chapters of the book (each of which contain a few reports on a topic) start with a cover page featuring a beautifully-drawn image some of which I reproduce in figure 4.9.

Since late 1960s low-level aerial survey techniques have used low-flying small aircraft to gather information on the natural resources. Different groups of people including wildlife biologists, cattle owners, and game wardens needed methods for collecting animal census data and habitat survey. In *LAST* we see the initial researchers in this area trying to negotiate the difficulties and unknowns of the subject. It is a rarity in history of science that we can witness the formation of a discipline in a single publication but that is exactly what we have here. The papers included discuss the equipments, field methods, analytical methods, data handling and processing, survey principles, the role of observers, biasses of different methods and so on.

We can see the above mentioned dual nature of surrogacy in these papers: (1) improving the theory relating the aerial data obtained to the actual animal population; (2) improving the related data gathering and processing techniques.

A caricature of the process is as follows: You observe animal herds from the plane and plug the number of animals observed in an equation to get the actual number of animals in the whole region. If only it was so easy. To begin with, how do you count animals from a fast flying plane? How do you record it? How do you train human observers or design capturing devices? How high does the plane fly? Which path does it take — a grid pattern or a spiral pattern or something else? How do the land formations and tree distributions factor in? How does the behaviour and daily/seasonal cycle of the particular animal species factor in? Is it possible to restrict the count to animals of particular size or age? Do you count animal

[46] See National Ground Water Association's information page http://www.ngwa.org/Fundamentals/studying/Pages/Groundwater-temperature's-measurement-and-significance.aspx. Retrieved on January 28, 2016.

[47] See https://en.wikipedia.org/wiki/Methods_of_detecting_exoplanets.

[48] The book is available from http://pdf.usaid.gov/pdf_docs/PNAAR331.pdf Accessed on February 1, 2016.

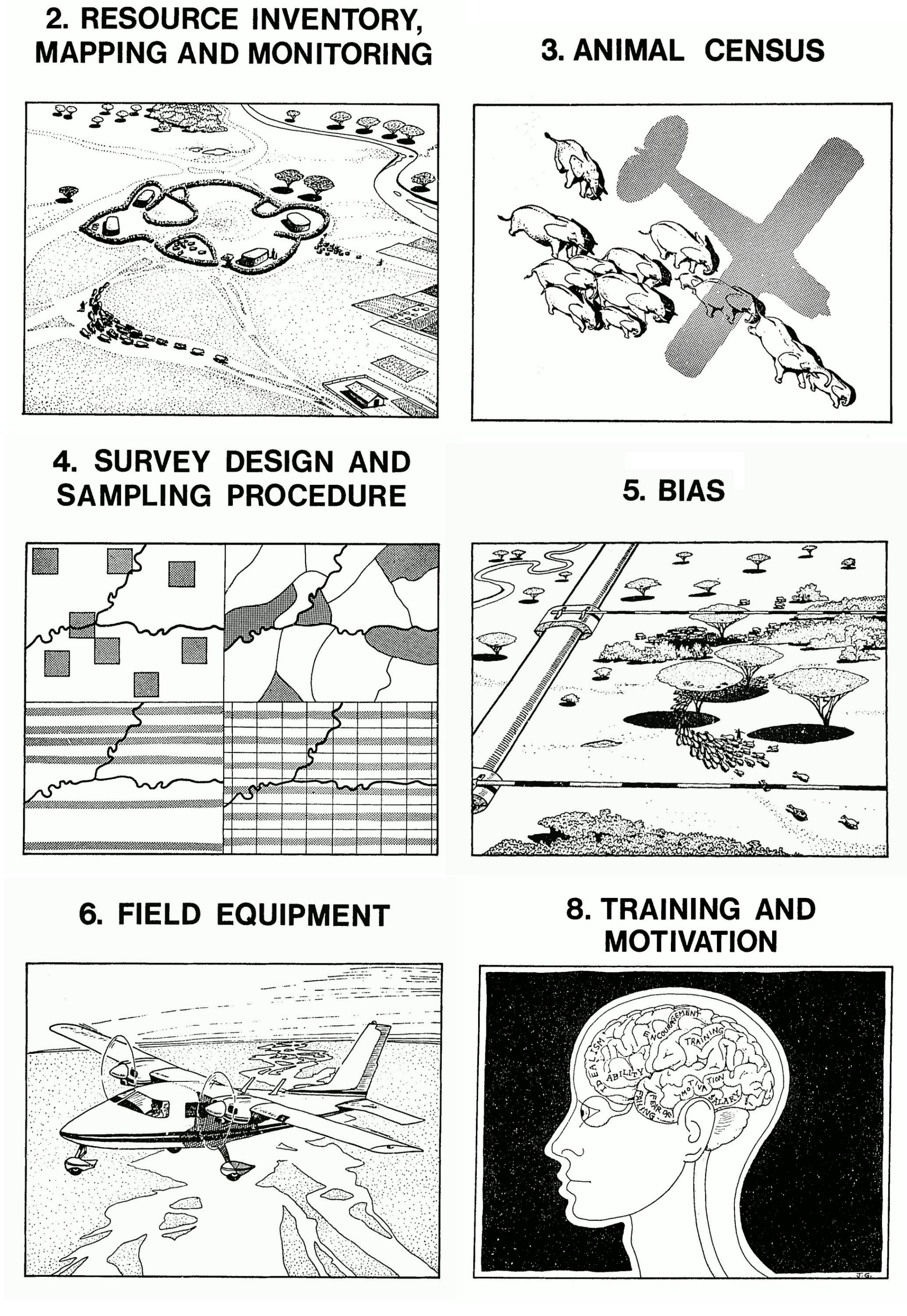



Figure 4.9: Images from selected parts of [ILCA, 1981] by an unidentified artist.

carcasses? What do you do with the collected data? How do you come up with an equation to put your data in? What statistical methods are suitable? Which sampling methods are available? Which one do you choose? These questions and more have to be thought of and worked on.

One of the important aspects of the work is to identify and mitigate biases involved as there can be various biasses at any stage of the aerial survey. Considerable number of papers in *LAST* discuss the possible biases and their potential elimination. Sometimes there is disagreement among researchers about the importance and/or method of elimination of a bias. For example, there is no agreement on whether sampling should be stratified or not, and also whether sampling should be or random or systematic, and so on. *LAST* shows how important the elimination of biases in this field but of course, as in all science, there is no guarantee that all biases can be discovered and eliminated.

If you recall my classification of biases at the end of section 2.13, papers in *LAST* are naturally of the third type: the way to diminish the bias is unknown but actively searched for. But of course there could be unknown biases in this research field.

**Surrogates in medicine.** In section 4.2, I mentioned how Claude Bernard suggested in 1860s a more accurate but hard to measure surrogate. The discussion about surrogates have not stopped there and the role of surrogates in current medicine cannot be overstated.

> Often, drugs are approved despite showing no benefit at all on real-world outcomes, such as heart attacks or death: instead, they are approved for showing a benefit on 'surrogate outcomes', such as a blood test, that is only weakly or theoretically associated with the real suffering and death that we are trying to avoid.
>
> This is best understood with an example. Statins are drugs that lower cholesterol, but you don't take them because you want to change your cholesterol figures on a blood test print-out: you take them because you want to lower your risk of having a heart attack, or dying. Heart attack and death are the real outcomes of interest here, and cholesterol is just a surrogate for those, a process outcome, something that we hope is associated with the real outcome, but it might not be, either not at all, or perhaps not very well. [Goldacre, 2013, 133–134]

If surrogates "only weakly or theoretically associated" with the real outcomes, then why are they used? Goldacre explains the pros and cons as follows:

> Often there is a fair reason for using a surrogate outcome, not as your only indicator, but at least for some of the data. People take a long time to die (it's one of the great problems of research, if you can forgive the thought), so if you want an answer quickly, you can't wait around for them to have a heart attack and die. In these circumstances, a surrogate outcome like a blood test is a reasonable thing to measure, as an interim arrangement. But you still have to do long-term follow-up studies at some stage, to find out if your hunch about the surrogate outcome was right after all. Unfortunately, the incentives for companies — which are by far the largest funders of trials — are all focused on short-term gains, either

> to get their drug on the market as soon as possible, or to get results before the drug comes off patent, while it still belongs to them .
>
> This is a major problem for patients, because benefits on surrogate endpoints often don't translate into real-life benefits. In fact, the history of medicine is full of examples where quite the opposite was true. [Goldacre, 2013, 134]

Here is only one of the examples known:

> Probably the most dramatic and famous comes from the Cardiac Arrhythmia Suppression Trial (CAST), which tested three anti-arrhythmic drugs to see if they prevented sudden death in patients who were at higher risk because they had a certain kind of abnormal heart rhythm. The drugs prevented these abnormal rhythms, so everyone thought they must be great: they were approved onto the market to prevent sudden death in patients with abnormal rhythms, and doctors felt pretty good about prescribing them. When a proper trial measuring death was conducted, everyone felt a bit embarrassed: the drugs increased the risk of death to such a huge extent that the trial had to be stopped early. We had been cheerfully handing out tablets that killed people (it's been estimated that well over a hundred thousand people died as a result). [Goldacre, 2013, 134–135]

As one other source puts it: "There had been spectacular failures of potential surrogates in medicine." [Lane and Barton, 2015, 162] Using surrogate outcomes is an important and necessary method that can lead to new insights. However, we must be watchful as it is too easy to generalize their results beyond the data. Unfortunately, in our time of "accelerated approval", in some cases surrogate outcomes do seem to have replaced real outcomes to the detriment of patients. For details see [Goldacre, 2013].

**Anecdotes.** The next topic I want to examine in this section are *anecdotes*. In the online dictionary oxfordlearnersdictionaries.com, one example sentence given for *anecdote* is "This research is based on anecdote, not fact". This usage is typical of scientists that denigrate a piece of evidence as anecdotal. What they mean is that information is not obtained in the right scientific setting. There are procedures (of course different in different sciences) of obtaining trustworthy evidence. Anecdotes break that procedures and therefore are suspect at best. Figure 4.10 is a brilliant parody of anecdotes.

Nevertheless, an anecdote is a free or low-cost information. It can potentially increase your evidence pool. Anecdotes bring the tension between quantity and quality to foreground.

Data dredging discussed in section 4.3 shows what can go wrong with anecdotes. Any data sample has anecdotal features and mistaking these to represent all possible data can be erroneous. But as discussed in that section, data dredging can be useful as well.

The role of anecdotes in a research tradition can change in time. For example, twentieth century witnessed a substantial change in medicine in this regard. In the beginning of the century there was the "art of medicine". New medical knowledge was obtained by good clinical judgements of physicians and researchers. But this anecdotal nature of medicine came under scrutiny later in the century. The most succinct history of this era that I could find is the following:

# Anecdotal evidence reliable? One man says "yes".

**A STUDY CONDUCTED YESTERDAY by a man on himself concluded that self-reported anecdotal evidence is, in fact, both reliable and relevant.**

The landmark study, conducted by Mark Mattingly of Virginia Beach in his apartment, concluded with 100% accuracy that data collected from personal experience can disprove other data conducted by reputable scientific institutions, thereby proving once and for all that "statistics can't be trusted".

In a press release Mr. Mattingly took aim at his detractors saying that "...this study shows what I've been telling people on the internet for years: all your fancy evidence and statistics don't mean nothing in the real world.".

A frequenter of internet forums, comment sections, and social media, Mr. Mattingly recounts that he was inspired to undertake the study when someone reportedly kept insisting that he provide evidence for his claims. "I think everyone's entitled to an opinion, and that my opinion is worth just as much as anyone else's" Mr. Mattingly said.

Academic types have criticised the study, and papers who are publishing it, saying that it lacks everything and makes no sense. When shown the study, Emeritus Professor James Albrecht of Carnegie Mellon University looked all confused and hopeless before making pining, guttural sounds.

Mr. Mattingly has responded saying that this is just the first of many studies he intends to conduct, and that a meta-analysis of people who have opinions and anecdotal experiences independent of controls, methodological rigor, blinding and peer review are soon to be published, adding further weight to his initial findings.

Figure 4.10: Published on Saturday 22, 2014 by yourlogicalfallacyis.com/anecdotal. Reformatted to one column and photo of "Mr. Mattingly in his apartment looking all smug" removed.


Figure 4.11: This Finnish stamp issued in 1975 hints at the statistical turn in medicine by means of the graphs in the background.

> Beginning in the late 1960s, several flaws became apparent in the traditional approach to medical decision-making. Alvan Feinstein's publication of *Clinical Judgment* in 1967 focused attention on the role of clinical reasoning and identified biases that can affect it. In 1972, Archie Cochrane published *Effectiveness and Efficiency*, which described the lack of controlled trials supporting many practices that had previously been assumed to be effective. In 1973, John Wennberg began to document wide variations in how physicians practiced. Through the 1980s, David M. Eddy described errors in clinical reasoning and gaps in evidence. In the mid 1980s, Alvin Feinstein, David Sackett and others published textbooks on clinical epidemiology, which translated epidemiological methods to physician decision making. Toward the end of the 1980s, a group at RAND showed that large proportions of procedures performed by physicians were considered inappropriate even by the standards of their own experts. These areas of research increased awareness of the weaknesses in medical decision making at the level of both individual patients and populations, and paved the way for the introduction of evidence-based methods.[49]

The result of this change is the currently accepted paradigm of Evidence-based medicine which gives much less importance to the "art" but rather stresses statistical methods, analysis of controlled trials, systematic reviewing and meta-analysis. (See figure 4.11.)

It is possible that anecdotes once dismissed as myth turn into objects of scientific study. One example is that of *transient lunar phenomena (TLP)* which are small apparent change on the surface of the Moon lasting only a short time. There are reports of TLP that go back to centuries but until very recently mainstream astronomy did not take these seriously. Why?

> The most significant problem that faces reports of transient lunar phenomena is that the vast majority of these were made either by a single observer or at a single location on Earth (or both). The multitude of reports for transient phenomena occurring at the same *place* on the Moon could be used as evidence supporting their existence. However, in the absence of eyewitness reports from multiple observers at multiple locations on Earth for the *same* event, these must be regarded with caution.[50]

In the last decade, astronomers stopped giving TLP the cold shoulder as recent technologies

[49] From https://en.wikipedia.org/wiki/Evidence-based_medicine. Accessed on February 13, 2015.

[50] From https://en.wikipedia.org/wiki/Transient_lunar_phenomenon. Accessed on February 13, 2015.


Figure 4.12: This stamp issued in 1932 by Newfoundland reflects the abundance of cod.

made TLP-research more viable. Today there is mainstream research on the topic though the jury still out on if and why TLP happen.

Sometimes science ignores anecdotes at its own peril. Newfoundland cod fishing industry collapsed in early 1990s because of overfishing. It turns out that fishery management in Canada was based on lopsided models. There were a number of unknown variables that made the scientific models biased in favour of continued fishing. But to the surprise of scientists, the cod population came to brink of extinction in a short time and the Canadian government had to announce a moratorium on cod fishing in 1992. (By the way, scientists also grossly underestimated the time needed for the cod population to recover.) This episode has led fishery science to become more reserved in its claims and look for ways to overcome the biases involved. What is interesting for us is that local fishermen were well aware of the problem but scientists did not take these "anecdotes" seriously. Local fishermen had a refined understanding of types of cod and were more attuned to signs of life-cycle changes than the scientific models did.[51]

It is also possible for anecdotes to be controversial in a science. As a case in point, there is a current debate about anecdotes and anthropomorphism in ethology. Do these concepts have a place in the study of animal behaviour? If yes, how and how much? If not, why not? There are eminent ethologists on the both sides of the controversy which does not seem to abate. We will see what role anecdotes are given when the dust settles. [See for example Mitchell et al., 1997]

To conclude, most of the time there is a price paid for quality and there is a place for other types of data in science.

[51]See [Bavington, 2009, 2011] for a discussion of the Newfoundland episode and [Kurlansky, 1998] for a general history of the cod fishing.

# CHAPTER 5

# PHILOSOPHY OF SCIENCE

Consider the question: *Would an alien from another galaxy like our science?* What kind of a question is this? Does it belong to philosophy of science? I do not think so. Granted, if you try hard enough, you can embed this question in some kind of philosophical context and turn it into a philosophy of science question (see figure 5.1); but, as it stands (here and today), it is not one. Why does it fail to be one?

- The alien question is not a suggestion to make our science likable to aliens. It does not propose a new ideal, goal, or policy we should try to incorporate into science.
- It is completely unrelated to the work of science. It does not connect to anything that could reflect in the practice of science in any way.
- The question is not related to any values practiced in science today. Science has some aesthetic values but none related to aliens.
- This is not a historical/sociological question as we do not know of any aliens that know of our science.
- It does not connect to other philosophy of science questions. It has no bearing on general philosophy of science or the philosophy of particular sciences.

Hence I conclude that the question is just a curiosity, not a genuine philosophy of science question.

Of course, today's curiosities can become tomorrow's serious business. Our environment, society, interests, and science can change in the future rendering this question a legitimate philosophy of science question. But doing philosophy of science *today*, we need to leave such questions aside as just curiosities.

In my opinion, there are some debates in philosophy of science circles that are a bit like the alien question. (1) Some of these questions are not philosophy of science questions. Like the alien question, they are mere curiosities. (2) Some questions resemble the alien question not necessarily in the sense that they are not genuine philosophy of science questions; but rather, they are discussed without taking considerations of the kind I listed above.

If one of the main aims of the philosophy of science is to understand science, then philosophers have to look at the practices, values, and history of science. Fiction, by definition, deals

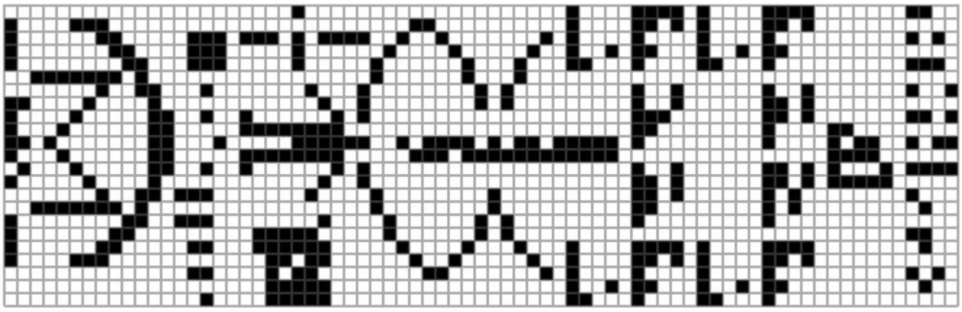


Figure 5.1: The message sent from the Arecibo radio telescope in Puerto Rico on 16 November 1974 towards the globular cluster M13 consisted of 1679 bits of information shown here graphically. The message contained various mathematical, chemical, biological, and astronomical information (see https://en.wikipedia.org/wiki/Arecibo_message for more information). As the wikipedia page states, "Because it will take 25,000 years for the message to reach its intended destination (and an additional 25,000 years for any reply),the Arecibo message was more a demonstration of human technological achievement than a real attempt to enter into a conversation with extraterrestrials." Nevertheless, this is a possible connection between our science and any aliens out there, though I doubt that it is a philosophically interesting one. [Image adapted from https://commons.wikimedia.org/wiki/File:Arecibo_message_bw.svg by Arne Nordmann (norro) and edited by Pengo and AnonMoos licensed under Creative Commons Attribution-Share Alike 3.0 Unported.]

with non-actual events and practices. A discussion that does not involve itself with the actual workings of science should not be called philosophy of science but rather science fiction.

In this chapter, I will look at some philosophical debates and show how one can clarify the discussion by looking at practices and values of science. One running theme of this chapter is the unfortunate tendency of philosophers to treat all fields of science as a unified whole and be indifferent to varied practices, values, and goals of these fields. The sections of this chapter deal with important themes in philosophy of science but I also deal with one less important issue in appendix D.

## 5.1 The Aim of Science

> By observing patterns in nature, by imagining the possible processes that could explain their origin and persistence, by doing experiments or making more observations that cast doubt on some of these possibilities and affirm others, we arrive at an ever changing and ever better understanding of how and why nature is as we see it. Research is intellectually and at times emotionally exciting. It is a creative, risky, difficult, totally consuming and immensely rewarding activity.
>
> From *The Smithsonian Tropical Research Institution* website

The list of philosophers who have written about *the aim of science* would be a *Who's Who* of philosophy of science. This is an interesting question on its own, but it has also been discussed as a part of other issues. Since the aim of an enterprise has much to guide it, it is a very important question. So, what is the aim of science?

**How can science have aims?** At the first sight, the aim of science question does not make any sense. It is *agents* that have aims. You can ask about aims of scientists or scientific institutions, but how can science in the abstract can have an aim? What do philosophers mean when they talk about the aim of science?

Karl Popper has a paper titled *The Aim of Science* [1979] in which he starts by clarifying how one can talk about the aim of science:

> To speak of "the aim" of scientific activity may perhaps sound a little naïve; for clearly, different scientists have different aims, and science itself (whatever that may mean) has no aims. I admit all this. And yet it seems that when we speak of science we do feel, more or less clearly, that there is something characteristic of scientific activity; and since scientific activity looks pretty much like a rational activity, and since a rational activity must have some aim, the attempt to describe the aim of science may not be entirely futile. [Popper, 1979, 191]

According to Popper, the aim of science is related to "something characteristic of scientific activity" — which is what exactly? For example, science is characteristically made by creatures with two ears. Can we conclude that the aim of science is having two ears? Obviously not. Since Popper does not say characteristic *what*, I am free to choose ears or whatnot.

I can think of only one interpretation of Popper that makes some sense: *There is some aim common to almost all scientific activity and that is what we call the aim of science.* The answer of *characteristic what?* is *characteristic aim*, so we are not free to choose ears. Clearly, there is a way to answer this question: look at a wide range of scientific activity and see if they share a similar aim. If yes, then that is the aim of science. Otherwise, there is no unique aim of science. As simple as this method sounds, the issue is complicated by the levels of aim as I shall explain now by using the example of the game of chess. We need some philosophical ammunition to deal with the aim of science question, and this digression about chess will help us develop them.

**What is the aim of chess?** Adopting J. D. Bernal's ideas which I will quote below, we can say that chess as an occupation may be considered to have three aims which are not mutually exclusive: the entertainment of the player and satisfaction of his naive curiosity of chess, the discovery and integrated understanding of chess, and the application of such understanding to the problems of human welfare. Let us call them the *psychological, rational, and social aims of chess*. These aims are easily observed in practice. Chess is found to be entertaining by a huge community. There are innumerable number of chess books and videos to educate, entertain, and improve the understanding of the game. The social side of chess is well established in the form of chess clubs, tournaments, web communities and so on. It has a number of beneficial traits to individuals and to society, and that is why we see chess programs in schools and prisons. It is very common in chess literature to read that one player aims for truth in chess. For example: "Chess for [Dutch chess Grandmaster] Timman is very much a search for truth — one gets the impression that he believes that top grandmasters are capable of playing near-perfect chess." [Burgess et al., 1998, 408] This fits well into rational aims. Does this trio of aims settle the initial question about chess? Not really. The trio of psychological, rational, and social aims of chess are examples of what I call the *weak aims*. These need to to be compared to the *strong aims*.

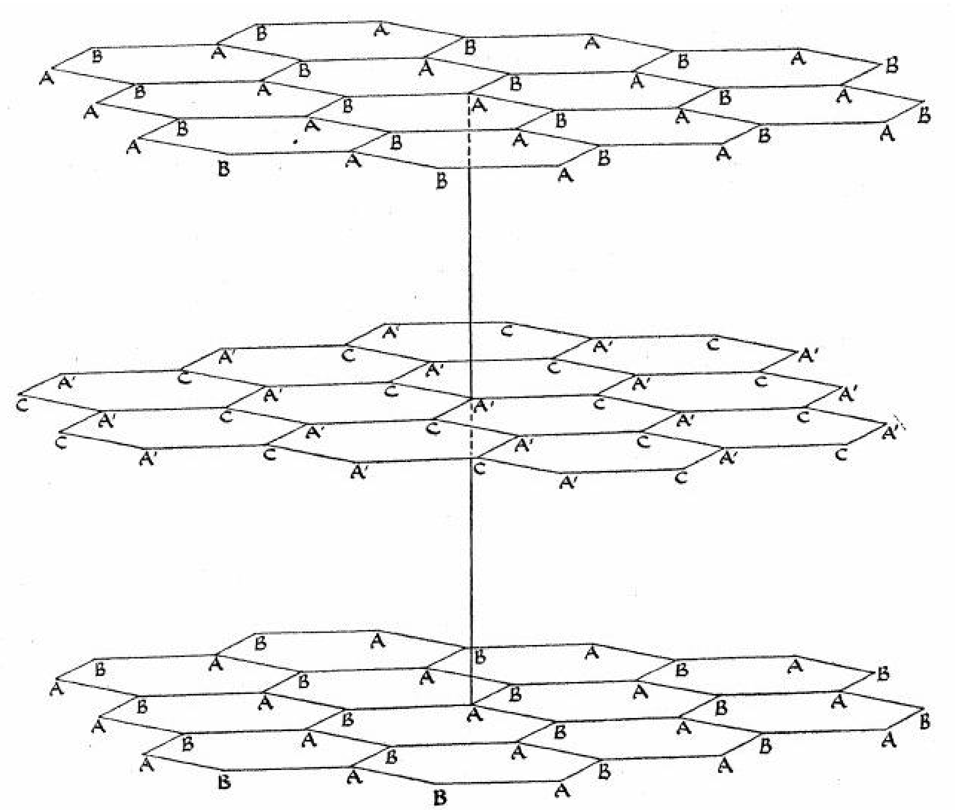


Figure 5.2: The structure of graphite. John Desmond Bernal (1901–1971) was one of the most influential British scientists of his time. From the publicity information of a book [Brown, 2006] on Bernal, we learn that "J. D. Bernal, known as *Sage*, was an extraordinary man and multifaceted character. A scientist of dazzling intellectual ability and a leading figure in the development of X-ray crystallography, he was a polymath, a fervent Marxist, and much admired worldwide. Although he himself never won a Nobel Prize, several of his distinguished students went on to do so, including Dorothy Hodgkin, Max Perutz, and Aaron Klug. . . . Bernal not only changed the course of science, but was witness to (and often a participant in) historical events (the Easter Rebellion, the Great Strike, the anti-fascist movement and pacific causes, civil defense, RAF bombing strategy, the planning for D-Day, post-war rebuilding, and nuclear weapons.) One of the few men familiar with Downing Street, the White House and the Kremlin, he left fascinating accounts of Churchill, Stalin, Mao Zedong, Louis Mountbatten and Picasso, as well as the century's greatest scientists." One of the scientific contributions of Bernal was the discovery of the structure of graphite [Bernal, 1924, 767] from which the above image is taken. Bernal also wrote extensively on history and nature of science, in particular the four volume text *Science in History* (1954).

The most important strong aim in chess is to mate the opponent. I will use strong-a for this ultimate aim and strong-b for the rest of strong aims. The chess game ends with mate when one side captures the king of the opponent. A player can resign before she is mated but this is because she understands that mate is unavoidable. Strong-b aims are the tactics and strategies employed in the game. For example, a player can try to gain material advantage, control a key square, improve weak pieces, establish open lines for her pieces, and so on. She may fail or be successful in such individual aims. If her strong-b aims lead to enough advantage over her opponent, she might achieve the strong-a aim and mate her opponent.

The difference between weak and strong aims are easy to observe but hard to define. In a sense, strong aims are lower level than weak aims. Weak aims are more about motivations and benefits, which are more likely to be shared with other games or behaviours. Strong aims are about how the game is played. One of the reasons I differentiate between these two types of aims is to emphasize that when we ask *what the aim of chess is* unqualifiedly, the answer should give strong aims, not weak aims. There are contexts in which a weak answer to such a question makes sense: If I ask you what is the aim of this chess game and you answer "to raise money to charity", that is a valid weak answer. But that question is not really about the game of chess but rather about an event. When the context does not shift the attention from the game, the aim of the game of chess is ultimately to mate the opponent and use tactics and strategy to achieve this. Since weak aims are shared among numerous other games and it is the strong aims that determine a game of chess, the latter type of aims are solely the ones we should consider as the aim of chess.

**Weak and strong aims in science.** After this long digression about chess, let us look at three aims of science introduced by J. D. Bernal (see figure 5.2) as follows:

> Science as an occupation may be considered to have three aims which are not mutually exclusive: the entertainment of the scientist and satisfaction of his naive curiosity, the discovery and integrated understanding of the external world, and the application of such understanding to the problems of human welfare. We may call them the psychological, rational, and social aims of science. [Bernal, 1939, 94]

It will be noticed that I wrote the above sentences about the aims of chess by changing Bernal's sentences *mutatis mutandis*. We can see the weak/strong distinction in science as well. What Bernal talks about are weak aims. His list is a good start but there are quite a number of aims that he leaves out. For example scientists aim to satisfy their egos, impress others, earn money, enjoy recognition, and so on. These should be added to his list of psychological aims. His two other categories also have glaring omissions but I will not bother to fill them in here. It is enough to have a rough idea about weak aims because what really matters as far as we are concerned are strong aims. Let me exemplify strong aims in science. (See also figure 5.3.)

The main strong aim of the string theory is to give a unified theory of all physical forces. Thus we may call this its strong-a aim. But in their day to day work, string theorists actually try to solve some minor theoretical problems which we can call strong-b aims. If enough of these aims are achieved, one day the ultimate aim of unification might be reached.

The strong-a aim of various "mercenary scientists" today is to discredit global warming. One of their strong-b aims can be to find alternative explanations of a particular climate data

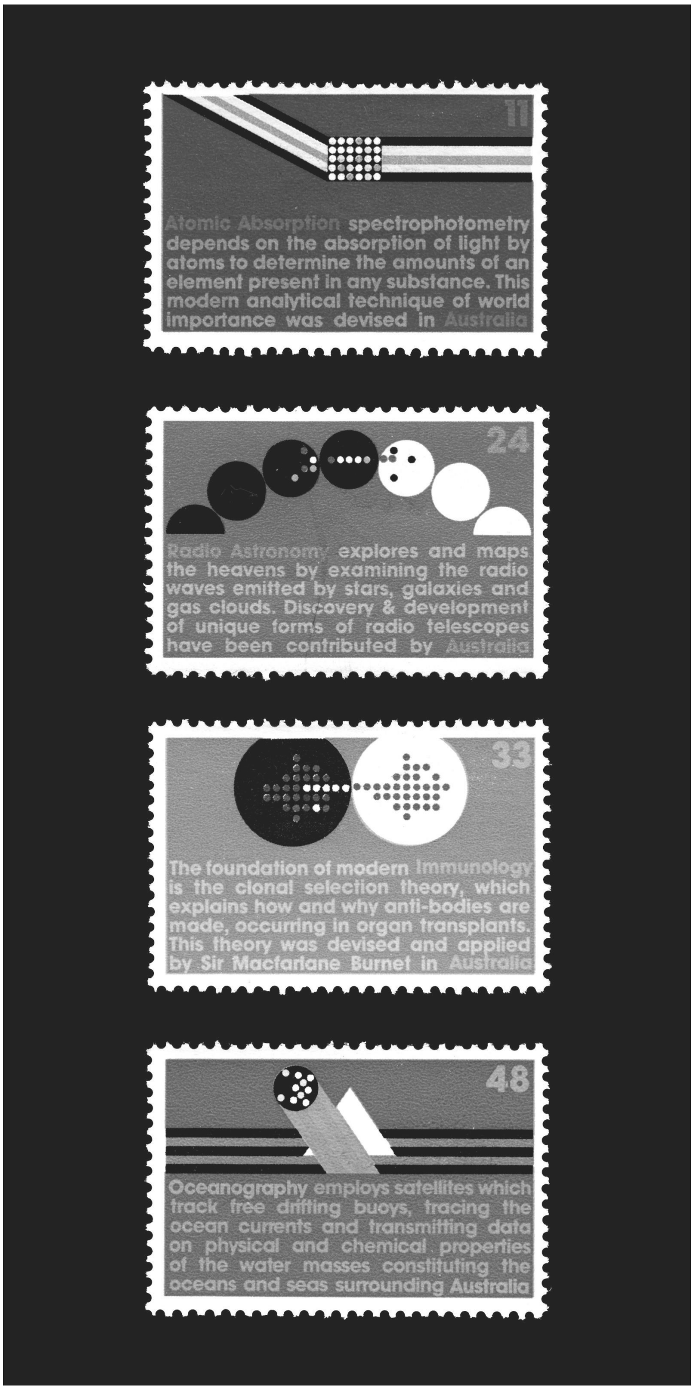



Figure 5.3: Australia issued a set of four stamps in 1975 which feature sciences that Australians contributed to. These are the only stamps that I know of which have a full paragraph on them explaining the methods and the aims of a science. Not surprisingly, weak aims are not mentioned.


Figure 5.4: World Meteorological Organization themed stamp issued by Switzerland in 1960.

that does not involve global warming. [See Oreskes and Conway, 2011]

The main aim of pharmaceutics is to improve the health of human beings by finding safe and effective drugs. This strong-a aim is not one that can be completely reached since (1) there will be always a possibility for improvement and (2) species change in time. In their daily work scientists do mundane things like understanding the properties/effects of a chemical or carrying out trials, and these are examples of strong-b aims.

Just as in chess, when we are asked the aim of a scientific discipline, the answer should be strong, not weak aims. It is safe to say that almost all scientists aim to earn money. Since this is a weak aim, we cannot say that the aim of meteorology is to earn money and we do not see money on meteorology themed stamps (see figure 5.4). The right answer is to study and understand the atmosphere. This is the main (strong-a) aim of meteorology. Individual research in meteorology has other strong-b aims and these can be investigated and found out.

In section 1.1, I wrote about different ways of investigating science and how different methods can improve our philosophical understanding of science. Until now, I advocated value analysis. There is another method that is essential for philosophy of science and it complements value analysis: *aim analysis*. Aim analysis finds out about the aim of scientific theories, researches, projects and so on. Since I equated aims with the strong ones, these are the main topics of aim analysis. The goal is to find strong aims, see if there are any to qualify as strong-a and how strong-b aims relate to strong-a aims.

The aim analysis can be quite easy sometimes. One can look at research grant applications or published work to see the aims. For example, one systematic review [Walsh et al., 2015] was made "To assess the effects of chlorhexidine-containing oral products (toothpastes, mouthrinses, varnishes, gels, gums and sprays) on the prevention of dental caries in children and adolescents." (See figure 5.5.) The authors clearly state the background, study characteristics, aim, and conclusion in their review. For the sake of completeness, here is their conclusion:

> We found little evidence from the eight trials on varnishes and gels included in this review to either support or refute the assertion that chlorhexidine is more

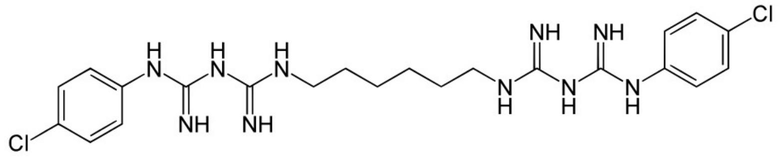


Figure 5.5: Chlorhexidine, which is an antibacterial in the World Health Organization's List of Essential Medicines, is used as an antiseptic and also in cosmetic and pharmaceutical products. [Image in the public domain from wikimedia.org.]

> effective than placebo or no treatment in the prevention of caries or the reduction of mutans streptococci [a bacteria that causes tooth-decay] levels in children and adolescents. There were no trials on other products containing chlorhexidine such as sprays, toothpastes, chewing gums or mouthrinses. Further high quality research is required, in particular evaluating the effects on both the primary and permanent dentition and using other chlorhexidine-containing oral products.

This example shows how easy it can be to find strong-b aims of research. But life is not always that easy, and finding out aims of a scientific discipline may require interviewing scientists, inspecting their practices, examining its history, and so on.[52]

Note that it is one thing to have an aim, it is another thing to have effective methods to reach those aims. This distinction will be important in section 5.3.

Consider for example "understanding the nature" — is this a weak or strong aim? It is a weak aim as it is too general and far removed from research. One can make the objection that, for example, the aim of understanding the effects of chlorhexidine falls under the aim of understanding the nature and so if one is a strong aim so must be the other. Not true. A researcher trying to understand the effects of chlorhexidine realizes the more general aim of understanding the nature *just because* he is trying to understand the effects of chlorhexidine. All other aims that fall under "understanding the nature" is completely neglected by that researcher, e.g. the aim of understanding the nature of space-time does not guide his chlorhexidine research. The aim of understanding the nature is realized in this case only by virtue of the aim of understanding the effects of chlorhexidine. Since only a single strong aim that falls under "understanding the nature" applies to the case and "understanding the nature" applies equally well to almost all scientific activity, "understanding the nature" is a weak aim.

I have mentioned in section 3.1 that the important values are the concrete ones: it is those values that affect the appraisal of theories and what scientists look at. Whether they know or care about (say) parsimony of theories of unrelated branches of science has no bearing on the parsimony of the theory in front of them. For example, to judge the parsimony of the life on Mars theory (section 2.6), a scientist need not look at the parsimony in the theory of mind issue (section 2.9) or the concept of parsimony in general. When it comes to aims, there is similarly a difference in importance of the type of aims.

The most important aims when it comes to determining the work of scientists are strong-b aims. It is those aims that determine what scientists do in their daily jobs. If the strong-b

[52] There may be even cases where the aim analysis may fail to find out the aims in a convincing matter or to distinguish strong aims from the weak ones.

aim in front of them is to determine the effectiveness of chlorhexidine, then that is what they do. Strong-a aims are more like rough guides. Among other things, they restrict the strong-b aims one can choose. For example, there is no place of the chlorhexidine study mentioned above in string theory. Weak aims are even more removed from the details of scientific work, and, most of the time, one can consider them to be negligible when discussing the work and practice of scientists. So the order of importance of aims for the philosophers of science are: First strong-b aims, then strong-a aims, and last, weak aims.

**Values as aims** Sometimes, (1) a value can become an aim, and (2) an aim can lead to a value. Let us see how.

(1) Values are not only important in theory evaluation, but also in theory construction. Scientists most of the time need to come up with theories that have particular values. They do not need any arbitrary theory, but one with specific values. Then obtaining that value becomes a strong aim. For example, as we have seen in section 2.5, one of John von Neumann's aims in constructing his quantum theory was to get a mathematically rigorous theory. In this case, the value of mathematical rigour has become an aim.

(2) Suppose that a group of scientists have an aim. It is only natural that they prefer theories that serve that aim and any such theory becomes valuable. Consider the following example:

> The fundamental problem in the design of a siphoning rain gauge is to ensure that the siphon empties the chamber quickly at some definite water level, without any dribbling; after 1875 there were several ingenious solutions. In 1886 the firm of Richard Frères submitted to G. J. Symons a siphoning rain gauge, but Symons found that the siphon dribbled, and returned it. The resourceful Jules Richard sent it back again, "cured". The cure was an electromagnet which, when 0.4 inch of rain had fallen, gave a strong push to the float, quickly starting the flow. Of course the requirement for a battery militated against the success of this idea. A simpler scheme, used a great deal in Germany, was the recording gauge devised by Gustav Hellmann and made by the Berlin firm of Fuess, which had a siphon with one rather long leg, made of a narrow glass tube. [Middleton, 1969, 148]

If your aim is to make a siphoning rain gauge which "empties the chamber quickly at some definite water level, without any dribbling" then obviously any design that achieves this better than another will be preferred. The aim immediately leads to a value. But note that there are other values of such a device: reliability, assembly price, etc.

We can conclude that there are important connections between aims and values and both of aim and value analysis are necessary to understand a particular scientific practice.

**What is the aim of science?** At last we have enough tools to tackle this question. Recall that I defined *the aim of science* as *the aim common to almost all scientific activity*. The use of the definite article "the" is apt if there is a *unique* aim common to all scientific activity.

One thing immediately clear is that there are numerous weak aims that are common to scientific activity. Scientists do what they do because of all kinds of weak aims: they enjoy their work, earn money, are entertained, satisfy their curiosity, satisfy their egos, discover new things, understand the world, explain stuff, reduce uncertainties, improve reliability of methods and theories, unify different branches of science, seek truth, improve human welfare, impress others, enjoy recognition, like being in a scientific community, and so on. Therefore we can answer the question *what is the weak aim of science?* — There is no unique weak aim of science but rather multitude of them.

What about strong aims? Is there any among them that deserves the title of *the strong aim of science?* We have seen above a few examples of strong aims from different fields. They differ so much among different fields of science that we cannot call them similar. In fact, one of the main reasons why scientists in unrelated traditions do different work is a result of their dissimilar strong aims. To put it another way, there are no strong aims common to different fields of science as this is what we observe and it is also defining: different strong aims are what makes them different fields to begin with.[53]

The definite article "the" is misplaced in the phrase "the aim of science" since it suggests a single aim of science. When we are talking about weak aims, there are more than one, in fact numerous, aims of science. When we are talking about strong aims, there is none that is common to all science, hence no strong aim of all science exists. Since the important aims in science are strong aims, I would say that science has no common aim at all (unless specifically asked about weak aims).

So what about philosophers talking about *the* aim of science? They are committing a sin I wrote about in the introduction to this chapter: treating all fields of science as a unified whole without actually looking at scientific practices, values, and goals. Science is so rich and varied in its goals that it is surprising how anyone can talk about "the aim of science".

**Plasticity of weak aims** For any given weak aim, there are myriad scientists, scientific communities, and scientific activity influenced by that aim. Consider any weak aim, e.g. "giving satisfactory explanations", "developing empirically-adequate theories", or "improving human-welfare". There are no doubt all kinds of theories that serve any such aim. For example, giving satisfactory explanations is important in both molecular chemistry and social anthropology. So there is a one-many relationship between weak aims and theories that share in them.

There is also a many-one relationship that is more important philosophically. Weak aims are *plastic* in the following sense: given any scientific theory or result, there are numerous weak aims that *could have* led to that theory or result. This is a counterfactual conditional which implies that if a weak aim (or a collection of them) has led to a theory, then you must not assume that only that weak aim could have lead to that theory. In fact, there are many more which could have led to that theory.

Kip S. Thorne [1994, 178–187] tells the story of the brilliant Soviet physicist Lev Davidovich Landau (see figure 5.6) who felt the heat of Stalin's Soviet regime in 1930s and was in fear for his life. Thorne writes that "In panic he [Landau] searched for protection. One possible protection might be the focus of public attention on him as eminent scientist," and Landau tried to come up with a scientific idea that would make into the news. Landau managed to achieve this goal by his work on stellar energy, but he was nevertheless prisoned for a year.

[53] There are different traditions in science that compete with each other as they have the same strong-a aims. As far as I know, even these traditions have different strong-b aims. For example, traditions in physics that aim to give a unified theory of quantum gravity have different strong-b aims, e.g. background independence. I cannot think of any separate traditions in science with exactly the same set of strong aims. It is of course course logically possible to have competing traditions with exactly the same strong aims and what makes them different is something else than strong aims. But in practice, a different set of strong aims is a *sine qua non* of tradition differentiation.

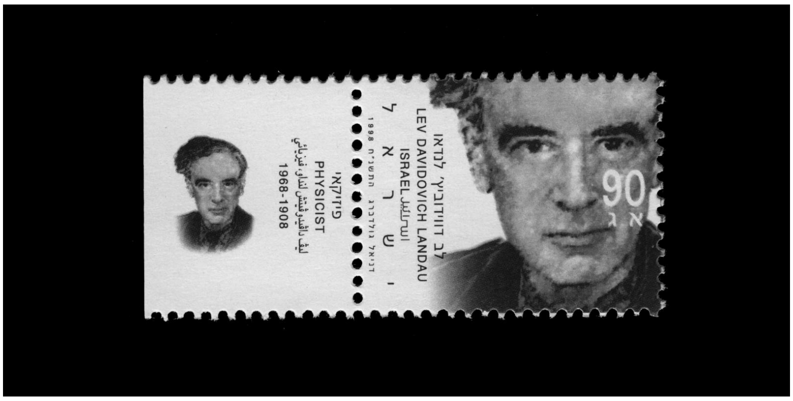



Figure 5.6: The prominent Soviet physicist Lev Davidovich Landau (1908–1968) on the stamp issued by Israel in 1998.

For our purposes, the details of Landau's work are not important as long as we grant that the strong aim of the work was to give an account of how stars generate energy and that one of the weak aims was to make news and avoid imprisonment.

Now clearly the work of Landau was shaped by his circumstances. But could Landau or another physicist have developed the same theory under different circumstances? Sure, why not? There is no fear of life requirement to become a good astrophysicist. Physics community did not take Landau's ideas as important because of his weak aims. While Landau was working under Stalin's shadow, his western colleagues were developing similar theories without the need to make news and avoid imprisonment. What they shared was not this weak aim, but the strong one.

Scientists or scientific communities come to similar theories with very different weak aims but cannot do so with different strong aims. To put it another way, weak aims are *plastic*, strong aims are not. This is why concentrating only on weak aims and taking them to be indispensable or defining feature of science or a particular scientific episode is mistaken.

**Two misguided approaches.** My discussion of the aim issue has consequences for two type of narratives.

(1) There are a number of philosophers who have made grand claims such as

- The aim of science is to give empirically-adequate theories. (Bas C. van Fraassen)
- The aim of science is to effectively solve problems. (Larry Laudan)
- The aim of science is to give satisfactory explanations. (Karl Popper)
- The aim of science is theoretical knowledge. (Ernan McMullin)
- The aim of science is to understand nature. (One of my teachers)
- The aim of science is to reduce bias. (My doppelgänger)

There are two problems with these claims. First, these approaches arbitrarily restrict aims to the weak ones.[54] The excessive importance given to weak aims is misplaced when it is the strong aims that guide scientific practices and determine the research topics.

Second, as we have seen in this section, there are numerous weak aims of science. The use of the definite article "the" in the above claims is simply false — one cannot use "the" unless there is a *unique* aim.[55]

What if somewhat all weak aims but one disappeared from the world? Would that remaining weak aim deserve the title of "the weak aim of science"? If all but one disappeared magically then of course the remaining would be the weak aim of science. To continue our story, what if in this scenario, the relevant linguistic community started calling this remaining weak aim "the aim of science"? Would it then become the aim of science? Only in name. You can call a lame duck whatever you want but this would not make it a dragon. To repeat the point made above again, weak aims are removed from the details of scientific work, and, most of the time, one can consider them to be negligible when discussing the work and practice of scientists. The order of importance of aims for the philosophers of science are: first strong-b aims, then strong-a aims, and last, weak aims.

In view of these points, a weak aim cannot be *the* aim of science.

(2) There are numerous historians of science who have made claims with regard to individual scientific episodes such as:

- Gender politics have affected Boyle's law ($PV = k$) as Boyle developed his gas law because of his political views about women. (Elizabeth Potter)
- Fluid mechanics is an underdeveloped area of physics because fluidity is associated with femininity which male scientists tend to avoid. (Luce Irigaray)

Even if we grant that gender politics and aversion to femininity are weak aims relatively pertinent to these cases, it does not follow that the science in question has been substantially determined by these aims. These views completely disregard strong aims and ignore the plasticity of weak aims.

**Serendipity.** Before I end this section, let me dispel the possible misunderstanding that all discoveries in science follow an aim. On the contrary, there are numerous unexpected discoveries. A scientist can come up with an accidental result that has nothing to with his original aims(s). Probably the most well-known accidental discovery in the history of science is Alexander Fleming's (see figure 5.7) discovery of penicillin.

> When I woke up just after dawn on September 28, 1928, I certainly didn't plan to revolutionise all medicine by discovering the world's first antibiotic, or bacteria killer, ... But I suppose that was exactly what I did.[56]

Royston M. Roberts [1989] collects a large number of accidental discoveries in science.

---

[54] I have explained above why "understanding nature" is a weak aim of science. By similar reasoning the other listed aims are also weak. They equally well apply to a wide range of sciences, do not determine the daily work of scientists, and they have no role in demarcating research traditions or branches of science.

[55] The uniqueness implications of definite descriptions are widely accepted though there might be problematic cases [see Ludlow, 2013]. But it is safe to say that a sentence of the form "The aim of science is to ..." implies that there is a unique aim of science.

[56] Fleming quoted in [Haven, 1994, 182].


Figure 5.7: A stamp issued by The Faroe Islands in 1983 depicting the discovery of penicillin by Alexander Fleming.

**Conclusion.** When someone asks you about the aim of a sport, game, or science, you should answer strongly (unless they are asking about a particular gathering). There are countless weak aims. There are numerous weak aims related to a single case. More importantly, weak aims are plastic. Strong aims are the ones that mainly determine science. Strong aims are not plastic. Strong aims help to separate different fields of study. Sciences, along with their aims, are so rich and diverse that only an aim analysis can find out aims of a research tradition. There is no "*the* aim of science" we can talk about.

## 5.2 Unbiasedness

Only the näive or dishonest claim that their own objectivity is a sufficient safeguard.

E. Bright Wilson, Jr. [1952, 44]

Objectivity has been an important issue in philosophy of science, and even more generally, in philosophy of knowledge. I suggest a better philosophical and scientific alternative notion over objectivity, namely unbiasedness.

**Objectivity** Objectivity is a concept that philosophers should abandon for a couple of reasons.

First, there are so many different conceptions and definitions of objectivity that it is impossible to make a principled discussion about it. I will not even bother to give some of these definitions here (see [Megill, 1994; Gaukroger, 2012; Daston and Galison, 2010; Resnik, 2006] for partial lists). There is no doubt about the fact that objectivity is a confused and unclear concept.

Second, objectivity is enforced upon science by some philosophers as a necessity criterion. I object that such a confused concept should be imposed upon science. I object that such philosophers have such a low opinion of science that they have a need to save it by a *deus ex machina* criterion. I object that instead of trying to understand real science, they resort to a fake construct. I will turn to this misunderstanding of science in section 5.5.

The death of objectivity is long overdue. What needs to replace it is *unbiasedness* which I will turn to now. But before that let me mention that even if you define 'objective' as being unbiased, you should still refrain from using the word 'objective' in philosophical contexts and stick to 'unbiased' since the latter does not have the connotations of the former.

I will structure this section around four questions.

**Is science unbiased?** The answer to this first question is straightforward. It is safe to assume that all branches of science are biased for at least following reasons:

(1) The first point is an observation: we see all kinds of bias in science, from gender bias to publication bias, from sampling bias to theoretical bias, science has it all. Science is full of bias but some fields like medicine are more prone to bias than other fields like physics.[57] How much and what kind of bias a field has is an empirical question. I have given one example of systemic bias in a branch of science in section 2.13.

(2) The second reason is a historical one: In the last 40 years, we learned a lot about biases and different types of them. The number of cognitive biases discovered since 1970s is astounding.[58] There is no reason to think that we know about all biases and probably, we have biases that we do not yet know about. Science can have biases we have no idea of.

(3) There are cases so complex that bias is unavoidable. "Bias is an inescapable element of research, especially in fields such as biomedicine that strive to isolate cause-effect relations in complex systems in which relevant variables and phenomena can never be fully identified or characterized." [Sarewitz, 2012] The cycling helmet issue I looked at in section 2.8 is an example. One needs to restrict the variables taken into consideration by importance and this introduces biases in the analysis.

(4) Scientists sometimes have to choose one of several competing theories to continue their own work even if it is premature to conclusively choose one of the theories over alternatives. They do not have the luxury to wait and see the eventual winner. In such situations, their choices can be biased reflecting their own interests or preferences. A scientist might not even see the vindication or rejection of this theory in his lifetime.

(5) When do scientists stop data collecting, experimentation, and analysis to settle a question or problem? This depends on the importance given to the problem. If our lives depend on it, we make sure to get the results right. But a problem deemed unimportant might be settled much easily. The choice between two theories involves some bias related to the importance given to the outcome. What is important is subjective, and this bias is necessarily a part of theory choice.

In the light of these points, all branches of science are biased one way or another. If so, a couple of related questions immediately come to mind.

[57] For bias in physics see for example [Jeng, 2006; Catena, 2014].

[58] Daniel Kahneman and Amos Tversky were forerunners in discovering cognitive biases and how humans handle risk. See [Kahneman, 2012] for a popular introduction to this field.

Figure 5.8: *An Experiment on a Bird in the Air Pump* (1768) by Joseph Wright of Derby (1734–1797) depicts a recreation of one of Robert Boyle's air pump experiments by a natural philosopher. [Image in the public domain from http://commons.wikimedia.org]

**Do biases in science change in time?** The biases we have and our conception of biases have changed in time. Not only we learn about new biases all the time, but also the importance we give to different biases change as well. As our conception and knowledge of biases change, those we eliminate from science change along. Hence biases in science change in time.

As mentioned in section 4.3, one interesting example given by Shapin and Schaffer [1985] is the role of witnesses in the seventeenth century pneumatic experiments. Invitation of observers to witness the experiments, which is immortalized in Derby's celebrated painting *An Experiment on a Bird in the Air Pump* (see figure 5.8), became a hallmark of unbiasedness in this era.

Another example is the increasing prominence achieved by double-blind experiments in the last hundred years. Single-blind experiments were employed since nineteenth century for their role in avoiding observer's bias. But experimenter's bias was still present in single-blind experiments and hence the need for double-blind experiments that avoids both biases. See [Stolberg, 2006; Kaptchuk, 2011, 1998] for history of blind experiments. But double-blind experiments do not eliminate all bias and there is ample opportunity for further improvement (see appendix B).

**Is science getting less biased?** I believe that scientists are getting better equipped to eliminate bias. We are learning more and more about biases and developing methods to eliminate them. For example, single-blind and double-blind experiments have been an effective tool to reduce bias in medicine, psychology, and social sciences (but see appendix B). Another success story is systematic review. But having the capacity to eliminate bias is one thing, actually doing so is another. Essentially, this is an empirical question that needs to be answered for a particular field of science.

For example, lately the bias in pharmaceutics came to public attention in western countries, and there are some initiatives to change the status quo. If these succeed we might see any rampant bias fade away. But we have already had the means to drastically reduce this bias for a very long time, though not used.

There is an obvious problem with determining the amount of bias in science: we can only observe those biases we know about. But, clearly, a decrease in known biases is necessarily a decrease in all. It makes sense to think that the more we know about biases, the more they stick out. The question is whether they are indeed pruned.

My offhand belief is that in most fields, the biases are decreasing. But without the requisite investigation of different fields to buttress my idea, I will not follow it.

**Should science aim to be unbiased?** There are a few reasons why I believe that it is important for science to strive to be unbiased.

(1) As I have explained in section 5.1, the most important and defining aims in science are the strong aims as they determine the day to day work of scientists. Biases most of the time get in the way of reaching those aims. For example, if your aim is to find out the frequency of a trait in a population, then your sampling better be unbiased.

(2) One weak aim of different branches of science is to reduce uncertainties. This weak aim might be realized differently in different fields. For example, in medicine, it is important to know the relative merits and disadvantages of different treatments. Uncertainty about which treatment to follow can even be fatal. A number of biases can lead to mistaken opinions about relative merits. One example is *comparator bias* which is the bias of choosing non-suitable traits when comparing treatments. See [Mann and Djulbegovic, 2013].

(3) Bias makes science unreliable. The accepted theories, results, explanations, etc. can change if the winds of bias blow from another direction. Bias can not only make it harder to achieve scientific aims, it can also leave us in doubt about our results.

Consider the evidence-based medicine revolution in the last few decades and its methods like systematic reviews. This movement has shown how previous biased methods have led to unreliable and even harmful treatments.

(4) A related point is that bias leads to irreproducible experiments or clinical trials.

(5) Bias leads to waste of time and resources.

(6) Science is practiced in a society and significantly public funded. Public not only expects science to improve human welfare, but also counts on it to be unbiased. Science is seen as an arbiter and we see its influence in all kinds of forums, from weighing on public policy to expert scientific testimony in courts. If science is to serve the people, then science owes it to the public to be unbiased. (See figure 5.9.)

(7) Bias can reduce trust in science and undermine the public support of science.

To summarize, science is biased, probably will remain somewhat biased forever. This does not mean that we cannot reduce bias. More importantly, science should aim to be unbiased (and this is a weak aim I endorse) even if it can never reach this aim completely.


Figure 5.9: This stamp issued by France in 1941 shows a typical perception of science. Our saviour Scientia defends us from the evil seven-headed Hydra representing cancer.

**Unbiased philosophy.** Having discussed bias in science, let me turn now my attention to philosophy of science which I believe to be more biased than science. This branch of philosophy is likely affected by numerous biases that also affect various natural and social sciences. But there is a further reason for my belief which is the track record of philosophy of science.

The emphasis of the twentieth century analytic philosophy of science has been to formulate what science *should be* rather than to understand it. This endeavor has come up with fascinating fictional sciences some of which I will touch upon in section 5.5 and appendix D. Unfortunately, science as practised by real scientists cannot be found among their hypothesized ones. Their views about science are biased by their normative emphases.

Diametrically opposite to analytic philosophy is another biased philosophy, though in a different way. This view starts with the justified premise that science is practised by real people and scientific communities but ends up with the simplistic view that all matters in science are societal values. "Scientific facts" are said to be constructed by a process of negotiation or power struggle. All other aspects of science are (sometimes acknowledged but) neglected and the history of science is reconstructed as a series of isolated episodes detailing how the work of a scientist was shaped by social and cultural circumstances of the era.

Both approaches lead to an impoverished view of science. One neglects real features of science for fictional ones, the other construes science as an episode from the book/television series *Game of Thrones*.

Both approaches are inflicted by a number of biases, the most important ones being cherry picking and hasty generalization. With a commitment to one value over others, the writings of these philosophers do great injustice to the richness of science by picking only the case studies that exhibit their favourite value. But there is even a further injustice carried out by a small but not insignificant number of philosophers of these camps: taking a case study from history of science and carefully reconstructing the story so as to avoid all other values. This process might be unintentional, but the result would make anyone working in The Ministry of Truth of Orwell's *1984* proud. I will reconsider these two camps in section 5.5.

Philosophy of science is biased, and more importantly, I do believe that (like science) it will always remain biased. But my contention is that philosophers of science should aim to be unbiased even if it is impossible to achieve this aim completely. By time, we can learn about biases in our philosophy of science and we will understand science better by eliminating them.

Some of the biases can be easier to eliminate than others. To start with, the cherry picking I mentioned above can be eliminated simply by *stopping* cherry picking, that is, by choosing a wide range of examples from science that show the variety of values in science; that is why I picked numerous recent case studies in chapter 2. But of course this is just a start.

The favourable reconstructions of science mentioned above can be challenged by other reconstructions. In fact, this is something I look forward to happen to me as well: If other philosophers offer alternative interpretations of my case studies, then this would probably lead to an improved and less biased understanding of the cases.

One of the reasons why I chose recent case studies in chapter 2 is that, as the saying by the British novelist and short story writer L. P. Hartley goes, "The past is a foreign country: they do things differently there." It is much easier to immerse oneself into the culture and works of a contemporary scientific community than a past one. The required interpretation that goes into a historical case study is much more involved and more likely to lead to biased analysis. To build a repertoire of values in science, it is best to slowly get one's feet wet by looking at such recent episodes. The historical case studies bring along the additional problem about the possibility of an unbiased history.[59] (As you can guess by now, I believe that historians should aim to be unbiased even if they cannot achieve this aim, and they can make progress. But I will not pursue this thought further here.)

Be it science, philosophy or history, it is possible to reduce biases in time and we should aim to get rid of them.

## 5.3 Rationality

> As a rationalist I feel that the word "rational" is one which indicates a high element of desirability, and I think it is much broader in its meaning than "deductive". In fact what appears to me to be the rational approach is to take designs which are in use already, to see what is achieved by these designs by consideration of the general aims to evaluate such designs, in a provisional way, and then to seek to find designs which improve on existing practice. Having found such designs the cycle should be repeated again.
>
> G. A. Barnard [1959]

Rationality has been a topic of dispute since antiquity and especially in the twentieth century. On the one side there are those who claim science is a rational enterprise and this confers credence on science. On the other side are those that claim that all belief systems are on equal

[59]This is more traditionally known as *the objectivity of history problem* but as I have made clear I prefer to use bias-language rather than objectivity-language.

footing and science is only one of them. According to this view, science is no more rational than (say) astrology. As I will argue below, both sides are mistaken.

I have heard of a conception of rationality employed by some: what is rational is by definition what is scientific. Furthermore, they use rationality as a demarcation criterion of science from non-science. It is lost on them that there is a vicious circularity in this endeavour. You cannot define what is rational to be what is scientific and vice versa at the same time. At least one has to give way. Actually, I argue that both statements have to be rejected.

Rationality should not be bestowed upon disciplines or agents by definition. It should not be an a priori matter whether a discipline is rational. Equating science and rationality strips rationality of any meaningful philosophical power: it becomes just another name for science. Besides, common-sense concept of rationality cannot be captured by such a move. Furthermore, what would happen if some irrational tenets were discovered in a science, say physics?[60] Then by definition physics would disqualify from being a science. Surely, this is at the very least an unpleasant outcome: some philosophers denying physics the status of science while scientific community and lay people calling it science as always.

To repeat the points: (1) Rationality is not something to be bestowed upon science (or any agents/disciplines) by fiat. It is an empirical matter whether any endeavour is rational or not. (2) Rationality is not a demarcation criterion of science from non-science. It is neither necessary nor sufficient for any work to be rational in order to be deemed scientific.

In the opening paragraph I mentioned those that claim that all belief systems are on equal footing and science is not any more rational than astrology. Hence this views also denies that rationality is an empirical matter.

The only way to settle whether a field of science is rational is to inspect and analyse that field of science. It might turn out that all of the sciences are rational or none are. Or it might turn out that some are rational, some are not.

What needs to be done is to employ a definition of rationality that brings forward the empirical character of its applicability. This is where discussions of rationality become troublesome. Definitions of rationality lack clarity. Rationality, like objectivity, means different things to different people (see figure 5.10) and this makes the issue of rationality a complex one. Here I will consider only one explication of rationality which can be made clear and in which values have an important role.

**The definition and the myth of rationality.** *Rationality* means to follow the best (available) methods, theories, and practices to create or acquire knowledge.[61] The claim is that most of science is quite good in this respect. If a new method comes up that shows promise, scientists will try it out and evaluate it. If it fulfills its promise, then it will be absorbed by science and will become part of mainstream science. Just to give an example, the double blind experiments were found to be effective in different fields of science and now have become a staple (see

[60] Indeed, there are episodes in history of physics that beg for the irrationality label: the episode of N-rays, and the fraud of Jan Hendrik Schön come to mind among others. Sometimes, I think that the whole history of particle physics in twentieth century is like a crazy roller coaster ride. But none of these episodes should deny the label "science" from physics.

[61] This conception of rationality can apply equally well to scientists and to science itself. If a science incorporates the best methods, etc. then the science is rational. If scientists incorporate the best methods, etc. then they are rational.


Figure 5.10: A Finland stamp issued in 1974 showing us what the Finnish postal authority thinks *rationalisation* looks like.

figure 5.11). Scientific methods are always compared to others and those that compare well to others are retained. Scientists might be a bit hesitant to adopt the new theories or methods but this is a sign of prudence and eventually science will adopt these new things. It might take another generation for the adoption as Max Planck's [1949, 33–34] oft-repeated quote reminds us: "A new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it." Nevertheless, science will catch up with the effective methods and theories, continually improving itself. Thus science assimilates the best methods, theories, and practices and so it is rational.

I will refer to the ideas expressed in the above paragraph as *the rationality of science picture* and make three points about this picture.

**A unified rational science.** The first point is that there is no one unified "science" which can be said to be rational. Science is claimed to be rational without actually looking at the practices and methods of the specific sciences. We are given a definition of rationality and then told a pleasant story about how science is rational. The problem I am discussing now is not about the definition but rather about the story. As I mentioned above, rationality of science is not something one can decide in an a priori fashion. There is no way to tell whether one branch of science is rational or not without analysing the branch in question. Rationality is something that needs to be investigated, not assumed. The above rationality picture generously paints all science rational by fiat. Let us look at some of the cases I presented in chapter 2.

Consider the nuclear models introduced in section 2.7. At first sight, this branch of physics might seem to be irrational. There are all kinds of analogical models, such as taking nucleus to be a drop or shell, and more importantly, these models contradict each other and experiments as well. How can such an endeavour be anything but irrational? But this is a mistaken way

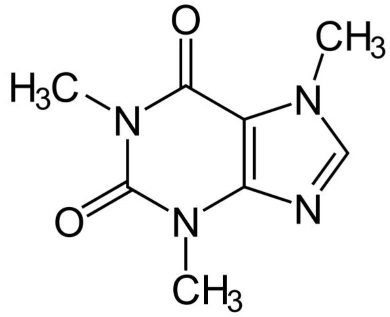


Figure 5.11: Caffeine molecule. One of the earliest uses of double-blind studies in the twentieth century was the investigation of the effects of caffeine by W. H. R. Rivers and H. N. Webber in their article “The action of caffeine on the capacity for muscular work” [1907]. They extol their methodology as follows: “In another important respect our work differs from all that previously recorded, in that on the normal or control days we have taken doses of a mixture which we were absolutely unable to distinguish from that containing the caffeine. In this way we have not only eliminated any possible influence of suggestion, but also the much more important sources of error arising from the sensory stimulation produced by swallowing the drug and from the interest due to the act of taking it.” Double-blind experiments turned out to be quite successful in eliminating or reducing a number of biases and became widely adopted in sciences. But double-blind experiments can have problems as well (see appendix B).

of thinking. The aim is to understand the nucleus. However analogical or contradictory the models are, they are the best of what is available now, and that is all that matters. Rationality is not about success, analogies, or contradictions; it is about finding the best and right tools for the job, improving them, and comparing them to new ones. Therefore what we have in this case is a rational enterprise.

What about the rationality of pharmaceutics? The aim of this research is to improve the health of human beings by finding safe and effective drugs. But as I have outlined in section 2.13, the current practices are a far cry from best possible methods to come up with such drugs. For example, changing how trials are carried out and published can lead to huge improvements. So can we conclude that pharmaceutics is irrational? Not necessarily. Let me turn to my second point about the above rationality picture to see why.

**Aims and rationality.** The definition of rationality as “following the best methods, theories, and practices to create or acquire knowledge” has a tacit assumption about the aims of the practice. What is best depends on what you are trying to do. Consider a tribe gathering around a fire to dance to fend off evil spirits. You should not jump the gun and call this an irrational act because there might be an ulterior motive of the dance, for example to bring together the community and forge closer relationships. It might me the case that such a dance is the best way of achieving community solidarity in that environment. To repeat the point, whether an endeavour is rational or not depends not only the methods, practices, etc., but also on the aims of the practice in question.

To return to the above example, if you take the aim of pharmaceutics to be “to improve the health of human beings by finding safe and effective drugs” then this research field is indeed an irrational one. Clearly, the best practices and methods to achieve this goal are not followed.

But this analysis forgets that pharmaceutics is a huge industry with the aim of maximizing their profits. They have huge incentives in making drugs that look *as if* they are safe and effective. One way of making drugs that look as if they are safe and effective is actually making drugs that are safe and effective. (That is why we all have these wonderful drugs today.) But if your aim is to improve your profits, then any method of making your drugs look safer and more effective than they really are is rational. The behaviour of the pharmaceutics industry is rational even though they do not follow the best practices and methods to improve the health of human beings by finding safe and effective drugs. They follow the best practices and methods as a business to improve their income and that is why they are rational. But one can make a strong case that the aim of the pharmaceutics as a science should be "to improve the health of human beings by finding safe and effective drugs". This makes pharmaceutics a rational business and an irrational science as it serves the aims of the former better than the latter.

I said above that I will make three points about *the rationality of science picture*. To recall, the first is to judge rationality of a science not by fiat but rather by investigation. This is a point that equally applies to all conceptions of rationality, not the particular one I employ. The second point was directed at the definition itself, showing an omission about *aims*. Unless you know the aims of the science, behaviour, etc. in question, you cannot decide if it is rational or not. To put it another way, a rationality analysis requires an aim analysis. Now I turn my attention about the third point which goes to the heart of the definition.

**What is "the best"?** There is a problem about the word "best" in the definition "to follow the best methods, theories, and practices to create or acquire knowledge." What does *best* mean here? Best in which respects? There are a lot ways a theory can be better than another theory and these need not favour the same theories.

Consider the theory of mind (ToM) issue discussed in section 2.9. The aim of this research is to find out if apes have some kind of ToM. Is it rational to attribute ToM to apes? This question is equivalent to which one of the alternative theories is better. As we have seen one side of the discussion has simplicity on their side and the other side has parsimony. Which one of the values simplicity and parsimony is the better one? This question is misleading. There is no a priori comparative hierarchy of values. You cannot expect an answer to a question like:

*On the scale of 0–100 give a number for each of the following values indicating how good is it for a theory to have that value: clarity, plausibility, familiarity, fruitfulness, usefulness, beauty, internal coherence, so on.*

This is of course the point Kuhn [1977a, 326] makes when he says that there is no weight function to decide the precise importance of values in theory appraisal. Scientists working in the same field can asses the values of a theory differently than each other. Furthermore, different values have different roles and importance in different fields. Therefore we cannot know what is considered to be good in a field without examining its values.

There have been those who have claimed that science is the pinnacle of rationality and there have been those that claim that rationality is a myth. Both camps are mistaken. To attribute (ir)rationality to a field of science, one must see that the best methods, theories, and practices are followed. This involves: (1) actually examining the field in question; (2)

deciding on the aims of the field; (3) subjecting the field to value analysis. If these steps can be concluded successfully, then we have an idea of what "good" means in this field and we can make rationality judgements.

**Consequences** There are a few consequences of my rationality analysis.

(1) The first is that the questions of the form "Is X rational?" (where X is a practice, research tradition, or part of science) is an acceptable question provided that X is selected narrowly enough such that the aims and values related to X can be identified. In particular, the question "Is science rational?" is an inadmissible question since it assumes a unity of aims and values in entire science which is simply absurd.

(2) Another consequence is that what is (ir)rational can change in time as the values, aims, and the alternative methods that serve that aim change. The ether theory was rational to use after Fresnel, but not after Einstein. The intelligent design theory was rational to use after Paley but not after the modern Darwinian synthesis.[62]

(3) Rationality of science or some other disciple cannot be decided on a priori grounds. You need to actually investigate and analyse the discipline to judge whether it is rational or not.

(4) It might not always be possible to decide on the rationality of a part X of science. This can happen in a number of ways:

(a) We might not understand or be unsure of the aims and values of X. We cannot know the best methods and practices in that case.

(b) If the problem or field is too complex, it might be hard to see the best theories or methods related to that problem. We saw an example of this in section 2.8: The helmet issue is so complex that there seems no completely rational way to asses the usefulness of helmets.

(c) It might not be possible to tell which side of a controversy is rational if there is no higher point of reference by which to judge these opposing theories (without actually arguing for one side or the other). To put it another way, to decide which side of a controversy is more rational, you might actually need the resolution of that controversy. In the case of ToM discussed in section 2.9, it is rational to hold that apes have some kind of ToM *if* simplicity is better than parsimony in this context. But this is exactly the bone of contention. It is not possible to tell which side of the controversy is rational since there is no higher point of reference to judge these opposing theories without actually arguing for one or the other.

One way out of this conundrum is to deem both ToM and anti-ToM views rational since they have equally valid claims to be the best theory to explain the behaviours of apes. But such an approach might weaken the concept of rationality so much that it can lose its punch.

**Overstating irrational elements in science.** Larry Laudan remarks that:

> If, however, science is predominantly irrational, then there is no reason to take its knowledge claims any more (or less) seriously than we take those of the seer, the religious prophet, the guru, or the local fortuneteller. [Laudan, 1977, 2]

Laudan is mistaken. I know that medicine is predominantly irrational as I discussed above. But, still, I would rather see a doctor than a guru, and I would rather get a pharmaceutical

[62] For an analysis of Paley's argument, see [Sober, 2000, chapter 2].

drug rather than snake oil. Why? Because just because medicine has irrational elements does not imply that it is on equal standing with quackery or homeopathy. It is rational to utilize medicine than the alternatives because it better serves the aim of improving the human health.

I find this "if irrational then no good" view not only only wrong, but also philosophically misleading: I suspect that it underlies much of the confused rationality debates. Some philosophers accept the "if irrational then no good" view and get worried that even a shred of irrationality would render all science useless. Their knee-jerk reaction is to equate science and rationality to avoid this outcome at all costs. But there is no need to panic and take such drastic measures. The assumption "if irrational then no good" is simply wrong.

> Of course scientists have goals, both individual and collective, and they employ more or less effective means for achieving these goals. So one may invoke an "instrumental" or "hypothetical" notion of rationality in explaining the success or failure of various scientific enterprises. But what is at issue is just the effectiveness of various goal-directed activities, not rationality in any more exalted sense which could provide a demarcation criterion distinguishing science from other human activities, such as business or warfare. What distinguishes science is its particular goals and methods, not any special form of rationality. [Giere, 2001, 41]

Diagnosing irrationality in science should be seen as a fortunate event to improve and remove hitherto hidden faults, not as a sign of doomsday. Deifying science to be necessarily rational would only hurt science itself.

**Conclusion.** In this section I investigated a particular definition of rationality and some of its consequences. But there are three features of rationality which must be defended independently of that particular definition, and I want to stress them. First, philosophers should avoid the habit of cobbling together all sciences and applying or denying the same label (rationality in this case) to science. Scientific traditions are so diverse that they cannot be uniformly grouped together, especially as far as the rationality issue is concerned. It is possible that one part of a given science is rational, another is not. Second, rationality cannot be decided by decree. This is an empirical problem. One must investigate if some tradition or behaviour is rational. The claims of (ir)rationality should be based on practices of a science. Finally, all irrational theories are not on par. There is no point in having an irrational fear of irrational elements in science.

## 5.4 Underdetermination

> An unhappy alternative is before you, Elizabeth. From this day you must be a stranger to one of your parents. Your mother will never see you again if you do not marry Mr. Collins, and I will never see you again if you do.
>
> Jane Austen, *Pride and Prejudice*

I will first define and give examples of underdetermination and then explain why it is a misguided concept from the value analysis perspective.

**Underdetermination of theory by evidence.** To begin with, *underdetermination* is short for *underdetermination of theory by evidence.* The idea is that evidence might not be enough to confirm a theory; rather, it might equally support different theories. Stathis Psillos mentions two senses of underdetermination:

> The claim that evidence underdetermines theory may mean two things: first, that the evidence cannot prove the truth of the theory, and second, that the evidence cannot render the theory probable. Let us call the first deductive underdetermination and the second inductive (or ampliative) underdetermination. Both kinds of claims are supposed to have a certain epistemic implication, namely that belief in theory is never warranted by the evidence. This is the underdetermination thesis. [Psillos, 2006, 575]

There is nothing to discuss about deductive underdetermination because it is just true: Every theory surpasses the evidence for it. What about inductive underdetermination? The answer depends on whether or not there are *empirically equivalent* theories. If there are incompatible theories that evidence equally supports, then the evidence cannot render one theory more probable. Let us consider examples of *logically* empirically equivalent theories — theories that are proven to be empirically equivalent. Psillos gives the following example:

> Yet it seems that there is a genuine case of empirical equivalence of theories of quantum mechanics. Alternative interpretations of the quantum-mechanical formalism constitute empirically equivalent but different theories that explain the world according to different principles and mechanisms. The most typical rivalry is between the orthodox understanding of quantum theory — the "Copenhagen interpretation," according to which a particle cannot have a precise position and momentum at the same time — and the Bohmian understanding of quantum theory — the hidden-variables interpretation, according to which particles always have a definite position and velocity, and hence momentum. On Bohm's theory, particles have two kinds of energy: the usual (classical) energy and a "quantum potential" energy. More recently, there have been three particularly well-developed theories (the Bohmian quantum mechanics, the many-worlds interpretation, and the spontaneous-collapse approach) such that there is no observational way to tell them apart. And it seems that there cannot be an observational way to tell them apart. This situation is particularly unfortunate, but one may respond that the ensued underdetermination is local rather than global; hence the possible skepticism that follows is local. [Psillos, 2006, 576–577]

I am not sure why Psillos downplays this case. But there are a number of other examples of underdetermination in theoretical physics. One of the most interesting papers about underdetermination is by J. Brian Pitts [2011] which discusses a number of different underdetermination cases in theoretical physics. Pitts concludes his paper with the words "As I read particle physics, the weight of examples of underdetermination above is fairly strong." Unfortunately, I cannot do justice to his examples here because they are over my head. Still, I want to mention a new type of underdetermination that Pitts establishes.

**One-parameter equivalence.** The oft-discussed version of underdetermination involves *exact empirical equivalence.* As its name suggests, this is when two theories have exactly the same empirical consequences. Pitts introduces different types of inexact empirical equivalence of which I will consider one, namely *one-parameter equivalence*:

> Let $\{(\forall m)T_m\}$ be a collection of theories labeled by a parameter $m$, where all positive values of $m$ are permitted. (One can admit a positive upper bound for $m$, but that change makes no difference.) Let $T_0$ be another theory of the same phenomena. If the empirical predictions of the family $\{(\forall m)T_m\}$ tend to those of $T_0$ in the limit $m \to 0$, then the family $\{(\forall m)T_m\}$ empirically approximates $T_0$ *arbitrarily closely*. Though $T_0$ is empirically distinguishable in principle from any particular element $T_i$ of $\{(\forall m)T_m\}$, yielding merely transient underdetermination between any two theories, $T_0$ is not empirically distinguishable from the entire family. At any stage of empirical inquiry, there are finite uncertainties regarding the empirical phenomena. If $T_0$ presently fits the data, then so do some members of $\{(\forall m)T_m\}$ for nonzero but sufficiently small $m$. While scientific progress can tighten the bounds on $m$ towards 0, human finitude prevents the bounds from being tightened to the point that all nonzero values of $m$ are excluded while $T_0$ is admitted. Thus for any stage of empirical science, there will be underdetermination between $T_0$ and elements of $\{(\forall m)T_m\}$ with $m$ close enough to 0, if $T_0$ is viable. The underdetermination between $T_0$ and part of $\{(\forall m)T_m\}$ is in this sense permanent. One can never exclude empirically all the $T_m$ theories with $m > 0$. [Pitts, 2011, 271]

For example, $T_0$ can be a theory that implies that a certain type of particle has no mass and $T_m$ implies that it has mass $m$. Pitts gives examples of a number of such theories. The point is that since there will be always a limit of our experimental capabilities, the massless theory $T_0$ cannot be empirically separated from the massive variants.

**Underdetermination of theory by all currently available evidence.** Generally when philosophers give examples of underdetermination, it is from physics. This is not surprising since the most abstract parts of theoretical physics seem to be unrelated to experience. But underdetermination is not exclusive to physics; indeed, I have already given such examples of underdetermination in chapter 2.

As I have argued in section 2.9, whether animals have theory of mind or not cannot be decided on empirical grounds. Experimental results matter as long as they make one hypothesis simpler or more parsimonious than the alternative. Given all the evidence, unless we employ other values like simplicity or parsimony, it is not possible to overcome underdetermination.

The helmet issue in section 2.8 shows that the evidence can be so complex and multifaceted that it is possible to interpret it in both ways. The choice between the alternatives depend in the end on cultural, psychological and political aspects.

The examples I gave so far are examples of underdetermination of theory by all *possible* empirical evidence. It follows from the discussion so far that underdetermination holds in this strong sense. But, there is a weaker version of underdetermination as well: the underdetermination of theory by all *currently available* evidence. This happens when all available evidence is not enough to choose between alternative theories.

Archaeology is one branch of science that seems to be littered with this kind of underdetermination. There is never a lack of theories about the past but also never enough evidence to conclusively choose one over the others. It is still not even clear to what degree humans had a violent past and the effects it had. In appendix E you can find a quotation from British archaeologist Paul Bahn which discusses gender roles and how there are different "possible explanations for archaeological data." For an amusing look at underdetermination in archaeology see figure 5.12.

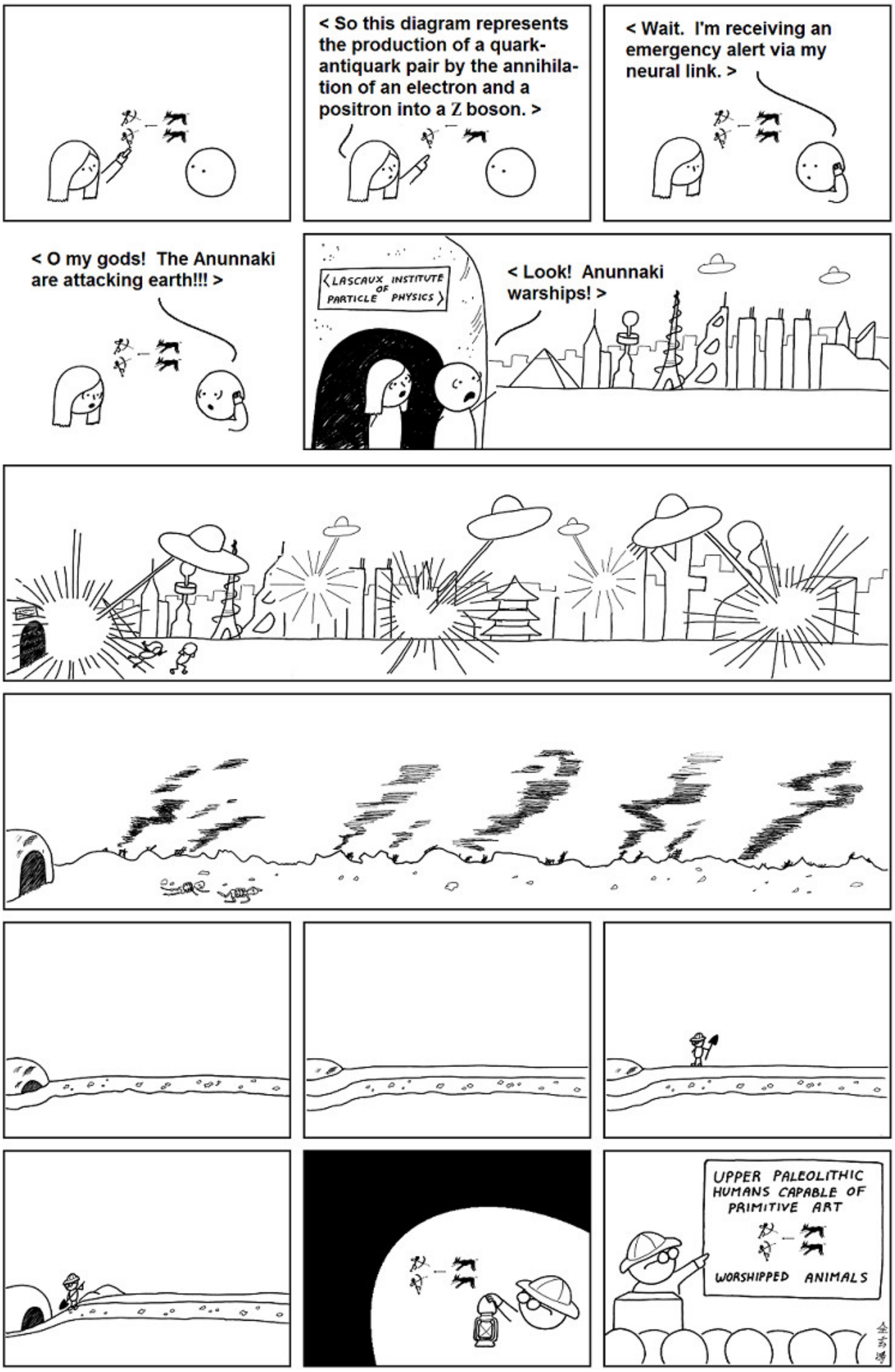



Figure 5.12: A comic from www.abstrusegoose.com which shows a different theory of cave paintings. [A pane removed for better fit. Image from www.abstrusegoose.com/525 licensed under Creative Commons Attribution-Noncommercial 3.0 United States License.]

The underdetermination by currently available evidence is definitely not exclusive to archaeology and happens a lot in science. Actually, I think it is the norm for three reasons.

(1) History of science shows us it is so. Scientists choose a theory before they have definitive evidence supporting it over its alternatives.

(2) Scientists need to continue their research and they cannot wait for an indefinite time period for conclusive evidence to accumulate. They choose a theory even if all currently available evidence is not enough to vindicate it. After all, there is no guarantee that such a vindication will be forthcoming.

The germ theory was accepted in the mid-nineteenth century by numerous scientists well before there was conclusive evidence for it. This theory turned out to be so fruitful, simple, and unifying that it gained popularity quickly.

(3) Evidence is not an independent judge of theories. Far from it, as we have seen in section 3.4, it is theory-laden.

(4) Evidence is not the only value in town. This is why in the beginning of this section I claimed that underdetermination is a misguided debate. To understand why let me elaborate what is wrong with the underdetermination debate.

**A misguided debate.** The underdetermination issue is based on the delusion about whether the only important value in theory appraisal is evidence or not. One side of the discussion tries to show that all that matters is evidence and the other side tries to show that the evidence might not be enough. The latter side is right in saying that the evidence might not be enough (and there are genuine cases of underdetermination) but this response misses the gist of the issue: Theory appraisal is a rich interplay of all kinds of values, only one of which is evidence. Even the way the underdetermination question is framed (is evidence enough?) shows the dismissal of other values.

There are a lot of cases of theory choice that shows that when the choice was made by the scientific community, it was made for reasons other than evidence. Evidence may turn out to support this choice later, but this does not change the fact that the choice made was not in favor of evidence. Even if evidence plays a role in acceptance of a theory, it is most likely that there are other values involved. For example as we saw in section 2.12 that even though the experiments were not enough to settle the dispute between the neuron theory and the reticular theory, scientific community at large accepted the neuron doctrine.

Underdetermination of theory by fruitfulness, underdetermination of theory by simplicity, etc. are as legitimate questions as underdetermination of theory by evidence. Of course, when you acknowledge the rich interplay of values this underdetermination talk becomes useless. Instead of talking of values undermining, what needs to be done is value analysis and seeing the full picture of values and how they are involved.

## 5.5 Enemies of Values and Practices

> "Drinking the Kool-Aid" is a figure of speech commonly used in North America that refers to a person or group holding an unquestioned belief, argument, or philosophy without critical examination. It could also refer to knowingly going along with a doomed or dangerous idea because of peer pressure. The phrase often times carries a negative connotation when applied to an individual or group.
>
> https://en.wikipedia.org/wiki/Drinking_the_Kool-Aid

There are some philosophies of science that do not bother to look at different values and practices in science. Naturally, they do not do a great job of capturing the intricacies of science. The aim of this section is to put three of these approaches under the microscope.

### 5.5.1 Confirmation Theory

There is a long tradition in analytical philosophy which tries to capture theory choice in a logical, mathematical, or algorithmical framework.[63] Generally, one philosopher would give such a formal account $A_1$ and another would give an example $E_1$ (real or constructed) that shows the inadequacy of $A_1$. The original account is then changed to $A_2$ which can deal with $E_1$ but this time some other example $E_2$ is found showing the problems with $A_2$. This *dance* continues for a while: candidate framework $A_n$ is shown to be inadequate by $E_n$. Then either all parties realize that no modification of $A_i$'s would be sufficient to capture how science works or they get bored with the dance and move on to new things. Then the account $A_1$ and its modifications $A_2, \ldots A_n$ are either completely forgotten or mentioned once in a blue moon as a pedagogical cautionary tale.

The problem with this dance is its futility. There is *no* logical, mathematical, or algorithmical account that would capture how theory choice in science works. None! There are a few reasons why:

(1) As I have repeated a few times already: Sciences are so rich and varied that these cannot be captured in an artificial framework. There is no uniformity of values or practices that could be captured in a formula.

(2) Theory choice is not made with rules but with values that simply are *not* rules. Hence they could not be made algorithmical. See section 3.2.

(3) Recall from section 3.1 where I said that "what is important in theory appraisal is not general values but the role of concrete values related to that theory." This means that any general account of theory appraisal is a non-starter.

If someone claims that a formal account $A$ captures theory choice in science, then one way to answer that claim is to find a counterexample $E$ that shows it does not work. But, really, there is no need to go through this song and dance every time: no formal account can ever work and that's that.

[63] These approaches have different names: confirmation theory, game theory, decision theory, formal epistemology, and whatnot.

### 5.5.2 Science Police

I have the following meaning of 'police' in mind: *people who seek to regulate a specified activity, practice, etc.*, for example, *the language police*.[64] The *science police* are the people who regulate what science is by means of some demarcating criteria. Most of the time, they have a *necessary criterion* that tells us a minimum condition that a science has to satisfy: If a practice A is scientific then it passes the criteria C. To put it another way, if a practice A fails to pass C, then A is not science. There are a number of such criteria considered by philosophers: rationality, objectivity, falsifiability and so on. I will look at one of these criteria, namely objectivity, yet, I believe that my remarks can be modified to cover some of the other criteria as well.

I will explain that the objectivity criterion (if a practice A is scientific then it must be objective) does not work as a demarcation criterion but first let me preempt a possible objection. In section 5.2, I praised unbiasedness in science as an important aim and here I argue against objectivity which is frequently thought to include unbiasedness. Is this not a contradictory stance? Not really. I defend a *bottom-up* approach of unbiasedness, not a *top-down* approach of objectivity.

Biases are something to be found out one at a time in scientific work. Only when we find about a bias and its effects do we have an option to eliminate it, and that is, *if* we can eliminate it. My approach starts with science and tries to find out and get rid of biases. Objectivity approach, on the other hand, starts with a notion of objectivity and filters what is science or not according to this notion. Let me further explain this difference by giving an interesting example from a recent book by the science journalist Matt Kaplan.[65]

> Recent work in animal behavior has revealed something fascinating: There are personality types in animals. Among fish in a single species, there are adventurous individuals, ready and willing to take risks, and there are more cautious and timid individuals, fearful of doing anything that could put them in danger. Similar variations in personality are starting to be found in birds and mammals too. A recent study led by Kathryn Arnold at the University of York revealed that when greenfinches were presented with brightly colored objects in their food, there was considerable variation in how long it took each bird to eat. When intriguing objects were attached to the birds' perches, a similar variation was found. Some birds quickly flew to explore the new toy while others stayed away.
>
> Being courageous or curious undoubtedly presents serious dangers. Ongoing studies indicate that fish with more daring personalities are more likely to nibble on bait on the end of a hook and risk-taking rodents more commonly end up in traps set by researchers. Yet having a personality that predisposes an animal to take risks can yield rewards. Courage can lead an animal to investigate previously unexplored locations where food is present, or it can lead to the discovery of well-hidden nesting areas that have yet to be found by any other members of the species. Such discoveries can lead to better health and better breeding

[64] http://dictionary.reference.com/browse/police.

[65] Different editions of Kaplan's book have different names. My reference is to the 2012 hard-cover edition.

> opportunities for the courageous animal that allow for its courageous genes to be passed along more readily to the next generation. [Kaplan, 2012, 4–5]

What makes this scientific development about animal behaviour interesting for us is the following unexpected consequence:

> As it happens, this has really screwed up lots of biological research. We have spent decades "thinking" we could get a reasonable sense of what animals are like by setting traps in the wild and then studying the animals that get caught. But if the animals that get caught are only the most daring individuals (or the most foolish) in a population, they are hardly giving us a reasonable sense of how a species behaves! [Kaplan, 2012, 5]

It turns out that biology has a sampling bias no one was aware of until now. Biologists were thinking for a long time that they were trapping random members of a species but we now know that their selection methods did not represent all members equally likely. Having discovered this sampling bias in their work, biologists can work on the steps to take to abate the bias. They might even choose to restrict the scope of their scientific claims from all members to a smaller set to deal with the bias.

This turn of events fits well with my *bottom-up* approach of unbiasedness. Biologists found out a bias in their work and now they have a chance to work on eliminating it. But the science police runs into serious problems. The following conclusions are inescapable from the viewpoint of top-bottom approach of objectivity.

(1) When it turns out that a "science" is not unbiased/objective, then it is not science by the objectivity criterion. In the example above, biology is not a science because we have found that it is not objective. There is no beating around the bushes with this outcome. If one accepts that a practice cannot be science without being objective, then the part of biology that studies animal behaviour is not a science. We cannot attribute the label "science" to biology.

(2) We were mistaken in saying that "biology is a science" all these years. Since it was not objective, our past attributions of the label "science" were mistaken.

(3) We cannot know what is science! Never! The reason is that a practice can be unobjective/biased without us knowing about it. Since we do not know what is objective, we cannot know what is scientific.

(4) The science police claim that all attributions of "science" by the scientists and the public are either wrong or unjustified. In the case of biology, scientific community and public is making a mistake in calling biology a science since it is not objective. Any other practice which scientists or public calls science is unjustifiably called so because we do not know what is objective.

(5) The science police also disallows some activities that one normally considers to be part of science. Scientists do not come up with perfect theories — they might have known or unknown problems including some biases.[66] Improvements will, hopefully, dispose of these problems. But the science police only allows objective work to be part of science. So you need to come up with perfect theories from the get-go.

I do not think that the science police has considered the consequences of their views thoroughly. The objectivity criterion makes all human practices unscientific as none is completely

[66] Recall the LAST example from section 4.5.


Figure 5.13: A 2007 South Korean miniature sheet with two stamps issued to commemorate the Year of Biology.

unbiased. Moreover, since it is likely that some kind of bias will be always present, we will never have science. Ironically, even if we had science nobody would know it as there is no way to be sure that there is no hidden bias left. The science police has managed to kill all science and any possibility of ever having science.

Now there are ways to modify the objectivity criterion. For example, some kind of biases can be allowed in science. Or, the condition “science cannot have any unobjective elements” can be weakened to “science cannot have any unobjective elements that we know of”. All these and similar weakenings of the criterion allows science to be unobjective or biased somehow. I think that letting some biases in does open the floodgates and strips all power from the criterion. I cannot see any meaningful weakening of this criterion.

The science police should stop trying to dictate what science is and enjoy science for all its glory, including the real scientific way of getting rid of biases: not by fiat, but by hard work.

**A corollary.** South Korea designated the year 2007 as the Year of Biology. They obviously did not do so to celebrate the bias in biology but instead they consider it a good and important science that needs to be endorsed: “The Korean government has designated 2007 as the Year of Biology in an effort to promote the significance of biology, a kind of basic science, and to elicit the Korean people’s interest in it.”[67] South Korea even issued a new stamp to commemorate the Year of Biology (see figure 5.13.)

The gist of the issue is that a lot of work done in biology deserves to be called *good science* that can be celebrated as the Korean government did despite the bias biology has. *A biased science can be good!* The discovery of a bias in a science does not immediately render it unscientific or

[67] From http://www.koreapost.go.kr/eng/html/woopyo/2007_07.jsp?contId=e1040104. Accessed on December 12, 2015.

useless junk. It might be quite detrimental or it might not be — depends on the case. That is one reason why we should stay away from hasty science policing that does not work at a case by case basis but deems all biased work unscientific. In the case of the bias found in biology, it does not change the fact that a lot of the research on animal behaviour was impressive, useful, innovative, and good science. The discovery of the bias gives scientists a chance to make it even better.

Let me translate what I have been saying into the objectivity language (even though I am against using it) in order to make it clear for the science police: A science can be *good* and *unobjective* (insert 'irrational' or 'unfalsifiable' or whatnot here) at the same time!

### 5.5.3 Social Constructivism

> "reality" cannot be used to explain why a statement becomes a fact, since it is only after it has become a fact that the effect of reality is obtained.
>
> Bruno Latour and Steve Woolgar [1986, 180]

Here are some tenets of social constructivism:

(1) We should be trying to understand real science as it is practised by real people. (Recall also Shapin's subtitle I mentioned in section 1.1.)

(2) Science is a product of our society and it is influenced by social factors.

(3) Societal values have an important role in science.

(4) Science is not objective.

(5) Furthermore, there is no such thing as objective reality.

(6) The rationality claims of the sciences are unfounded.

These sentences do not exhaust social constructivism and I deliberately formulated them in a way that make them superficially similar to some of the claims I made in this text. If I agree with them then I must be a social constructivist, right? Not really.

To begin with, there is only one item in the above list that I would unqualifiedly accept and that is (1). Philosophy of science should not be an exercise in creating science fiction as I previously mentioned in this chapter and also in appendix D. I have problems with all of the other statements, especially the way they are taken to the extreme by social constructivism.

Jumping to (4), it looks like something that I could sign under. That might be the case provided that my proviso about unbiasedness is added. I believe that science should aim to be unbiased and for the most part, it does. Just because I throw objectivity out of the window does not mean that I open the door to biases of all kind (including social biases). Getting rid of biases is not something high on the list of social constructivists.

When I read the sentence (5), the robot inside me wants to yell "does not compute". Objectivity is one of the most unclear concepts and it is especially hard to understand what is claimed here. If we have no access to reality, then how could we know the nature of reality? In such a reading, the statement becomes self-defeating. The most charitable reading of this statement is that evidence is independent from theory and values do not exist. I mentioned the theory-ladenness and value-ladenness of evidence before and I would agree with this reading of (5). But I doubt this is the intended meaning and there seems to be a confusion between epistemic access and ontology.

The claim (6) is simply absurd. As I explained in 5.3, it is actually the blanket claims of (ir)rationality of all science that are unfounded. It is rational to see a doctor than a guru, and to get a pharmaceutical drug rather than snake oil.

Looking at (2), science is indeed a product of our society and it is influenced by social factors. But it does not end there. As discussed in section 3.1, social, cultural, biological, cognitive, environmental, and historical factors have a say.

Finally, (3) says that societal values have an important role in science. Sure, societal values do indeed have their roles, and as I claimed in 3.5, understanding them is one of the more salient issues in philosophy of science. But we must not forget that there are all kinds of other values in science. I don't know what kind of black magic it is, but social constructivists are unable to see any other kind of value in science. Naomi Oreskes remarks that

> Historical case studies can illustrate how the development of a particular idea — including our best science — reflects the constraints of historical situations, and in recent years historians have produced many such studies. But in many such contextualized histories of science, social context is a kind of miasma that pervades scientific thinking in an intangible and ultimately inexplicable manner. The evidence for the role of social forces in the production of scientific knowledge is almost always circumstantial. [Oreskes, 1999, 4–5]

These circumstantial role of societal values are exaggerated by social constructivists at the expense of other values. In the end, their cases fail to be convincing.

One of the asymmetric properties of social constructivists' science analysis is the importance they give to weak aims and the neglect of strong aims. As I have emphasized in section 5.1, the prominent aims are the strong ones. It is these types of aims that determine the area of research and day to day work of researchers. Moreover, weak aims are plastic and their connection to research is not something to be assumed but rather demonstrated. This is one of the things that made Oreskes' analysis of the history of continental drift persuasive.

The tendency to block out non-social factors and non-societal values coupled with the putting weak aims at the forefront gives a lop-sided and impoverished analysis of science.

To sum up, I agree with (1), disagree with (6), puzzled about (5), and find (2–4) to be a small patch from a much a bigger picture.

# CHAPTER 6

# CONCLUSION

"Begin at the beginning," the King said, very gravely, "and go on till you come to the end: then stop."

Lewiss Caroll, *Alice in Wonderland*, 1865

The words "exegesis" and "eisegesis" related to interpretation of a (religious) text can be adapted for our purposes. Both words, originally from Greek, mean to interpret. But there is a crucial difference: Literally, the former means to "to lead out" and the latter means "to lead into". *Exegesis* is the interpretation of a text based on a careful analysis aiming to uncover the meaning of the text. *Eisegesis*, on the other hands, is the interpretation of a text aiming to find preconceived ideas in the text. This occurs when one "reads into the text" introducing his presuppositions, agendas, or biases into the interpretation.

By analogy, we can extend these two types of interpretation to science as well. What I have defended in this thesis is that we carry out an exegesis of science and stay away from eisegesis. If we want to understand how science works, we *cannot* achieve this by insisting that there is a single value (be it evidential, societal, or something else) operational in science. Neither neglecting those aspects of science that do not fit into one's philosophical commitments, nor replacing real science with hypothetical or ideal science (whatever that may be) would help our understanding.

Consider the old dispute nature vs. nurture or its more recent reincarnation genetics vs. environment. We now know that the truth lies in between the two extremes. But telling that the right answer is the middle-ground is not by itself satisfactory. (Recall figure 1.2.) Anyone can tell you that both nature and nurture viewpoints are mistaken. What is important is to tell *why* or *how* it is. What needs to be established is how exactly different factors work together and influence us. Today there are a number of disciplines that are paving the way to a detailed understanding.

Similarly, I have argued that the two extreme views related to values in science are mistaken. But this would be a trite comment on its own. What is important is to find a way to increase our philosophical understanding of science. The methods I have propounded in this text, namely value analysis, aim analysis, and more generally, looking at the practices of science, are crucial for our comprehension.

Clark Glymour starts the preface of his book *Theory and Evidence* [1980] as follows: "If it is true that there are but two kinds of people in the world — the logical positivists and the god-damned English professors — then I suppose I am a logical positivist." But this dichotomy is a false one. We do not have to be either blinded positivists or dogmatic relativists. We do not have to look at science spellbound with a philosophical shackle. Instead, we can read values and practices from science. The *philosophy by imposition* needs to be replaced by *philosophy by analysis*.

I have claimed in section 5.2 that science should aim to be unbiased and the same goes for philosophy of science as well. I do not claim that we can wave a magic band and immediately get rid of all biases. On the contrary, I think that we can *never* eliminate all biases from philosophy or science. But this does not mean that philosophy by analysis is impossible. Philosophers should aim to be unbiased and do their best to analyse science. The more we learn about biases, the more we will eliminate them. Even if we our science is inflicted by some biases, we will still understand science infinitely better using analysis than using the alternative, which is philosophy by imposition.

Today we know so much more about biases than a century ago. No doubt, we can deal with them better today and, for sure, we will do better in the future. There might be some socio-biological threshold that we can never pass, but having only a baseline level of bias is much better than rampant bias. After all, understanding science is something to be carried out by *us*, and our particular socio-biological nature will be always relevant. This is consistent with aiming to be unbiased.

I preceded the philosophy part of this thesis by numerous case studies. These show that it is possible to read values and practices off science. Furthermore, they show how varied the values and practices can be. It is a futile attempt to capture this variety with a blind commitment to a few values. This is why those who commit to eisegesis end up with surreal ideas about science. When one actually analyses values, aims, and practices of science, these ideas can be seen for what they are: science fiction.

Eisegesis makes one blind to the diversity of sciences. When one is programmed to see only some types of values or practices, unsurprisingly, he does not recognize the richness of sciences. The result of this monotonous outlook is the simplistic philosophy some of which we have seen in chapter 5.

The way out of this conundrum is neither to put science on so high a pedestal that we cannot investigate it, nor to dismiss it as a communal positive feedback loop. Philosophers should look at the practices, values, aims of the sciences without a commitment to a grandeur philosophy of what science is or how it works. The road to understanding science is paved with biases, but it is not a vicious circle. The more we analyse science and also our previous analyses, the more faithful our image of science will get to reality.

# BIBLIOGRAPHY

Abramson, John. *Overdosed America: The Broken Promise of American Medicine*. Harper Perennial, paperback edition with a new preface edition, 2013.

Alexander, L. G. *Longman English Grammar*. Longman, 1988.

Andrews, Kristin. Chimpanzee theory of mind: Looking in all the wrong places? *Mind and Language*, 20(5):521–536, 2005.

Angell, Marcia. *The Truth About the Drug Companies: How They Deceive Us and What to Do About It*. Random House, 2004.

Arnett, Jeffrey J. The neglected 95%: Why American psychology needs to become less American. *American Psychologist*, 63(7):602–614, 2008.

Ashtekar, Abhay. The winding road to quantum gravity. *Current Science*, 59(12):2064–2074, 2005.

Baggott, Jim. *Farewell to Reality: How Modern Physics Has Betrayed the Search for Scientific Truth*. Constable, 2013.

Bahn, Paul. *Archaeology: A Very Short Introduction*. Oxford University Press, 2012.

Bailer-Jones, Daniela. Models, metaphors and analogies. In Machamer, Peter and Silberstein, Michael, editors, *The Blackwell Guide to the Philosophy of Science*. Wiley-Blackwell, 2002.

Barnard, G. A. Discussion of Kiefer, 1959. *Journal of the Royal Statistical Society Series B*, 21: 311–312, 1959.

Barrett, Louise. *Beyond the Brain: How Body and Environment Shape Animal and Human Minds*. Princeton University Press, 2011.

Bavington, Dean. Science manages the sea. In Cayley, David, editor, *Ideas on the Nature of Science*. CBC, 2009.

Bavington, Dean. *Managed Annihilation: An Unnatural History of the Newfoundland Cod Collapse*. UBC Press, 2011.

Begley, C. Glenn and Ellis, Lee M. Raise standards for preclinical cancer research. *Nature*, 483 (7391):531–533, 2012.

Berlucchi, Giovanni. Some aspects of the history of the law of dynamic polarization of the neuron. from William James to Sherrington, from Cajal and Van Gehuchten to Golgi. *Journal of the History of the Neurosciences*, 8(2):191–201, 1999.

Bernal, John Desmond. The structure of graphite. *Proceedings of the Royal Society of London. Series A*, 106(740):749–773, 1924.

Bernal, John Desmond. *The Social Function of Science*. The Macmillan Company, 1939.

Bernard, Claude. *An Introduction to the Study of Experimental Medicine*. Dover Publications, 1865. Originally published in French. My page references are to the 1957 Dover print.

Bhattacharjee, Yudhijit. *The Mind of a Con Man*. The New York Times, 2013.

Biber, Douglas; Conrad, Susan, and Leech, Geoffrey. *Longman Student Grammar of Spoken and Written English*. Longman, 2002.

Biddle, Justin and Winsberg, Eric. Value judgements and the estimation of uncertainty in climate modeling. In Magnus, P. D. and Busch, Jacob, editors, *New Waves in Philosophy of Science*, pages 172–197. Palgrave Macmillan, 2010.

Bird, Alexander. Thomas Kuhn. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/archives/win2013/entries/thomas-kuhn/, winter 2013 edition, 2013.

Bissell, Mina. The risks of the replication drive. *Nature*, 503(7476):333–334, 2013.

Blackburn, Simon. *Oxford Dictionary of Philosophy*. Oxford University Press, second revised edition, 2008.

Bogard, D. D. and Johnson, P. Martian gases in an Antarctic meteorite? *Science*, 221(4611): 651–654, 1983.

Börjars, Kersti and Burridge, Kate. *Introducing English Grammar*. Hodder Education, second edition, 2010.

Bristol Post, The. Hanham mum: My son's cycling helmet saved his life after crash. 2014. Available from http://www.bristolpost.co.uk/Hanham-mum-son-s-cycling-helmet-saved-life-crash/story-21117394-detail/story.html. Posted on May 21, 2014. Accessed on April 10, 2015.

Broughton, Geoffrey. *Penguin English Grammar A–Z: Exercises for Advanced Students*. Penguin Books, 1990.

Brown, Andrew. *J. D. Bernal: The Sage of Science*. Oxford University Press, 2006.

Brush, Stephen G. *Transmuted Past: The Age of the Earth and the Evolution of the Elements from Lyell to Patterson*. Cambridge University Press, 1996.

Bueno, Otávio. Dirac and the dispensability of mathematics. *Studies in History and Philosophy of Modern Physics*, 36(3):465–490, 2005.

Burgess, Graham; Nunn, John, and Emms, John. *The Mammoth Book of the World's Greatest Chess Games*. Robinson Publishing, 1998.

Call, Josep. Chimpanzee social cognition. *Trends in Cognitive Sciences*, 5(9):388–393, 2001.

Carroll, Lewis. *The Annotated Alice*. W. W. Norton & Company, 2000. Introduction and Notes by Martin Gardner.

Catena, Riccardo. Analysis of the theoretical bias in dark matter direct detection. *Journal of Cosmology and Astroparticle Physics*, 2014(09), 2014.

Chang, Hasok. *Inventing Temperature: Measurement and Scientific Progress*. Oxford University Press, 2004.

Chang, Hasok. Operationalism. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/archives/fall2009/entries/operationalism/, fall 2009 edition, 2009.

Chow, Timothy Y. A beginner's guide to forcing. *http://arxiv.org/abs/0712.1320*, 2008.

Cohen, Paul J. The independence of the continuum hypothesis. *Proceedings of the National Academy of Science*, 50:1143–1148, 1963.

Colombo, Matteo; Hartmann, Stephan, and van Iersel, Robert. Models, mechanisms, and coherence. *The British Journal for the Philosophy of Science*, 2014.

Colyvan, Mark. Indispensability arguments in the philosophy of mathematics. In Zalta, Edward N., editor, *The Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/archives/spr2015/entries/mathphil-indis/, spring 2015 edition, 2015.

Committee, of Tilburg/Groningen/Amsterdam. *(Stapel Investigation Report) Flawed science: The fraudulent research practices of social psychologist Diederik Stapel*. https://www.commissielevelt.nl/, 2012. Accessed on April 28, 2015.

Crease, Robert P. and Mann, Charles C. *The Second Creation: Makers of the Revolution in Twentieth-Century Physics*. Macmillan Publishing Company, 1986.

Dasch, Pat and Treiman, Allan. *NASA LPI Slide Set: Ancient Life on Mars*. Lunar and Planetary Institute, 1997.

Daston, Lorraine J. and Galison, Peter. *Objectivity*. Zone Books, 2010.

Davies, James. *Cracked: Why Psychiatry is Doing More Harm Than Good by Davies*. Icon Books, 2013.

de Souza, Romualdo. Indiana University C460 Nuclear Chemistry Lecture Notes, 2014. Available from http://courses.chem.indiana.edu/c460/lecture_notes.asp.

Dennett, Daniel C. Intentional systems in cognitive ethology: The "Panglossian paradigm" defended. In *The Intentional Stance*, pages 237–268. The MIT Press, 1987. Originally published as an article in 1983.

Dick, Steven J. and Strick, James E. *The Living Universe: NASA and the Development of Astrobiology*. Rutgers University Press, 2004.

Dieudonne, Jean A. Review of The Prehistory of the Theory of Distributions by Jesper Lützen. *The American Mathematical Monthly*, 91(6):374–379, 1984.

Dirac, Paul. *The Principles of Quantum Mechanics*. Oxford University Press, 1958. First edition published in 1930.

Duhem, Pierre. *To Save the Phenomena: An Essay on the Idea of Physical Theory from Plato to Galileo*. The University of Chicago Press, 1969. Originally published in 1908 in French.

Easterly, William. *The Tyranny of Experts: Economists, Dictators, and the Forgotten Rights of the Poor*. Basic Books, 2013.

Edwards, Paul N. *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. The MIT Press, 2010.

Errington, Timothy M.; Iorns, Elizabeth; Gunn, William; Tan, Fraser Elisabeth; Lomax, Joelle, and Nosek, Brian A. Science forum: An open investigation of the reproducibility of cancer biology research. *eLIFE*, 2014. http://elifesciences.org/content/3/e04333.

Farley, John. *The Spontaneous Generation Controversy From Descartes to Oparin*. The Johns Hopkins University Press, 1977.

Fitzpatrick, Simon. Doing away with Morgan's Canon. *Mind and Language*, 23(2):224–246, 2008.

Flegel, Ilka and Söding, Paul. Twenty-five years of gluons. *CERN Courier*, November 12, 2014.

Foster, John Bellamy; Clark, Brett, and York, Richard. *The Ecological Rift*. Monthly Review Press, 2010.

Frank, Philipp G. *Philosophy of Science: The Link Between Science and Philosophy*. Prentice-Hall, 1962. Reprinted by Dover Publications, 2004.

Fritzsch, Harald and Gell-Mann, Murray, editors. *50 Years of Quarks*. World Scientific, 2015.

Garvey, Brian. *Philosophy of Biology*. McGill-Queen's University Press, 2007.

Gaukroger, Stephen. *Objectivity: A Very Short Introduction*. Oxford University Press, 2012.

Gelfert, Axel. Manipulative success and the unreal. *International Studies in the Philosophy of Science*, 17(3):245–263, 2003.

Gell-Mann, Murray. A schematic model of baryons and mesons. *Physic Letters*, 8(3):214–215, 1964.

Gibson, E. K.; McKay, D. S.; Thomas-Keprta, K. L.; Wentworth, S. J.; Westall, F.; Steele, A.; Romanek, C. S.; Bell, M. S., and Toporski, J. Life on Mars: evaluation of the evidence within Martian meteorites ALH84001, Nakhla, and Shergotty. *Precambrian Research*, 106:15–34, 2001.

Giere, Ronald N. Cognitive approaches to science. In Newton-Smith, W. H., editor, *A Companion to the Philosophy of Science*. Blackwell, 2001.

Glymour, Clark. *Theory and Evidence*. Princeton University Press, 1980.

Godfrey-Smith, Peter. *Philosophy of Biology*. Princeton University Press, 2014.

Goldacre, Ben. *Bad Pharma: How Medicine is Broken, and How We Can Fix It*. Fourth Estate, revised and updated edition edition, 2013.

Goldacre, Ben and Spiegelhalter, David. Bicycle helmets and the law. *British Medical Journal*, 346, 2011.

Goldman, Alvin I. Theory of mind. In Margolis, Eric; Samuels, Richard, and Stich, Stephen P., editors, *The Oxford Handbook of Philosophy Cognitive Science*. Cambridge University Press, 2012.

Goldstern, Martin. A taste of proper forcing. In di Prisco, Carlos Augistino.; Larson, Jean A.; Bagaria, Joan, and Mathias, A. R. D., editors, *Set Theory: Techniques and Applications, Curaçao 1995 and Barcelona 1996 Conferences*. Kluwer Academic Publishers, 1998.

Goozner, Merrill. *The $800 Million Pill: The Truth behind the Cost of New Drugs*. University of California Press, 2005.

Gould, Stephen Jay. The validation of continental drift. In *Ever Since Darwin: Reflections in Natural History*. W. W. Norton & Company, 1979.

Gould, Stephen Jay. Pildown revisited. In *The Panda's Thumb*. W. W. Norton & Company, 1980.

Grayling, A. C. Simplicity. In Brockman, John, editor, *This Idea Must Die: Scientific Theories That Are Blocking Progress*, pages 9–10. Harper Perennial, 2015.

Griffiths, David. *Introduction to Elementary Particles*. Wiley-VCH, second edition, 2008.

Guillery, R. W. Relating the neuron doctrine to the cell theory. Should contemporary knowledge change our view of the neuron doctrine? *Brain Research Reviews*, 55(2):411–421, 2007.

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

Hallam, Anthony. *Great Geological Controversies*. Oxford University Press, second edition, 1989.

Hanson, Norwood Russell. *The Concept of the Positron: A Philosophical Analysis*. Cambridge University Press, 1963.

Hare, Brian; Call, Josep, and Tomasello, Michael. Do chimpanzees know what conspecifics know? *Animal Behaviour*, 61(1):139–151, 2001.

Harris, Henry. *Things Come to Life: Spontaneous Generation Revisited*. Oxford University Press, 2002.

Haven, Kendall. *Marvels of Science: 50 Fascinating 5-Minute Reads*. Libraries Unlimited, 1994.

Hawking, Stephen. *Black Holes and Baby Universes*. Bantam Press, 1993.

Heinric, Bernd. The biological roots of aesthetics and art. *Evolutionary Psychology*, 11(3):743–761, 2013.

Henrich, Joseph; Heine, Steven J., and Norenzayan, Ara. The weirdest people in the world? *Behavioral and Brain Sciences*, 33:61–83, 2010.

Hewings, Martin. *Advanced Grammer in Use*. Cambridge University Press, 1999.

Holmes, Arthur. *Principles of Physical Geography*. Thomas Nelson and Sons Ltd, 1944.

Horn, D. Quarks in the bootstrap era. In Fritzsch, Harald and Gell-Mann, Murray, editors, *50 Years of Quarks*, pages 105–113. World Scientific, 2015.

Hu, Chenming Calvin. *Modern Semiconductor Devices for Integrated Circuits*. Prentice Hall, 2009.

ILCA, . *Low-level Aerial Survey Techniques: Report of an International Workshop Held 6–11 November 1979 Nairobi, Kenya*. International Livestock Centre for Africa Publications, 1981.

Jeffreys, Harold. *The Earth: Its Origin, History and Physical Constitution*. Cambridge University Press, sixth edition, 1976.

Jeng, Monwhea. A selected history of expectation bias in physics. *American Journal of Physics*, 74:578–583, 2006.

Johnson, George. *Strange Beauty: Murray Gell-Mann and the Revolution in Twentieth-Century Physics*. Vintage, 2000.

Kahneman, Daniel. *Thinking, Fast and Slow*. Penguin Books, 2012.

Kaplan, Matt. *Medusa's Gaze and Vampire's Bite: The Science of Monsters*. Scribner, 2012.

Kaptchuk, Ted J. Intentional ignorance: A history of blind assessment and placebo controls in medicine. *Bulletin of the History of Medicine*, 72(3), 1998.

Kaptchuk, Ted J. A brief history of the evolution of methods to control of observer biases in tests of treatments. http://www.jameslindlibrary.org/, 2011.

Kassirer, Jerome P. *On the Take: How Medicine's Complicity with Big Business Can Endanger Your Health*. Oxford University Press, 2005.

Kaxiras, Efthimios. *Atomic and Electronic Structure of Solids*. Cambridge University Press, 2003.

Kerr, Richard A. A lunar meteorite and maybe some from Mars. *Science*, 220(4594):288–289, 1983.

Kerr, Richard A. Martian meteorites are arriving. *Science*, 237(4816):721–723, 1987.

Kerr, Richard A. Requiem for life on Mars? support for microbes fades. *Science*, 282(5393): 1398–1400, 1998.

Knoll, Andrew H. *Life on a Young Planet: The First Three Billion Years of Evolution on Earth*. Princeton University Press, 2003.

Kronz, Fred and Lupher, Tracy. Quantum theory: von Neumann vs. Dirac. In Zalta, Edward N., editor, *The Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/archives/sum2012/entries/qt-nvd/, summer 2012 edition, 2012.

Kuhn, Thomas S. *The Structure of Scientific Revolutions*. The University of Chicago Press, 1962.

Kuhn, Thomas S. *The Structure of Scientific Revolutions*. The University of Chicago Press, second edition, 1970.

Kuhn, Thomas S. Objectivity, value judgment, and theory choice. In *The Essential Tension: Selected Studies in Scientific Tradition and Change*. University Of Chicago Press, 1977a.

Kuhn, Thomas S. The historical structure of scientific discovery. In *The Essential Tension: Selected Studies in Scientific Tradition and Change*. University Of Chicago Press, 1977b.

Kuhn, Thomas S. The essential tension: Tradition and innovation in scientific research. In *The Essential Tension: Selected Studies in Scientific Tradition and Change*. University Of Chicago Press, 1977c.

Kurlansky, Mark. *Cod: A Biography of the Fish that Changed the World*. Penguin Books, 1998.

Lakatos, Imre. Falsification and the methodology of scientific research programmes. In Worrall, John and Currie, Gregory, editors, *Philosophical Papers of Imre Lakatos*, volume 1, pages 8–101. Cambridge University Press, 1978a.

Lakatos, Imre. History of science and its rational reconstructions. In Worrall, John and Currie, Gregory, editors, *Philosophical Papers of Imre Lakatos*, volume 1, pages 102–138. Cambridge University Press, 1978b.

Lane, Peter and Barton, Philip. Use of surrogates in medicine: ideas that may be useful for ecology. In Lindenmayer, David; Barton, Philip, and Pierson, Jennifer, editors, *Indicators and Surrogates of Biodiversity and Environmental Change*. CRC Press, 2015.

Latour, Bruno and Woolgar, Steve. *Laboratory Life: The Construction of Scientific Facts*. Princeton University Press, 1986.

Laudan, Larry. *Progress and Its Problems: Towards a Theory of Scientific Growth*. University of California Press, 1977.

Leighton, Robert B. *Principles of Modern Physics*. McGraw-Hill Book Company, 1959. Available from https://archive.org/details/PrinciplesOfModernPhysics.

Lester, Mark. *English Articles and Determiners Up Close*. McGraw-Hill, 2013.

Ludlow, Peter. Descriptions. In Zalta, Edward N., editor, *The Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/archives/fall2013/entries/descriptions/, fall 2013 edition, 2013.

Lützen, Jesper. *The Prehistory of the Theory of Distributions*. Springer Verlag, 1982.

Mann, Howard and Djulbegovic, Benjamin. Comparator bias: why comparisons must address genuine uncertainties. *Journal of the Royal Society of Medicine (Supplement)*, 106(1):30–33, 2013.

McGarity, Thomas O. and Wagner, Wendy E. *Bending Science: How Special Interests Corrupt Public Health Research*. Harvard University Press, 2008.

McGrew, Roderick. *Encyclopedia of Medical History*. McGraw-Hill Book Company, 1985.

McKay, David S.; Gibson, Everett K.; Thomas-Keprta, Kathie L.; Vali, Hojatollah; Romanek, Christopher S.; Clemett, Simon J.; Chillier, Xavier D. F.; Maechling, Claude R., and Zare, Richard N. Search for past life on Mars: Possible relic biogenic activity in Martian meteorite ALH84001. *Science*, 273(5277):924–930, 1996.

McMullin, Ernan. Values in science. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1982(4):3–28, 1982.

Megill, Allan. Introduction: Four senses of objectivity. In Megill, Allan, editor, *Rethinking Objectivity*. Duke University Press, 1994.

Michaels, David. *Doubt is Their Product: How Industry's Assault on Science Threatens Your Health*. Oxford University Press, 2008.

Middleton, W. E. Knowles. *Invention of the Meteorological Instruments*. The Johns Hopkins Press, 1969.

Mitchell, Robert W.; Thompson, Nicholas S., and Miles, H. Lyn. *Anthropomorphism, Anecdotes, and Animals*. State University of New York Press, 1997.

Moore, Gregory H. The origins of forcing. In Drake, F. R. and Truss, J. K., editors, *Logic Colloquium '86*, pages 143–173. Elsevier, 1988.

Morgan, C. Lloyd. *Introduction to Comparative Psychology*. Walter Scott, 1894.

Morgenbesser, Sidney, editor. *Philosophy of Science Today*. Basic Books, 1967.

Newell, Peter. *Globalization and the Environment: Capitalism, Ecology and Power*. Polity Press, 2012.

Nickles, Thomas. Discovery. In Newton-Smith, W. H., editor, *A Companion to the Philosophy of Science*. Blackwell, 2001.

Nickles, Thomas. Scientific revolutions. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/archives/win2013/entries/scientific-revolutions/, winter 2013 edition, 2013.

Noorden, Richard Van. Parasite test shows where validation studies can go wrong. *Nature News*, 2014. http://www.nature.com/news/parasite-test-shows-where-validation-studies-can-go-wrong-1.16527.

Nunn, John. *Understanding Chess Middlegames*. Gambit, 2011.

Oerter, Robert. *The Theory of Almost Everything: The Standard Model, the Unsung Triumph of Modern Physics*. Pi Press, 2006.

Olive, K. A. and others, . Review of Particle Physics. *Chinese Physics C*, 38:090001, 2014.

Oreskes, Naomi. *The Rejection of Continental Drift: Theory and Method in American Earth Science*. Oxford University Press, 1999.

Oreskes, Naomi and Conway, Erik M. *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*. Bloomsbury Press, 2011.

Pietsch, Wolfgang. A revolution without tooth and claw. redefining the physical base units. *History and Philosophy of Science*, 46:85–93, 2014.

Pitts, J. Brian. Permanent underdetermination from approximate empirical equivalence in field theory: Massless and massive scalar gravity, neutrino, electromagnetic, Yang-Mills and gravitational theories. *British Journal for the Philosophy of Science*, 62(2):259–299, 2011.

Planck, Max. *Scientific Autobiography and Other Papers*. Philosophical Library, 1949.

Plantinga, Alvin. Methodological naturalism? In Pennock, Robert T., editor, *Intelligent Design Creationism and Its Critics*. The MIT Press, 2001.

Polchinski, Joseph. All strung out? *American Scientist*, 2007. Available from http://www.americanscientist.org/bookshelf/pub/all-strung-out. Accessed on April 25, 2015.

Popper, Karl R. The aim of science. In *Objective Knowledge, An Evolutionary Approach*. Oxford University Press, 1979.

Popper, Karl R. *Unended Quest*. Routledge, 2002. Originally published as a book chapter in 1974.

Povinelli, Daniel J. and Vonk, Jennifer. Chimpanzee minds: suspiciously human? *Trends in Cognitive Sciences*, 7(4):157–160, 2003.

Povinelli, Daniel J. and Vonk, Jennifer. We don't need a microscope to explore the chimpanzee's mind. *Mind and Language*, 19(1):1–28, 2004.

Premack, David and Woodruff, Guy. Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 4(4):515–629, 1978.

Principe, Lawrence M. *The Aspiring Adept*. Princeton University Press, 1998.

Prinz, Florian; Schlange, Thomas, and Asadullah, Khusru. Believe it or not: how much can we rely on published data on potential drug targets? *Nature Reviews Drug Discovery*, 10(712), 2011.

Psillos, Stathis. Underdetermination thesis, Duhem-Quine thesis. In Borchert, Donald, editor, *Macmillan's Encyclopedia of Philosophy*, volume 9, pages 575–578. Macmillan, second edition, 2006.

Putnam, Hilary. Indispensability arguments in the philosophy of mathematics. In Caro, Mario De and Macarthur, David, editors, *Philosophy in an Age of Science: Physics, Mathematics, and Skepticism*. Harvard University Press, 2012.

Ramón y Cajal, Santiago. *Comparative Study of the Sensory Areas of the Human Cortex*. 1899.

Reich, Eugenie Samuel. *Plastic Fantastic: How the Biggest Fraud in Physics Shook the Scientific World*. Palgrave Macmillan, 2009.

Reichenbach, Hans. *The Rise of Scientific Philosophy*. University of California Press, 1951.

Reicher, Maria. Nonexistent objects. In Zalta, Edward N., editor, *The Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/archives/win2014/entries/nonexistent-objects/, winter 2014 edition, 2014.

Resnik, David B. *The Price of Truth: How Money Affects the Norms of Science*. Oxford University Press, 2006.

Rickles, Dean. Quantum gravity: A primer for philosophers. In Rickles, Dean, editor, *The Ashgate Companion to Contemporary Philosophy of Physics*. Ashgate Publishing, 2008.

Righter, Kevin and Gruener, John. *Lunar Meteorite Compendium*. NASA, 2007.

Riordan, Michael. *The Hunting of the Quark: A True Story of Modern Physics*. Simon & Schuster, 1987.

Rivers, W. H. R. and Webber, H. N. The action of caffeine on the capacity for muscular work. *The Journal of Physiology*, 36(1), 1907.

Roberts, Royston M. *Serendipity: Accidental Discoveries in Science*. John Wiley & Sons, 1989.

Rost, Peter. *The Whistleblower: Confessions of a Healthcare Hitman*. Soft Skull Press, 2006.

Rothstein, Hannah R.; Sutton, Alexander J., and Borenstein, Michael, editors. *Publication Bias in Meta-Analysis*. John Wiley & Sons, 2005.

Sabin, Paul. *The Bet: Paul Ehrlich, Julian Simon, and Our Gamble over Earth's Future*. Yale University Press, 2013.

Sarewitz, Daniel. Beware the creeping cracks of bias. *Nature*, 485(7397):149, 2012.

Sarkar, Sahotra. Diversity: A philosophical perspective. *Diversity*, 2(1):127–141, 2010.

Savage-Rumbaugh, Sue; Shanker, Stuart G., and Taylor, Talbot J. *Apes, Language, and the Human Mind*. Oxford University Press, 1998.

Sears, Derek W. G. Oral histories in meteoritics and planetary science — XVI: Donald D. Bogard. *Meteoritics & Planetary Science*, 47(3):416–433, 2012.

Segrè, Emilio. *From X-rays to Quarks: Modern Physicists and Their Discoveries*. Dover Publications, 2007.

Shapin, Steven. *Never Pure: Historical Studies of Science as if It Was Produced by People with Bodies, Situated in Time, Space, Culture, and Society, and Struggling for Credibility and Authority*. Johns Hopkins University Press, second edition, 2010.

Shapin, Steven and Schaffer, Simon. *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*. Princeton University Press, first princeton paperback printing with corrections, 1989 edition, 1985.

Shockley, William. *Electrons and Holes in Semiconductors: With Applications to Transistor Electronics*. D. Van Nostrand Company, 1950.

Simonyi, Károly. *A Cultural History of Physics*. RC Press, 2012.

Smith, George Davey and Ebrahim, Shah. Data dredging, bias, or confounding. *BMJ*, 325 (7378):1437–1438, 2002.

Smolin, Lee. *The Trouble With Physics: The Rise of String Theory, The Fall of a Science, and What Comes Next*. Houghton Mifflin Company, 2006.

Smoryński, Craig. *Adventures in Formalism*. College Publications, 2012.

Sober, Elliott. *Philosophy of Biology*. Westview Press, second edition, 2000.

Stanford, P. Kyle. *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*. Oxford University Press, 2010.

Stolberg, Michael. Inventing the randomized double-blind trial: the Nuremberg salt test of 1835. *Journal of The Royal Society of Medicine*, 99(12):642 – 643, 2006.

Strick, James E. *Sparks of Life: Darwinism and the Victorian Debates over Spontaneous Generation*. Harvard University Press, 2000.

Svesko, Andy. A detailed introduction to string theory. Available from http://people.oregonstate.edu/~sveskoa/, 2013.

Taylor, G. Jeffrey. An adulterated Martian meteorite. *PSR Discoveries*, July 1999.

Thomas, Anthony W. and Weise, Wolfram. *The Structure of the Nucleon*. Wiley-VCH, 2001.

Thorne, Kip S. *Black Holes and Time Warps: Einstein's Outrageous Legacy*. W. W. Norton & Company, 1994.

Tomasello, Michael; Call, Josep, and Hare, Brian. Chimpanzees versus humans: It's not that simple. *Trends in Cognitive Sciences*, 7(6):239–240, 2003a.

Tomasello, Michael; Call, Josep, and Hare, Brian. Chimpanzees understand psychological states — the question is which ones and to what extent. *Trends in Cognitive Sciences*, 7(4): 153–156, 2003b.

Tong, David. *Lectures on String Theory*. http://arxiv.org/abs/0908.0333v1, first edition, 2009.

Tong, David. *Lectures on String Theory*. http://arxiv.org/abs/0908.0333v3, third revised edition, 2012.

Triarhou, Lazaros C. and del Cerro, Manuel. Ramón y Cajal erroneously identified as Camillo Golgi on a souvenir postage stamp. *Journal of the History of the Neurosciences*, 21:132–138, 2012.

Tukey, John W. The future of data analysis. *The Annals of Mathematical Statistics*, 33(1):1–67, 1962.

van Fraassen, Bas C. *The Scientific Image*. Oxford University Press, 1980.

Veltman, Martinus. *Facts and Mysteries in Elementary Particle Physics*. World Scientific, 2003.

von Neumann, John. *Mathematische Grundlagen der Quantenmechanik*. Springer-Verlag, 1932.

Walsh, T.; Oliveira-Neto, J. M., and Moore, D. Chlorhexidine treatment for the prevention of dental caries in children and adolescents. In *Cochrane Database of Systematic Reviews, Issue 4, Article No. CD008457*. 2015.

Wegener, Alfred. *The Origin of Continents and Oceans*. Dover Publications, 1966. English translation of the fourth German edition, 1929.

Weiskopf, Daniel A. The theory-theory of concepts. In *The Internet Encyclopedia of Philosophy*. http://www.iep.utm.edu/th-th-co/, 2011.

Wilson, Jr., E. Bright. *An Introduction to Scientific Research*. McGraw-Hill Book Company, 1952. My page references are to the 1990 Dover reprint.

Woit, Peter. *Not Even Wrong: The Failure of String Theory and the Search for Unity in Physical Law*. Vintage, 2007.

Woodham, Roger. *BBC Learning English: Zero Article*. 2011. Available from http://www.bbc.co.uk/worldservice/learningenglish/grammar/learnit/learnitv198.shtml. Accessed on April 10, 2015.

Woodward, Arthur Smith. *The Earliest Englishman*. Watts & Co, 1948.

Worrall, John. Thomas Young and the "refutation" of Newtonian optics: a case-study in the interaction of philosophy of science and history of science. In Howson, Colin, editor, *Method and Appraisal in the Physical Sciences: The Critical Background to Modern Science, 1800–1905*. Cambridge University Press, 1976.

Worrall, John. Fresnel, Poisson and the white spot: The role of successful predictions in the acceptance of scientific theories. In Gooding, David; Pinch, Trevor, and Schaffer, Simon, editors, *The Uses of Experiment*. Cambridge University Press, 1989.

Zweig, George. An SU(3) model for strong interaction symmetry and its breaking II. CERN preprint, 1964.

Zweig, George. Origins of the quark model. Invited talk, Baryon 1980 Conference, 1980.

# INDEX

# APPENDICES

# Appendix A

# Verification Bias

The final report of the joint committee [Committee, 2012] on the Stapel investigation (see section 2.10) mentions a number of flaws in the publications of Diederik Stapel, only one of which is *verification bias* which is "the use of research procedures in such a way as to 'repress' negative results by some means." The following list is a verbatim copy of the variety of ways this bias arises in the work of Stapel.

- An experiment fails to yield the expected statistically significant results. The experiment is repeated, often with minor changes in the manipulation or other conditions, and the only experiment subsequently reported is the one that did yield the expected results. It is unclear why in theory the changes made should yield the expected results. The article makes no mention of this exploratory method; the impression created is of a one-off experiment performed to check the a priori expectations. It should be clear, certainly with the usually modest numbers of experimental subjects, that using experiments in this way can easily lead to an accumulation of chance findings. It is also striking in this connection that the research materials for some studies shows the use of several questionnaire versions, but that the researchers no longer knew which version was used in the article.

- A variant of the above method is: a given experiment does not yield statistically significant differences between the experimental and control groups. The experimental group is compared with a control group from a different experiment — reasoning that "they are all equivalent random groups after all" — and thus the desired significant differences are found. This fact likewise goes unmentioned in the article.

- The removal of experimental conditions. For example, the experimental manipulation in an experiment has three values. Each of these conditions (e.g. three different colours of the otherwise identical stimulus material) is intended to yield a certain specific difference in the dependent variable relative to the other two. Two of the three conditions perform in accordance with the research hypotheses, but a third does not. With no mention in the article of the omission, the third condition is left out, both in theoretical terms and in the results. Related to the above is the observed verification procedure in which the experimental conditions are expected to have certain effects on different dependent variables. The only effects on these dependent variables that are reported are those that support the hypotheses, usually with no mention of the insignificant effects on the other dependent variables and no further explanation.

- The merging of data from multiple experiments. It emerged both from various datasets and interviews with the co-authors that data from multiple experiments had been combined in a fairly selective way, and above all with benefit of hindsight, in order to increase the number of subjects to arrive at significant results.

- Research findings were based on only some of the experimental subjects, without reporting this in the article. On the one hand "outliers" (extreme scores on usually the dependent variable) were removed from the analysis where no significant results were obtained. This elimination reduces the variance of the dependent variable and makes it more likely that "statistically significant" findings will emerge. There may be sound reasons to eliminate outliers, certainly at an exploratory stage, but the elimination must then be clearly stated.

- Conversely, the Committees also observed that extreme scores of one or two experimental subjects were kept in the analysis where their elimination would have changed significant differences into insignificant ones; there was no mention anywhere of the fact that the significance relied on just one or a few subjects.

- Finally entire groups of respondents were omitted, in particular if the findings did not confirm the initial hypotheses, again without mention. The reasons given in the interviews were ad hoc in nature: "those students (subjects) had participated in similar experiments before"; "the students just answered whatever came into their heads", but the same students had in the first instance simply been accepted in the experiment and the analysis. If the omitted respondents had yielded different results they would have been included in the analyses.

- The reliabilities of the measurement scales used and the reporting thereof were often handled selectively, and certainly to the experimenters' advantage, in the sense of confirming the research hypotheses. It is impossible to "test" hypotheses with unreliable measurement instruments, and therefore, for example, the reliability was estimated for a part of the research group, with the unreported omission of "unreliable" subjects. Or items were selected differently for each study in such a way, which did lead to a reliable instrument, but with no awareness that that this was achieved at the expense of the mutual comparability of studies.

- Where discrepancies were found between the reliabilities as reported by the researcher (usually the alpha coefficient) and as calculated by the statisticians, the reported values were usually conspicuously higher. If the reliability of a dependent variable or a covariate was not reported, its value often turned out to be too low relative to the accepted standard (e.g. less than 0.60).

- Sometimes the reliability was deliberately not reported, in particular if it was extremely low. For instance, a co-author reported that the supervisors urged that the data be sold as effectively as possible, and discouraged attempts to undermine the data, which might make editors and reviewers suspicious. If they asked any questions the missing data would be provided later.

- There was also selective treatment of the measurement scales, depending on what it was required to prove. Variously one or two dimensions (underlying variables) were used with the same set of items in different experiments, depending on what was most expedient in the light of the research hypotheses.

- The following situation also occurred. A known measuring instrument consists of six items. The article referred to this instrument but the dataset showed that only four items had been included; two items were omitted without mention. In yet another experiment, again with the same measuring instrument, the same happened, but now with two different items omitted, again without mention. The only explanation for this behaviour is that it is meant to obtain confirmation of the research hypotheses. It was stated in the interviews that items that were omitted did not behave as expected. Needless to say, "good" ad hoc reasons were given, but none were mentioned in the publication, and the omissions could be ascertained only by systematically comparing the available survey material with the publication.

- When the re-analysis revealed differences in significance level (p values) when applying the same statistical tests, it was usual for the values reported in the articles to be "more favourable" for the researcher's expectations. Similarly, incorrect rounding was also found, for example: $p = 0.056$ became $p = 0.05$.

# Appendix B

# Bad Trials

Medicine aims to find safe and effective treatments. One might think that double-blind clinical trials are the pillars of unbiased research in medicine but there are ways to taint them and these are exhibited in a chapter titled *Bad Trials* in Ben Goldacre's book *Bad Pharma* [2013]. In this appendix I will give a list of short quotations describing different types of problems Goldacre reveals. This list should be seen as an appetizer and for the full course see the detailed and well-exemplified chapter of Goldacre. The paragraph headings below are taken verbatim from the subsection headings of Goldacre's chapter.

**Outright fraud.** One obvious way to ruin the validity of a trial is to commit fraud. Goldacre notes that "fraud is also fairly rare, as far as anyone can tell. The best current estimate of its prevalence comes from a systematic review in 2009, bringing together the results of survey data from twenty-one studies, asking researchers from all areas of science about malpractice. Unsurprisingly, people give different responses to questions about fraud depending on how you ask them. Two percent admitted to having fabricated, falsified or modified data at least once, but this rose to 14 percent when they were asked about the behaviour of colleagues. A third admitted other questionable research practices, and this rose to 70 percent, again, when they were asked about colleagues." [Goldacre, 2013, 174–175]

**Test your treatment in freakishly perfect "ideal" patients.** "In the real world, patients are often complicated: they might have many different medical problems, or take lots of different medicines, which all interfere with each other in unpredictable ways; they might drink more alcohol in a week than is ideal; or have some mild kidney impairment. That's what real patients are like. But most of the trials we rely on to make real-world decisions study drugs in unrepresentative, freakishly ideal patients, who are often young, with perfect single diagnoses, fewer other health problems, and so on." [Goldacre, 2013, 176] This causes problems when you apply the results to everyday patients.

**Test your drug against something rubbish.** It is possible to make a new treatment look good by "testing it against something that doesn't work very well" or testing against an older drug "given at an unusually high dose, which means it has worse side effects by comparison." [Goldacre, 2013, 180]

**Trials that are too short.** Trials are often brief "because companies need to get results as quickly as possible, in order to make their drug look good while it is still in patent, and owned by them." [Goldacre, 2013, 181] But the long-term effects can be different from the short-term ones.

**Trials that stop early.** "If you stop a trial early, or late, because you were peeking at the results as it went along, you increase the chances of getting a favourable result. This is because you are exploiting the random variation that exists in the data." [Goldacre, 2013, 184] If a new drug trial shows discouraging results after a certain period of time, then stopping the trial earlier makes the drug look much more favourable.

**Trials that stop late.** Sometimes "prolonging a trial — or including the results from a follow-up period after it — can dilute important findings, and make them harder to see." [Goldacre, 2013, 187]

**Trials that are too small.** Large trials are necessary if the aim is to detect a small difference between two treatments and also "The benefit of more participants is that it evens out the random variation among them." [Goldacre, 2013, 192] Small trials can bring forth accidental properties forth.

**Trials that measure uninformative outcomes.** Trials do not always directly test the effects of a drug but rather measure correlated variables. If these surrogate outcomes are inappropriate or weak representatives then the trials might "fail to measure real-world outcomes".

**Trials that bundle their outcomes together in odd ways** "Sometimes, the way you package up your outcome data can give misleading results. For example, by setting your thresholds just right, you can turn a modest benefit into an apparently dramatic one. And by bundling up lots of different outcomes, to make one big 'composite outcome', you can dilute harms; or allow freak results on uninteresting outcomes to make it look as if a whole group of outcomes are improved." [Goldacre, 2013, 194]

**Trials that ignore drop-outs.** Ignoring patients that drop out of trials when analysing the data can skew the results: "as soon as you start to think about why patients drop out of treatment in trials, the problems with this method start to become apparent. Maybe they stopped taking your tablets because they had horrible side effects. Maybe they stopped taking your tablets because they decided they didn't work, and just tipped them in the bin. Maybe they stopped taking your tablets, and coming to follow-up appointments, because they were dead, after your drug killed them. Looking at patients only by the treatment they took is called a 'per protocol' analysis, and this has been shown to dramatically overstate the benefits of treatments, which is why it's not supposed to be used." [Goldacre, 2013, 198–199]

**Trials that change their main outcome after they've finished.** "If you measure a dozen outcomes in your trial, but cite an improvement in any one of them as a positive result, then your results are meaningless. Our tests for deciding if a result is statistically significant assume that you are only measuring one outcome. By measuring a dozen, you have given yourself a dozen chances of getting a positive result, rather than one, without clearly declaring that. Your study is biased by design, and is likely to find more positive results than there really are." [Goldacre, 2013, 200]

**Dodgy subgroup analyses.** "If your drug didn't win overall in your trial, you can chop up the data in lots of different ways, to try and see if it won in a subgroup: maybe it works brilliantly in Chinese men between fifty-six and seventy-one." [Goldacre, 2013, 205]

**Dodgy subgroups of trials, rather than patients.** "You can draw a net around a group of trials, by selectively quoting them , and make a drug seem more effective than it really is. When you do this on one use of one drug, it's obvious what you're doing. But you can also do it within a whole clinical research programme, and create a confusion that nobody yet feels able to contain." [Goldacre, 2013, 210]

**"Seeding trials".** "Sometimes, trials aren't really trials: they're viral marketing projects, designed to get as many doctors prescribing the new drug as possible, with tiny numbers of participants from large numbers of clinics." [Goldacre, 2013, 212]

**Pretending it's all positive regardless.** "At the end of your trial, if your result is unimpressive, you can exaggerate it in the way that you present the numbers; and if you haven't got a positive result at all, you can just spin harder." [Goldacre, 2013, 216]

The above list points to the possible problems that could occur in clinical trials. Goldacre not only presents them in much more detail, but also he shows that this is an actual list of problems by giving numerous examples.

# Appendix C

# xkcd # 882

Here is a transcription of the xkcd comic from http://xkcd.com/882/ by Randall Munroe given in figure 4.7 on page 124.

[Panel 1. Ponytail runs up to another person, who then points off-panel to the scientists.]

Ponytail: *Jelly beans cause acne!*
Another: *Scientists! Investigate!*
Scientists: *But we're playing Minecraft! ... Fine.*

[Panel 2. Two scientists. Cueball has safety goggles, Megan has a sheet of notes.]

Cueball: *We found no link between jelly beans and acne ($p > 0.05$).*

[Panel 3. Back to the original two.]

Another: *That settles that.*
Ponytail: *I hear it's only a certain color that causes it.*
Another: *Scientists!*
Scientists: *But Miiiinecraft!*

[20 near identical small panels]

Cueball: *We found no link between purple jelly beans and acne ($p > 0.05$).*
Cueball: *We found no link between brown jelly beans and acne ($p > 0.05$).*
Cueball: *We found no link between pink jelly beans and acne ($p > 0.05$).*
Cueball: *We found no link between blue jelly beans and acne ($p > 0.05$).*
Cueball: *We found no link between teal jelly beans and acne ($p > 0.05$).*
Cueball: *We found no link between salmon jelly beans and acne ($p > 0.05$).*
Cueball: *We found no link between red jelly beans and acne ($p > 0.05$).*
Cueball: *We found no link between turquoise jelly beans and acne ($p > 0.05$).*
Cueball: *We found no link between magenta jelly beans and acne ($p > 0.05$).*
Cueball: *We found no link between yellow jelly beans and acne ($p > 0.05$).*
Cueball: *We found no link between grey jelly beans and acne ($p > 0.05$).*
Cueball: *We found no link between tan jelly beans and acne ($p > 0.05$).*
Cueball: *We found no link between cyan jelly beans and acne ($p > 0.05$).*

Cueball: *We found a link between green jelly beans and acne ($p < 0.05$).*
Off-panel: *WHOA!*
Cueball: *We found no link between mauve jelly beans and acne ($p > 0.05$).*
Cueball: *We found no link between beige jelly beans and acne ($p > 0.05$).*
Cueball: *We found no link between lilac jelly beans and acne ($p > 0.05$).*
Cueball: *We found no link between black jelly beans and acne ($p > 0.05$).*
Cueball: *We found no link between peach jelly beans and acne ($p > 0.05$).*
Cueball: *We found no link between orange jelly beans and acne ($p > 0.05$).*

[Newspaper front page.]

*NEWS Green Jelly Beans Linked To Acne! 95% Confidence. Only 5% chance of coincidence!*

# Appendix D

# Science of Tables

> It is allowed on all hands, that the primitive way of breaking eggs, before we eat them, was upon the larger end; but his present majesty's grandfather, while he was a boy, going to eat an egg, and breaking it according to the ancient practice, happened to cut one of his fingers. Whereupon the emperor his father published an edict, commanding all his subjects, upon great penalties, to break the smaller end of their eggs. The people so highly resented this law, that our histories tell us, there have been six rebellions raised on that account; wherein one emperor lost his life, and another his crown.
>
> Jonathan Swift, *Gulliver's Travels*

Consider the following 3x3 table with axis names $y_1$, $y_2$, $y_3$ horizontal on the top and $x_1$, $x_2$, $x_3$ vertical on the left:

| | $y_1$ | $y_2$ | $y_3$ |
|---|---|---|---|
| $x_1$ | 1 | 1 | 0 |
| $x_2$ | 0 | 1 | 1 |
| $x_3$ | 1 | 0 | 1 |

Can you guess what it is a table of?

It is a table of working hours of different cleaning crews at a local company. The $x_1$, $x_2$, $x_3$ are different crews and $y_1$, $y_2$, $y_3$ are certain time periods. The symbols 1, 0 are used to show working and not-working crews at those times respectively.

No, I lied to you. Actually, it is a table showing whether or not two groups of animals can be seen at a particular water source at the same time. Here $x_1, x_2, x_3, y_1, y_2, y_3$ are different animal groups and for example $x_2$ and $y_1$ have never been observed at sharing the water source as their table entry has 0.

I lied again. Actually, I did not intend this table to be a table of something. It is a table I came up with just to show that it is possible to represent different things with the same table. A table just by itself is a table, a mathematical object, nothing more. It can stay like that, or you can interpret it as depicting some relation in a science or in the world. But it can be recycled, that is, used or interpreted as something else. It can have infinitely many different interpretations.

Now having settled what tables are and how they can be interpreted, let us look at a philosophical view. According to the proponents of (the model theoretic version of) the view called the *semantic conception of theories*, all of science can be formulated using tables (including infinite ones). They claim that tables are far better than other mathematical things and we must strive to use them more. Science is in principle can be done using only tables, that is, you can reconstruct science using tables and no other mathematical objects.[68] Tables are inherently more advantageous to use than other mathematical objects. They make a big deal out of the claim that table-only science is possible and make all kinds of philosophical conclusions (though they do not always agree on them).

Actually, these philosophers do not talk about tables but the things called *models* in mathematical logic. A *model* is essentially a collection of tables any of which can be infinite. Since infinite tables cannot be drawn, you need some mathematics to express them. So models are technically a bit more complicated than tables, but in spirit they are the same. Therefore, for the sake of simplicity, I will continue talking about tables rather than models. Moreover, using the word "model" runs into the risk of confusing models in mathematical logic with the models in science (see below).

Since science is full of interpretations, be it one part of it in another or in the world, and tables have none, the interpretations of them must be supplied. But according to these table-lovers, all must be done using tables. Thus, I guess, the interpretation of a table must be given in another table. But then, you can see where this is going, the second table needs an interpretation table as well. So it is tables all the way down!

You might think that these table-lovers are just a few crackpots but you would be mistaken. Not only their list includes some eminent philosophers, but also there are other kinds as well. You see, just as there are table-lovers, there are lovers of other kinds of mathematics as well: state spaces, partial structures, axiomatic systems, and whatnot. These philosophers think that that part of mathematics is inherently (philosophically) better than the others and it is (philosophically) preferable to reconstruct science using that mathematics. This whole debate seems as if it is a tale from Gulliver's Travels (figure D.1).

Here is another such tale: All of science can be made by people running in place. Running in place is a good exercise and healthier than sitting all day. It will surely promote better science. Yes, some types of scientific activities might be harder while running, but these difficulties can easily be overcome using suitable gadgets. All science we have today could in principle be achieved by running people. (Actually, I really think that this is much more likely than writing science using only tables.) So let us be all running-science-lovers and use this to make important philosophical conclusions about science.[69]

Why is this tale ridiculous? Because running is not reflected in the practice of science in any way. There is no (realized or ideal) value that tells scientists to run when they do their work. Who cares if science can be in principle made by only running scientists? There may be health benefits for scientists during running in place while doing science but for sure not an

[68] Some proponents of semantic conception of theories seem to write as if they mean that models *are* the structure of scientific theories. Since this is obviously wrong, I will only consider the weaker view that science can be reconstructed using models.

[69] I can think of a number of *very important* philosophical consequences of running-science, but I do not want to prolong the tale and I will save you the details.

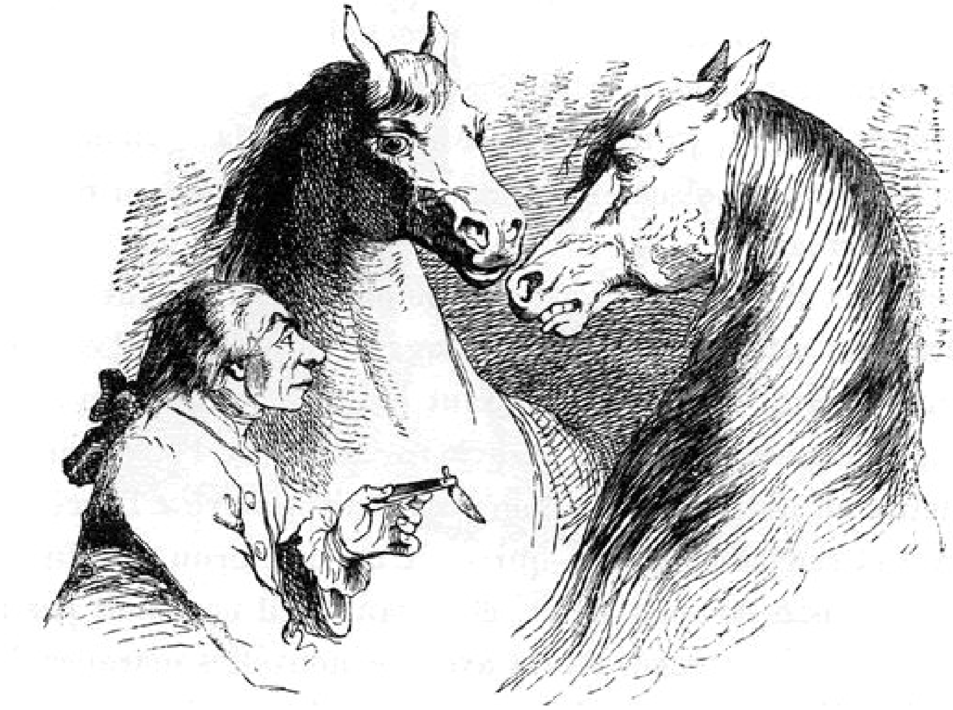


Figure D.1: Gulliver in discussion with Houyhnhnms (1856 lllustration by J. J. Grandville). Probably, they are discussing whether science should be reconstructed using only tables or only axiomatic systems. [Image in the public domain from http://commons.wikimedia.org.]

iota of improvement for science itself.

Similarly, who cares if science can be in principle formulated using only tables, state spaces, or whatever. Is there a such value in science? No. Is there anything in the practice or results of science that even suggests such a value? No. Is there a practical advantage of using only one kind of mathematics? No.

As we have seen in section 2.5, scientists consider mathematics as a useful tool. By itself mathematics is devoid of any interpretation and there is no reason that some branch of mathematics is indeed better or true to nature than any other. Scientists use mathematics as they see fit, and they use all kinds of mathematics that they find useful and simple. They might use even absurd mathematics as we have seen in section 2.5. There is no point in claiming that you can do all your repair and maintenance jobs using just a single tool (even if it is true and even if it is in principle). Tools are good or bad depending on use. Just as there is no philosophical advantage in using the spoke wrench to repair your electrical appliances, there is none in restricting science to a single mathematical tool.

Interpreting the great physicist Paul Ehrenfest's views on the role of mathematics in physics, George Johnson [2000, 67] writes that "It was dangerously easy to get carried away playing with numbers and forget that the object of the game was to talk about the real world." Supporters of the semantic conception of theories pay no heed to this lesson and get carried away with tables.

How can philosophers defend such a misinformed and misguided approach to science as the semantic conception of theories? I can only speculate to why.

First, I think that some philosophers mistake the importance of the models in science for the importance of the tables (logical models). Just because two things have the same name does not mean that they have equal importance.[70] There are all kinds of models in science that has nothing to do logical models, for example model organisms like fruit flies.

Second, they might be confusing the interpretation of models in the mathematical sense and in the representational sense. The logical (Tarskian) notion of interpretation of a logical model is just a mathematical formalism: one does not provide a representation in the world doing such an interpretation. Logical interpretation still leaves the representation problem intact.

Third and most importantly, instead of looking at how science works, these philosophers have the audacity to think that they can decide which parts of mathematics science can employ in an a priori fashion. The reconstruction of science using tables is an artificial program not sensitive to the practice in any way.

Philosophy of science needs to understand science itself, not to come up with pretend science. A narrative disconnected from the practice of science can only be science fiction, not philosophy of science.

[70] For an introduction to models in science, see [Bailer-Jones, 2002].

# Appendix E

# Archaeology

The following paragraphs which are from archaeologist Paul Bahn's book *Archaeology: A Very Short Introduction* [2012] touch on gender bias and underdetermination in archaeology.

> It is true, and worth stressing, that scholars have often treated some activities as exclusively male — notably hunting, stone toolmaking, and rock art — whereas ethnography shows that women often do these things too. Male scholars either were ignorant of this fact, or chose to ignore it, and the result was a skewed version of the past. But the feminists themselves, far from shunning this practice (while justifiably complaining about it), do exactly the same by ignoring or brushing aside examples of men carrying out 'female' activities. In any case, the realization that women made stone tools will hardly produce compelling insights. Tools tell us nothing about gender: even if some future analytical technique were to detect traces of pheromones or copulins on a stone tool, or blood residues that could be identified as male or female, this would merely tell us which sex was the last to touch it; it would reveal nothing about which sex made or habitually used it.
>
> Any detailed knowledge we have about which sex did what comes from ethnohistory and ethnography, not from archaeology. There is no alternative to reconstructing the past in this way, combining modern observations with the archaeological data. But how far can ethnography help to 'find' women in the past?
>
> The basic problem is that ethnography can usually provide a number of possible explanations for archaeological data. It has been pointed out that even a rich female burial doesn't necessarily indicate that the occupant had any power; it could merely reflect her husband's wealth (and the opposite is equally applicable to a rich male burial, of course).
>
> In fact it is hard to see how the respective roles of men, women, or indeed children (who are now starting to be noticed too!) could be determined from the tenuous evidence provided by archaeological excavation. The most important message of gender archaeology is that archaeology is about people — not just about men, and not just about women either. [Bahn, 2012, 90–91].

In the following paragraph Bahn discusses objectivity and the presence of societal values.

Of course, the presentation of the past to the world at large is a big responsibility, especially as it cannot be done objectively. We used to think that it could, that it was simply a matter of laying out our finds with some explanatory texts in glass cases or in books for the public's delectation. However, in recent years, as archaeologists have indulged in intense self-examination thanks to the interest in theory and thanks to being attacked from all sides, they have come to realize that, through their choice of artefacts, themes, and approaches, they are constantly projecting messages that reflect their own prejudices and beliefs, or those of their society, religion, politics, or of a general world view — all under the influence of the archaeologists' own backgrounds, upbringing, and education, their social status, their interests, teachers, and friends, their political and religious beliefs, and their alliances and enmities: all these things colour their version of the past, while the actual evidence often takes a back seat. [Bahn, 2012, 93–94].

# Appendix F

# Curriculum Vitae

**PERSONAL INFORMATION**

Name, Surname: Raşit Hasan Keler
Nationality: Turkish (TC)
Date and Place of Birth: 27 March 1977, Ankara
email: hasan.keler@gmail.com

**EDUCATION**

| Degree | Institution | Year of Graduation |
|---|---|---|
| Ph.D. | METU Philosophy, Ankara | 2016 |
| M.Sc. | METU Mathematics, Ankara | 2002 |
| B.Sc. | METU Mathematics, Ankara | 1999 |
| High School | Ankara Koleji | 1994 |

**WORK EXPERIENCE**

| | |
|---|---|
| 1996–2005 | Mathematics tutor |
| 2007–2008 | Popular science columnist |
| 1998–2016 | Web designer |

**FOREIGN LANGUAGES**

Native Turkish, Advanced English

**ACADEMIC INTERESTS**

Science and Values, History of Science, Philosophy of Science, Logic, Set Theory

# Appendix G

# Turkish Summary

## Giriş

Bilim felsefesinin farklı amaçları vardır (s. 1). Bu amaçlardan biri ve tezimin amacı bilimin doğasını daha iyi ve açık bir şekilde anlamaktır. Bu tez bilimdeki uygulamalarla, özellikle de bilim insanlarının ve bilimsel toplulukların bilimsel kuramlarla, araştırma gelenekleriyle (research traditions), modellerle, kullandıkları araçlarla, verilerle, hipotezlerle, teorik çerçevelerle ve kullandıkları her türlü şeyle ilgili ne tercih ettikleri hakkındadır.

Bilimsel kuramların seçiminde rol alan meziyetlerin (virtues) araştırılması tezimde hatırı sayılır ölçüde bir yer tutuyor. Bu meziyetler şu sorularla ilgilidir: Neden bilim insanları ya da bilimsel topluluklar kuramları kabul veya red eder? Bilim insanları bilimsel kuramların hangi özelliklerini beğenir, hangilerini beğenmez. İyi bir kuramın meziyetleri nelerdir? Bilim insanları kuramları nasıl değerlendirir? Neden bir kuram yerine bir başkasını seçerler? Kuram seçim ilkeleri disiplinler arasında farklı mıdır? Bunlar zamanla değişir mi? Ortak eğilimler var mıdır? Bu meziyetlerin birbirleriyle ilişkisi nedir?

Bu meziyetler sadece kuram değerlendirmede değil, aynı zamanda kuram yapımında da önemli etkenlerdir. Bilim insanları gelişigüzel yeni kuramlar yaratmaz, ancak bazı meziyetlere sahip kuramlar yaratmaya çalışırlar.

Bu tezde bilimde kullanılan bir tür unsura (örneğin kuramlara) dikkatimi kısıtlamadım ve kuramlar, modeller, araçlar, veri vb. gibi her türlü unsura baktım. Ancak tekrar tekrar "kuramlar, modeller, araçlar, veri vb." yazmak hantal olacağı için genelde bu uzun liste yerine bir tanesini kullandım. Örneğin, kullandığım "fizikçiler basit kuramları tercih eder" cümlesi "fizikçiler basit modelleri tercih eder" anlamına da gelir. Öyleyse benim kuramlar vb. hakkında yazdıklarım aslında diğer birçok unsuru da içerir. Ancak çoğu zaman bu diğer unsurları ayrıca belirtmeyeceğim.

Tabii ki, bilimlerde meziyetlerin başka türden önemi olan taraflar da vardır. Örneğin, bilim insanları için davranış kuralları dürüstlük ve doğruluk gibi farklı değerleri içerir. Değerlerin etkin olduğu başka alanlar neyin araştırılacağı, çalışmanın veya araştırmanın amaçları, soruların nasıl sorulacağı, kaynakların nasıl ayrılacağı, sonuçların nasıl yayımlanacağı ve sunulacağı, deneylerin nasıl yapılacağı ve sonlandırılacağıdır. Kuram/veri/modelleri değerlendirmekle (appraisal) ilişkili olmadıkları sürece bu değerleri incelemeyeceğim. Şu andan itibaren "değer" ve "meziyet" kelimelerini yalnızca değerlendirme ile ilgili anlamına kısıtlayacağım. Açıkça aksi belirtilmedikçe "değer" ve "meziyet" kelimelerini eş anlamlı kullanacağım.

Meziyetlere olan bu sınırlı ilgi bana kelime kullanımında başka bir kısıtlamaya gitmek için de fırsat veriyor: Bu metinde meziyeti olanlar her zaman bir kuram, veri, veya modeldir, asla insanlar veya kurumlar değildir. Bilim insanlarının (veya kurumların) motivasyonlarından ve amaçlarından bahsedeceğim ancak değer atfını sadece kuram, veri gibi bilimsel unsurlara (bilimsel çalışmanın verimliliği, verilerin basitliği gibi) kısıtlayacağım. Bu kullandığım dildeki ikililik, bilimdeki bütün değer merkezli sorunlarla ilgilenmediğimi açıkca ortaya koyuyor.

Bilimde kuram seçiminde değerlerin önemi, sınırlı bir dereceye kadar Kuhn öncesi dönemde bilim felsefecileri tarafından kabul edilmişti. Daha sonra Kuhn'un The Structure of Scientific Revolutions [1962] ve Objectivity, Value Judgment, and Theory Choice [1977a] makaleleri bu değerlerin önemini açık bir şekilde ortaya koydu. Kuhn sonrası dönemde Kuhn'un çok popüler olmasıyla beraber bilimde değerler hakkında tartışmalarda büyük bir artış oldu. Ama değerler üzerine yazıların çoğu aslında kuram seçilimi ile ilgili olmayan değerler ile ilgili konular üzerinde durdular. Değerler hakkında söylenebilecek bütün şeylerin Kuhn'dan beri ifade edildiğini düşünmek yanıltıcı olur. Bir çok sorun halen sonuçlandırılmamıştır.

Değerlerle beraber bilimdeki pratiğe bakmak bilimi anlamak için önemli olsa da maalesef felsefeciler arasında bunlara gerekli dikkat çoğu zaman verilmemiştir. Bu eksiklik bu tezde uygulamalar ve değerlere daha kapsamlı ve ayrıntılı bir bakışla dolduruluyor. Bilimdeki uygulamalar ve değerleri inceledikten sonra çeşitli felsefi tartışmalarla ilgileniyorum. Bilim felsefesi sorunlarının çoğuna bilimdeki pratikler ve değerler ışığında bakarsak nasıl ilerleme kaydedilebileceğini göreceğiz.

Bilim ve bilimsel faaliyetleri araştırmak için farklı felsefi yollar vardır. Tarihyazımı çalışmaları, sosyolojik incelemeler, retorik analizi, feminist çalışmalar, Popper'ın yanlışlanabilirlik yöntemi ve her türlü diğer yaklaşımlar bize bilimin daha iyi anlaşılmasını sağladığı iddiasındadır. Lakatos [1978b] bu farklı felsefi yaklaşımları birbirleri ile rekabet halinde görür. Bence bu yaklaşımlar rakip olsa bile mutlaka birbirlerine ters düşmek zorunda değildir. Farklı yöntemler bilimi daha iyi anlamamıza yol açabilir.

Ben bilimi anlamak için değer analizi (value analysis) adını verdiğim başka bir yaklaşım öneriyorum. Bu yöntem bir kuramın sahip olduğu özel değerlerini belirlemeyi, bunları genel değerlerle karşılaştırmayı, diğer kuramların değerleriyle karşılaştırmayı ve bilimsel faaliyetlerle ilişkisini kurmayı gerektirir. Değer analizi, bilim analizine münhasır değildir ve başka alanlarda da uygulanabilir. Ama burada ben bir iki ufak örnek dışında bilime uygulamakla yetineceğim.

Bütün felsefi yaklaşımlar gibi değer analizinin de zayıf ve kuvvetli olduğu taraflar vardır. Bu tezde daha çok kuvvetli taraflarından bahsetsem de bazı zayıf yönlerine de değiniyorum. Ancak yukarıda bahsettiğim gibi felsefenin bir tek yaklaşımla yetinmemesi gerektiğini düşünüyorum. Değer analizinin zayıf tarafları diğer yaklaşımlarla desteklenebilir.

Farklı yaklaşımların beraberce bilimi anlamakta kullanılması bunların sonuçlarını birleştirmekten ibaret değildir. Bir yöntemin sonucu başka birinin verisi olabilir. Örneğin, değer analizi değerlerini belirlemek için tarihsel ve sosyolojik çalışmalardan yararlanabilir (ve bu ben bu metinde sıkça yararlanıyorum). Bu bir özyinelemeli bir işlemdir ve farklı yaklaşımlar sonsuza kadar birbirlerini etkileyebilirler.

Bilimi anlamak için tüm bu farklı analizlere bakmak ve bunları birleştirmek gerekir. Hem bazı felsefeciler hem de bazı sosyologlar arasında değer analizin ihmal edildiğine sıkça tanık oluyoruz. Bir taraf, tek geçerli değerleri bulgusal (evidential) değerler olarak görür. Diğer taraf ise sadece toplumsal değerleri kabul eder. Ben bu seçimlerin bilimden yeterince haberdar bir şekilde yapıldığını sanmıyorum. Tek bir değeri kabul etmek bu iki dalın da bilime yaklaşımını çarpıtmıştır. Benim bilime yaklaşımım bu iki ters görüşten faydanlansa da onlar arasında bir uzlaşma arayışı değildir.

Bilimin sağlıklı bir analizi için yapılması gereken bilimdeki değerleri varsaymak değil, tam aksine inceleyip görmektir. Bilimi anlamak için bilime ve her türlü bilimsel faaliyetlere bakmak gerektiğini savunuyorum.

Toparlarsak, bu tezin amacı bilimindeki değerleri ve pratikleri incelemek, "bilimin nasıl işlediğine gerçekten bakalım" felsefesini öne çıkarmak, ve bunların bilim felsefesindeki farklı problemlere etkilerini ortaya koymaktır.

Farklı yazarlar tarafından bilimde olduğu iddia edilen değerlerin listesini yaparsak basitlik, açıklık, inandırıcılık, otorite ile uyum, sağduyu ile uyum, kullanışlılık, güzellik, gibi değerlerden oluşan uzun bir liste elde ederiz (bkz. s. 7). Ancak bunun gibi listeler felsefeciler için hatırlatıcı kaba listelerden öte gidemez. Her bilimsel geleneğin (research tradition) kendine has değerleri vardır ve başka geleneklerle bunlardan bazılarını farklı derecelerle paylaşabilir. Ancak bir gelenekte geçerli olan değerleri o geleneği incelemeden bilemeyiz. Aynı şekilde bilimsel faaliyetler de inceleme sonucu bilinebilir. Dolayısıyla ikinci bölümde verdiğim örnek vakalar bilimdeki farklı değer ve faaliyetlere dikkat çekmesi açısından tezimin temelini oluşturur.

## Bilimden Örnek Vakalar

Bu bölüm on dört farklı örnekten oluşuyor.

İlk örnek matematiğin alt dalı olan kümeler kuramındaki zorlama (forcing) tekniğine ait. Dünyada bu teknik iki farklı dil sistemi ile ifade ediliyor: Cohen (Boolean) ve İsrail notasyonu. Bu durum bir kaç açıdan dikkat çekici. Ülkesel farklılıkların bilimdeki yeri hakkında fikir veriyor. Bilimdeki her seçimin önemli bir değer çekişmesine dayanmadığını gösteriyor. Ayrıca bir kuramın farklı dillerle ifade edilebileceğini ortaya koyuyor.

İkinci örnek linguistikten alınmış ve İngilizce gramerdeki "zero article" kavramı ile ilgili. Bu örnek var olmayan bir objeye varmış gibi referans yapıldığı bir durumu sergiliyor. Dilbilimciler psikolojik ve pedagojik nedenlerle bu kullanımı tercih ediyorlar. Bir kurguyla çalışmayı bu meziyetlerinden ötürü gerçek duruma tercih ediyorlar.

Üçüncü örnek SNC tipi meteorlarla ilgili ve felsefi açıdan bir çok ilginç yanı var. Kuram yükü (theory-ladenness), sağduyusal ve kabul edilmiş varsayımların bilimdeki önemi, çarpışan kuramlar ve veriler gibi farklı noktalara dikkat çekiyor.

Dördüncü örnek fizikte quasi-parçacıkların rolünü inceliyor. Matematiksel olarak basit ve kullanışlı olmaları nedeniyle fizikte yer aldıklarını gösteriyor.

Beşinci örnek kuantum fiziğine Dirac tarafından sokulan delta fonksiyonu ve daha yakın zamanlarda sicim kuramında geçen bir denklemle ilgili. Fizikçilerin matematiğe pragmatik yaklaşmalarını ve işlevsel bulurlarsa doğru olmayan matematik kullanmaktan bile kaçınmayacaklarını gösteriyor.

Altıncı örnek Mars'a ait olduğu düşünülen canlı mikro-fosilleri ile ilgili. Tutumluluk (parsimony) ve basitlik (simplicity) değerlerinin çatışmasını gözler önüne sergiliyor. Hipotezlerin değerlendirilmesinde riskin de rol alabileceğine dikkat çekiyor.

Yedinci örnek nükleer fizikteki çekirdek modellerinin analojik yapısına dikkat çekiyor. Bu modellerin kapsamlarının sınırlı olması ve birbirleriyle uyuşmadıkları yanlar olması fizikçilerin daha genel bir model bulmaya çalışmalarına rol açıyor. Ayrıca bilim insanlarının her zaman kuramlarına harfi harfine inanmadıklarını gösteriyor.

Sekizinci örnek bisikletçilerin kask takmasının zorunlu kılınmasının faydalı mı olacağı sorusunun ne kadar karmaşık bir soru olduğunu gösteriyor. Sadece verilere bakarak bu soruya cevap vermek mümkün görünmüyor. Kuramın veriler tarafından belirlenememesi (underdetermination of theory by evidence) olgusuna bağlanıyor.

Bir sonraki örnek insanımsı maymunların zihinsel kapasiteleriyle ilgili bir tartışmada tutumluluk ve basitlik değerlerinin çatışmasına dikkat çekiyor. Verilerin dolaylı yoldan ancak etkili olabileceği tespit ediliyor.

Onuncu altbölümde bilim tarihinden iki sahtekarlık vakası üzerinde duruluyor. İlki 1912'de bulunan Piltdown kafatası ve ikincisi yakın zamanda sosyal psikolog Diederik Stapel'in uydurma makaleleri hakkında. Bilim felsefesi açısından ilginç yanları verilerin nasıl kuramlar ışığında yorumlandığını ve kuramların da nasıl zamanın arkaplanından etkilendiğini göstermesi.

Sıradaki altbölüm kuarkların ortaya atılması ve kabulu arasında geçen döneme bakıyor. Zaman içinde parçacık fizikçilerinin kuarklara bakışlarındaki değişiklikleri ve kuarkların matematiksel varlık (mathematical entity) özelliklerini kaybetmesini anlatıyor. Bu vakanın felsefi açıdan dikkat çekici bir çok yanlarından bazıları şunlar: (1) arkaplan veya rehberlik eden varsayımların (guiding assumptions) güçlü etkileri olması; (2) matematiksel varlık olmanın kuramdan bağımsız olması; (3) icat-keşif ayrımının erimesi; (4) soyut fiziğin daha fazla kuram yüklü olması; (5) fizikçilerin kuramlarına bir mecaz olarak bakmaları.

Nöron doktirininin doğumu ve yüzyıl boyunca nasıl değiştiğini konu alıyor bir sonraki örnek. Benim yayılmış kuram inşaası (extended building) dediğim uzun ve karmaşık bir süreç sonucu kuram oluşuyor ve değişiyor. Veri ve kuram arasındaki karmaşık ilişkiye ve de bilimde indirgemeciliğin rolüne dikkat çekiyor. Ayrıca bilim insanlarının kuramlarının limitlerini ve faydalarını zaman içinde tekrar gözden geçirerek ona göre davrandıklarını gösteriyor.

On üçüncü örnek tıbbın kara lekesi sayabileceğimiz ilaç geliştirilmesine bakıyor. Önyargıların ve sapmaların (bias) ne yazık ki yoğun bulunduğu bu dalda ki sadece bir problemi ele alıyor: yayın sapması (publication bias). Bu sorunun farklı çözümlerinin bilinmesine rağmen nasıl ortadan kaldırılmadığına değiniyor.

Son örnek jeolojideki kıta kayması (continental drift) kuramının Avrupa ve Amerika'da farklı karşılanması ile ilgili. Kuramın Wegener tarafından ortaya konulmasından sonraki yaklaşık 50 yıllık sürede Avrupalı jeologlar bu kurama Amerikalı meslektaşlarına göre çok daha sıcak ve kabul edilebilir bakıyorlar. Neden? Naomi Oreskes'in bu soruya cevabı iki kıtadaki jeologların sadece metodolojisinde değil, aynı zamanda bilim anlayışında önemli farklılıklar olduğunu ortaya koyuyor. Bu örnek sosyokültürel farklılıkların bilimde ne kadar etkili olabileceğini gösteriyor.

## Değerlerin Felsefesi

Bu bölümde bilimsel değerlerin farklı açılardan felsefi incelenmesini yapıyorum.

Birinci altbölümde değerlerle ilgili bir çok kısa tespit yapıyorum.

(a) Öncelikle tek ve birleşik bir bilimden bahsetmek yanlış olur. Bilimde değer ve pratikler farklı araştırma geleneklerinde değişiklikler gösteriyor. Bilim zengin ve çeşitli bir yelpaze sunuyor bize. Bilimi homojen ve bir bütün gibi görmek bir çok felsefi soruna yol açtığını göreceğiz.

(b) Bilimde değerleri göz önüne alırken somut (concrete) ve genel (general) değerler arasındaki ayrım önemlidir. Genel değerler "bilim insanları basit kuramları tercih eder" savındaki gibi soyut kullanılan değerlerdir. Somut değerler ise belirli bir kuramın meziyetini dile getirirler. Bilimde etkili olan değerler somut değerlerdir.

(c) Değerler siyah-beyaz değildir. Yani bir çok dereceleri ve çeşitleri olabilir. Bilim insanları bir kuramın potansiyel değerleriyle de ilgilenebilir.

(d) Değerlerin oluşumunda sosyal ve kültürel faktörler, biyolojik faktörler, çevresel faktörler ve tarihsel faktörler etkilidir. Bunlar bir somut değerin oluşumunda o duruma has bir şekilde yol açarlar.

(e) Bilimsel değerlerin mümkün olan bir kategorizasyonu şudur: pragmatik, kuramsal, epistemik, sosyal. Ancak bu ayrım tutarlı bir şekilde yapılamaz ve unutulması daha iyi olur.

İkinci altbölüm değer analizi ve ilgili bazı konuları ele alıyor.

(a) Bazı felsefeciler bilimdeki değerlerin ne olduğuna önsel (a priori) bir şekilde karar veriyorlar. Bu tutum hiç gerçekçi olmayan bir bilim felsefesine yol açıyor. Bu yaklaşımı değer analizi (value analysis) adını verdiğim yaklaşımla değiştirmek lazım. Dar anlamıyla değer analizi bir kuramla veya araştırma programıyla ilgili değerleri araştırmayı içerir. Bunun için sosyolojik ve tarihsel metodlara da başvurabilir. Önemli olan değerleri önceden var saymamaktır. Bu dar anlamıyla değer analizini diğer kuram ve araştırma geleneklerinin değerleriyle ilişkilendirerek genişletebiliriz. Bulunan somut değerler başka kuramların somut değerleri ile karşılaştırılabileceği gibi genel değerlerle de bağlantılandırabilir. Bilimsel faaliyetler ve değerler arasındaki ilişkiler tespit edilebilir. Bu daha genel projeye geniş anlamda değer analizi diyebiliriz. Elde edilen bulgu ve sezgiler bilim felsefesindeki sorunlarla yüzleşmek için kullanılabilir.

(b) Arkaplan değerleri bir kuramın seçiminde rol almayan ama yine de o kuramda mevcut olan değerlerdir. Bunları bulmak hem kendi başına hem de değer analizinin sağlığı açısından önemlidir. Etkili olan değerleri belirlemek için arkaplan değerlerini ayıklamak gerekir. Bazen bir değer bir başkasına bağlı olarak yaşayabilir. Mesela matematiksel olarak basit fizik kuramları ileride diğer kuramlara daha kolay gömülebilirler, yani daha dinamiktirler. Ancak her zaman hangi değerlerin esas, hangilerinin ikincil olduğunu tespit etmek mümkün olmayabilir.

(c) Değer analizinin dikkate alması gereken bir başka nokta ise değerlerin zaman içindeki değişimidir. Şu an ağır basan bir değer ileride zayıflayabilir. Kuramın tamamen yeni bir değeri olabilir. Bir arkaplan değeri su yüzüne çıkabilir. Kuramın şu anda var olan değerleriyle kabul veya reddedildiği zamandaki değerlerini karıştırmamak gerek. Mesela bir çok fizik kuramı kullanışlı olmasından önce başka nedenlerle kabul edilmiştir.

Üçüncü altbölüm iki konuya eğiliyor.

(a) Analojiler (benzeşmeler) bir alandaki kavramları diğer bir alana uygular. Çoğu zaman yerleşmiş bir kuram veya modelin bazı parçaları yeni bir kuram veya model içinde kullanılır. Analojilerin bizi ilgilendiren tarafı değerleri taşıyor olmalarıdır. Bir V özelliğine sahip A kaynak modeli ile B hedef modeli arasında analoji kurulmuşsa V özelliği B'ye aktarılabilir. Mesela basit bir kaynak model kullanılıyorsa hedef modelin de basit olması muhtemeldir. Bir analojide genelde değerler kaynak modelden hedef modele aktarılsa da nadiren tersi de olabilir. Mesela hedef modelde faydalı bulunan bazı unsurlar kaynak modelde de işe yarayabilir. Analojiler değer transferi için önemli bir araçtır ve hangi analojilerde hangi değerlerin aktarıldığı araştırılarak bulunabilir.

(b) Bazen bir model veya kuram bir başkasını içerebilir. Analojilerin tersine bu tür kapsamalar her zaman değer transferine neden olmaz. Mesela karmaşık bir modelin basit bir alt modeli, diğer kuramlarla çelişen bir kuramın tutarlı bir altkuramı olabilir. Bu konu sicim kuralının günün birinde standart modeli içerecek şekilde gelişeceği umuduyla da ilgilidir. Bir iddiaya göre bu kapsama sicim kuramına standart modelin bazı değerlerini aktaracaktır. Ancak bu genel bir kural olarak doğru değildir ve sicim kuramının standart modeli nasıl kapsayacağına bağlıdır.

Dördüncü altbölümde kuram, veri ve keşif arasındaki karmaşık ilişkilere bakıyorum.

(a) Öncelikle Thomas Nickles'ın bahsettiği "ilk seçilim" olgusunu ele alalım. Bilim insanları her zaman sonsuz kuramlar uzayından sadece sonlu ve az sayıda olanıyla ilgilenebileceği için, bu az sayıdaki kurama nasıl eriştikleri en sonda sahip olacakları kuram üzerinde önemli bir etkiye sahiptir. Bilim dalının tarihçesi, yöntemleri, ilgi alanları, kuramlar arası ilişkiler hep Nickles'ın ilk aşamasını etkiler. Ancak göreceğimiz gibi keşfin kuram üzerindeki etkisi bu ilk aşamayla sınırlı değildir.

(b) Bir tanıma göre bir veri cümlesinin kuram yüklü (theory-laden) olması onun ancak bazı kuramlar ışığında anlamlı gelmesi demektir. Ben kuram yüklü olmayı bu tanımdan biraz daha genel olarak gözlem ve verilerin kuramsal varsayımlardan etkilenmesi anlamında da kullanıyorum. Bunun bir sonucu kuramsal varsayımların bir grup gözlemi diğer bir gruba göre teorik olarak daha çekici olduğu için tercih etmemize neden olabilir. Bir başka sonuç ise yarışan kuramlar arasında seçim yapmak için veri toplamak beklenenden daha az başarılı olabilir çünkü bilim insanları verileri destekledikleri kurama uygun bir şekilde yorumlayabilirler. Bunun örneği olarak bilim tarihindeki cansızdan canlı oluşumu (spontaneous generation) kuramını sunuyorum.

(c) Gözlemsel ve deneysel veriler bazen aşırı derecede kuram yüklüdürler: uygun kuramlar yerinde olmadan önce bu verilerin kimse farkında olmayabilir veya ihmal edilmiş olabilirler. Mesela fizikçiler pozitron'un Anderson tarafından keşfinden önce sonradan pozitron olarak yorumlayacakları parçacıkların izine rastladıklarında bunu deneysel hataya bağlıyorlardı.

(d) Bir kuramla ilgili gözlem ve veriler her zaman sorunsuz bir şekilde bir araya getirilemez. Bunun nedeni veriler ya birbiriyle ya da başla kuramlarla çatışabilir. Bu gibi durumlarda kuram çatışmayı çözebilecek şekilde verileri yorumlamak zorundadır ancak bunun başarılabileceğinin garantisi yoktur.

(e) "Bilim insanları bir kuram ortaya atar ve bu bağımsız bir şekilde test edilir" görüşünün yanlış olduğunu gösteren durmlardan biri benim yayılmış kuram inşaası (extended building) dediğim durumdur. Bazen kuramlar uzun ve karmaşık bir süreç içinde oluşur ve bu sırada veri, kuram, deney, keşif, icat birbirlerini etkileyerek ve hatta başka kuramların da etkisinde kalarak değişir.

(f) Bazen keşif ve kuram yükü bilimin rasyonalitesi konusuna bağlanır. İddiaya göre bilimde pekiştirme (justification) keşiften (discovery) tamamen ayrıdır ve özellikle de pekiştirme sadece bağımsız verilere dayanıyorsa bilim rasyoneldir. Bu rasyonalite anlayışına katılmadığımı aşağıda belirteceğim. Daha önemlisi, bu bölümde gördük ki pekiştirme ve keşif ayrık değildir ve veriler de kuramdan bağımsız değildir.

Bu bölümün son altbölümü sosyal değerlerle ilgili.

(a) Bazı felsefecilerin ve bilim insanlarının bilimde sosyal değerler olmasına karşı duydukları kaygıyı anlamak zor değil. Eğer bilim soyopolitik yapı ve kuvvetlere dayanırsa, bu yapılar ve kuvvetler değişince ne olur? Muhtemeldir ki, bilim de beraber değişir. Bu ciddi bir sorun oluşturur çünkü bilimin sağlam ve her toplumda yinelenebilir olması beklenir. Bu nedenledir ki sosyal değerlerden muaf bir bilim anlayışı bir çok felsefeciye çekici gelmiştir. Günümüzde böyle bir bilim anlayışının bir hayalden öte olmadığını kabul etsek dahi hala cevaplanması gereken sorular kalmıştır. Bilimdeki sosyal değerleri inceleyip yapısını anlamak gerekir. Özellikle hangilerinin bilimin sağlamlığı önünde bir engel oluşturduğunu incelemek gerekir.

(b) Bilimde sosyal nedenlerle kabul edildiği ilk akla gelen geleneklerdir (conventions). Metinde bunları inceleyip çoğu zaman bilimin sağlamlığına bir engel teşkil etmediklerini tespit ediyorum. Ancak istisnaları olabilir.

(c) Bilimin bir toplumdan diğerine (veya bir araştırma geleneğinden rakibine) göre değişmesinin ve sağlamlığını kaybetmesini en önemli nedeni temel varsayımlardır. Farklı temel varsıyımlar farklı bilimlere yol açabilir. Problem eğer kabul edilen varsayımlar birbiriyle çelişiyor ise artar.

(d) Bilim gruplarının sahip olduğu felsefi varsayımlar da bilimin ayrılmasına neden olabilir.

(e) Farklı bilim yapma metodları da bir toplumdaki bilimi farklılaştırabilir. Yukarıda anlattığım kıta kayması kuramında farklı ekollerin nasıl oluştuğu buna örnektir.

(f) Bir problem o kadar karmaşık olabilir ki ilgili faktörleri bir şekilde azaltmadan yol alınamıyabilir. (Bisiklet kaskları ile ilgili bölümde bunu gördük.) Belirsizlikleri azaltmak için bazı seçimler yapmak gerekebilir. Hangi değişkenlerin ne ölçüde göz önüne alınacağı böyle durumlarda sosyopolitik kararlarla belli olabilir. Bu konu daha genel bir konu olan önem konusuna bağlanabilir. Çoğu zaman bilim insanları yaptığı işleri, farklı faktörleri önem sırasına dizmek, verdikleri önem doğrultusunda davranmak zorunda kalırlar. Ancak önem sosyopolitik olarak belirlenir ve toplumdan topluma değişebilir.

(g) Koruma biyolojisi (conservation biology) gibi bazı alanlarda normatif öğeler bilimin ayrılmaz bir parçasını oluşturur.

(h) Bilim insanları çoğu zaman bilimsel hipotezlerin nihai kabulunu beklemeden kullanmaları gerekir. Bu temel varsayımlar konusunun özel bir alt durumu olarak görünebilir.

(i) Özel çıkar grupları sosyoekonomik ve politik güçlerini bilime baskı uygulamak için kullanabilirler.

(j) Sonuç olarak sosyal değerler bilimin bir parçasıdır. Ama felsefeci ve sosyologların işi bunu tespit ile bitmez. Bilimdeki sosyal değerlerle ilgili daha çok bilgi edinilmelidir. Hangilerinin bilimin sağlamlığı önünde engel veya potansiyel engel olabileceğini araştırmak gerekir. Eğer engel olanları varsa da bunun nasıl ortadan kaldırabileceğimiz bulunmalıdır.

## Değerlerin Çatışması

İdealde bütün değerler birbiriyle uyumludur. Bir kuram aynı anda basit, tutumlu (parsimonious), yararlı, güzel, vb. olabilir. Ancak pratikte kuramlar bütün bu değerlere sahip olamaz. Değerler çatışır. Bu bölümde çatışan değerlere bakacağız.

İlk altbölüm tutumluluk (parsimony) ve basitlik (simplicity) değerlerinin çarpısması hakkında.

(a) Bu çarpışmanın nedeni tutumluluğun düşük seviyeli (low level) hipotezleri desteklerken basitliğin yüksek seviyeli (high level) hipotezleri desteklemesidir. Bu çatışma nasıl sonuçlanır? Önemli olan verilen düşük seviyeli açıklamaların yapaylığıdır. Eğer arka arkaya yapay açıklamalar veriliyorsa bunlara genel bir açıklama tercih edilebilir.

(b) İlgili bir konuda bilimin sonlandırılması (science-stopping) konusudur. Yüksek seviyeli bir hipotezi kabul etmek alternatif düşük seviyeli hipotezleri aramayı sonlandırır. Bu aceleci olabileceği için hemen basit hipotezleri kabul etmemek için bilimde farklı höristikler vardır.

İkinci altbölüm hassaslık (precision) ile bağlantılı bazı değerlere bakıyor.

(a) Hassaslık ve netlik bazen çarpışabilir. Detaylar veya hesaplar içinde boğulmaya başlandıkça netlik veya açıklık kaybolmaya başlar.

(b) Hassaslık ve kaynak kullanımı arasında bir bağlantı vardır. Ancak eldeki kaynakların gerektirdiği kadar hassas olunabilir.

(c) Kesinlik (accuracy) ve elverişlilik (convenience) bazen çatışabilir. Mesela bir ölçümü daha kesin yapmak ölçüm metodlarını zorlaştırabilir.

Üçüncü altbölüm tekrarlanabilirlik (reproducibility) üzerine. Son yıllarda tıp ve psikolojinin bazı dallarında yapılan deneylerin tekrar yapıldığında aynı sonuçları vermemesi bu meseleyi önemli hale getirdi.

(a) Öncelikle her tekrarlanamayan deney kötüye alamet olmak zorunda değildir. Şu ana kadar bilim insanlarının farkında olmadığı faktörleri bize öğretebilir.

(b) Bilimdeki sahtekarlıklar genelde tekrarlanabilirlik ile sonuçlanır.

(c) İlaç endüstrisinde gördüğümüz gibi temel güdü para kazanmak olursa tekrarlanabilirlik ikinci plana atılabilir.

(d) Yenilikçi fikirler ve yeni teknolojiler sorunsuz kullanılmaya başlanana kadar tekrarlanmalarında sorun olabilir.

(e) Bazen elde bulunan bir veri seti bilinmeyen bağıntılar için taranabilir. Elde edilen sonuçlar bize yeni şeyler öğretebileceği gibi bu veri setine has, tekrarlanamaz öğeler de verebilir.

(f) Tekrarlanabilirliğin elden gittiği durumlarda kuramın uygulama alanı kısıtlanarak bu sorun çözülebilir.

Dördüncü altbölüm Kuhn'un "temel gerilim" (essential tension) adını taktığı bilimdeki bir çatışma ile ilgili.

(a) Yenilikçi ve gelenekçi düşünceler arasında bir "temel gerilim" vardır. Gelenekçi düşünce bize zamanın testinden geçmiş ve güvenilir olan fikirlere ve metodlara yakın durmamızı söyler. Zaman, kaynak ve insan gücünden tasarrufa yol açar. Bilinmeyen durumlarla karşılaşıldığında bize hakkından gelmemiz için dolu bir alet çantası verir. Yenilikçilik ise bize eskinin yeterli olmadığı durumlarla başetmemizi sağlayabilir. Bu gerilim esas metinde anlattığım gibi değerlerle de ilgilidir.

(b) Kuhn'un anlattığı gibi yenilikçi yaklaşımlar bilimsel devrimler sırasında sıkça görülebilir. Ancak yenilikçiliği sadece devrimlere ait görmek yanlış olur.

Beşinci ve son altbölüm verilerin ve deneylerin kalitesi (quality) konusuna giriyor.

(a) Mümkün olduğu kadar fazla ve kaliteli veri sahibi olmak istiyebiliriz. Fakat kaliteyi arttırmak beraberinde teknik veya ekonomik zorluklar getirebilir. Amacımıza göre kaliteden fire vermek zorunda kalabiliriz.

(b) Bilimde bazen ölçmek istediğimiz bir değişken yerine onunla bağıntılı "vekil" (surrogate) denilen başka bir değişken ölçmemiz gerekebilir. Bunlar daha kolay ölçülebildiği veya daha uygun olduğu için tercih edilebilirler. Genelde vekiller esas değişkeni ölçmek kadar kaliteli olmaz. Vekillerin ne kadar başarılı olduklarının iki kriteri vardır: esas ve vekilin bağıntısını veren kuram ve vekilin ölçülebilme kolaylığı.

(c) Anektodlar bir araştırma geleneğinin güvendiği bilgi edinme metodları dışında elde edilen verilerdir. Kalitesi düşük olsa da faydalı olabilirler. Bunların nasıl değerlendirileceği araştırma geleneğine göre değişir.

## Bilim Felsefesi

Bilim felsefesinin bir çok sorununu bilimsel faaliyetleri ve değerleri daha yakından inceleyerek göz önüne alırsak çözümüne yaklaşabiliriz. Beş altbölümün her birinde tezin önceki bölümlerinde geliştirdiğim fikirler ışığında bilim felsefesinin problemlerini açıklığa kavuşturacağım.

İlk altbölümde bilimin amacı meselesini ele alıyorum.

(a) Öncelikle "bilimin amacı" denilen şeyin ne demek olduğunu inceliyorum. Vardığım sonuç şu: bütün bilimlerde ortak bir amaç varsa buna bilimin amacı diyebiliriz. Ancak böyle bir amaç olup olmadığına bakmadan önce geliştirmem gereken fikirler var. Soruya geri döneceğim.

(b) Bilimde zayıf ve kuvvetli amaçlar arasında bir ayrım yapmak gerek. Kuvvetli amaçlar yapılan araştırmanın (research) konusunu tespit eden amaçlardır. Bilim insanlarının işlerini belirler. Diğer araştırmalardan farkını çizer. Zayıf amaçlar ise bilim insanlarının motivasyonları gibi dolaylı yoldan araştırmaya yol veren ve başka araştırmalarla da paylaşabilen amaçlardır. Dolayısıyla bilim felsefecileri için önemli olan bilimi anlamak ise önem vermeleri gereken kuvvetli amaçlardır.

(c) Bilimin (tek) amacı zayıf bir amaç olamaz çünkü bunlar hem sayıca çoktur hem de bilimsek faaliyetlere etkileri önemsizdir.

(d) Bilimin (tek) amacı kuvvetli bir amaç olmaz çünkü bunlar daldan dala farklılık gösterir.

(e) Zayıf amaçlar plastikdir: verilen herhangibi bir bilimsel kuram veya sonuca yolaçan çok sayıda zayıf amaç olabilirdi. Yani zayıf bir amacın neden olduğu bir sonucun sadece bu zayıf amaçla olabileceği düşünülmemelidir.

İkinci altbölümde nesnellik (objektiflik) ve sapmasızlık (unbiasedness) konuları üzerinde duruyorum.

(a) Nesnellik iki nedenle bilim felsefecilerinin bir kenera atıp unutmaları gereken bir kavramdır. Öncelikle nesnellik net bir kavram değildir. Bin farklı kişi bin farklı şekilde tanımlar. İkincisi, bazı bilim felsefesi gelenekleri tarafından bilimin başına yapay bir gereklilik kriteri olarak sarılmıştır. Bunun sorunlarına metinde değiniyorum. Nesnellik kavramının yerini alması gerek kavram sapmasızlıktır.

(b) Bilimde sapmalar mutlaka vardır. Dalına göre çeşidi ve miktarı değişir ve bilim dalını incelemeden bunları bilemeyiz.

(c) Hem bilimdeki sapmalar hem de bizim sapmalar üzerine olan düşüncelerimiz zaman içinde değişmiştir.

(d) Bilimde sapmalar azalıyor mu? Şüphe yoktur ki sapmalar hakkında gitgide daha çok bilgi sahibi oluyoruz. Ancak sapmaları kaldıracak araçlara sahip olmak ile bu araçları kullanıp sapmaları kaldırmak faklı şeylerdir. Bu araştırılması gereken bir sorudur.

(e) Metinde açıklanan nedenlerden ötürü bilim sapmasız olmayı amaçlamalıdır.

(f) Bilim felsefesi de bilimler gibi sapmalıdırlar. Hatta bilim felsefesinin yirminci yüzyıldaki durumuna bakarsak ne yazık ki sapmaların pek fazla olduğunu görürüz. Felsefeciler de kendi dallarını sapmasız yapmayı amaçlamalıdırlar.

(g) Sapmasızlık asla tam olarak erişilemez ama bu yine de onu amaçlamamıza engel olmaz. Katedilecek her yol faydalıdır.

Üçüncü altbölümde rasyonalite meselesini tartışıyorum.

(a) Rasyonalite ile bilim bazen özdeş tutuluyor. Bu bilime yeni bir isim vermekten başka bir işe yaramaz. Rasyonalitenin bilimden bağımsız bir tanımı olmalı ve daha sonra bu tanım farklı bilimlerde geçerli mi diye bakılmalıdır. Bilimin rasyonel olup olmaması a priori bir mesele olmamalı.

(b) Rasyonalitenin bir tanımı şudur: bilgi edinmek için (eldeki) en iyi metod, kuram ve uygulamaları kullanmak. Bu tanımı incelediğimde aklımdakine yakın buluyorum ancak bazı şeyleri hasıraltı ettiğini görüyoruz. Bunları açıklığa kavuşturduğumuzda ortaya çıkan rasyonalite olgusu bütün bilimi ne rasyonel ne de rasyonelliğini kesinlikle bilinebilir kılıyor.

(c) Neyin rasyonel olduğunu anlamak için amaç ve değer analizi yapmak zorunludur.

Dördüncü altbölündeki konu kuramın veriler tarafından belirlenememesi (underdetermination of theory by evidence), kısaca KVTB, olgusu.

(a) KVTB verilerin farklı kuramları eşit miktarda pekiştirmesinden (justification) kurtulamıyacağımızı söyler. Bu doğruysa bazen eldeki verilerin pekiştirdiği farklı farklı kuramlar bulabiliriz. Sonucunda pekiştirme sadece verilerle yapılamaz.

(b) Metinde tartıştığım ve örneklendirdiğim üzere KVTB gerçek bir olgudur. Bilimde veriler farklı kuramları eşit destekleyebilir. Hatta bu istisnai bir durum değildir.

(c) KVTB tartışması kısır bir tartışmadır. Veriler tek önemli değer değildir. Kuramların bir çok meziyeti varken sadece birine saplanıp diğerlerini görememenin sonucu olan bir tartışmadır.

Beşinci ve son altbölümde üç yanlı ve sorunlu görüşe değiniyorum.

(a) Pekiştirme kuramı (confirmation theory) Bilim felsefesindeki kuramların tek meziyetinin veriler tarafından desteklenmesi olduğunu düşünen bir grup felsefecinin bu destek bağıntısını matematiksel bir şekilde ifade etme çabalarıdır. Bu nafile bir çabadır. Bilimler çok zengin ve farklıdır ki bir formülle ifade edilemezler. Dahası kuram seçiminde kurallar değil değerler vardır. Dahası somut (concrete) ve genel (general) değerler arasında kuram seçimi için önemli olan somut değerlerdir. Dolayısıyla genel bir pekiştirme kuramı yapılamaz.

(b) Benim "bilim polisi" adını taktığım bazı felsefeciler neyin bilim olup olmadığını nesnellik, rasyonalite, yanlışlanabilirlik gibi kriterlerle belirlemek istiyorlar. Bu yapay kriterlerin zararlı olduğunu anlatıyorum. Bırakın işe yaramayı, tam aksine neyin bilim olup olmadığını bilememize yol açarlar.

(c) Sosyal oluşturmacılık (social constructivism) bilimde sosyal faktör ve değerleri öne çıkaran, nesnelliği ve rasyonaliteyi çöpe atan bir görüştür. Bu bölümdeki analizimde vardığım sonuç oluşturmacıların kabul edebileceğim iddialarda bulunsalar bile resmin bütününe bakmaktan uzak kaldıklarıdır.

## Sonuç

Bir (dini) metnin yorumlanması ile ilgili "exegesis" ve "eisegesis" kelimeleriyle bu tezde yazdıklarım arasında bir bağ kurulabilir. Yunanca kökenli olan bu iki kelime, yorumlamak veya tefsir anlamına gelir. Ama aralarında önemli bir fark var: ilk kelimedeki "ex" öneki dışına anlamına gelirkern "eis" öneki ise içine anlamı taşır. Exegesis metnin anlamını ortaya çıkarmak amacıyla dikkatli bir analize dayalı yorumlamaktır. Eisegesis ise daha önceden sahip olunan fikileri metinde bulmayı amaçlayın bir yorumlamadır. Yorumlayıcı kendi gündemini, savunduklarını, önyargılarını metinde bulmak için uğraşır.

Bu kelimeleri benzetme ile bilim felsefesine taşıyabiliriz. Benim bu tezde savunduğum bilimin felsefesi eisegesis değil exegesis ile yapılması gerektiğidir. Bilimi anlamak istiyorsak yapmamız gereken bütün bilim dallarını aynı küfeye koyup aralarındaki farkları göz ardı etmek olmasa gerek. Daha önceden var olduğunu düşündüğümüz istisnai değerleri ve uygulamaları incelemeden bilimlere atfederek yol katedemeyiz. Felsefi varsayımlarımızla uyuşmayan değer ve pratikleri göz önüne almaya hazır olmalıyız. Gerçek bilim yerine varsaydığımız ideal bir bilimin felsefesini yapmak bizim bilimi anlayışımıza katkıda bulunmaz.

Nasıl bilim sapmasız (unbiased) olmayı hedeflemeliyse bilim felsefecileri de sapmasızlığı temel amaçlarından biri yapmalıdır. Hiç bir zaman bütün sapmalardan kurtulamayacagımızı düşünüyorum. Ancak yine de hedefimiz bu olmalı. Çünkü gitgide daha az sapmalı bir felsefe geliştirmek mümkündür. Minimum düzeyde bir sapmanın günümüzdeki durumdan çok daha iyi olacağı açıktır. Bilimdeki uygulamaları ve değerleri inceleyerek yapılan bir felsefe ezbere yapılanına göre bilim doğası hakkında bize daha çok şey öğretecektir.

# Appendix H

# Tez Fotokopisi İzin Formu

**ENSTİTÜ**

| | |
|---|---|
| Fen Bilimleri Enstitüsü | |
| Sosyal Bilimler Enstitüsü | |
| Uygulamalı Matematik Enstitüsü | |
| Enformatik Enstitüsü | |
| Deniz Bilimleri Enstitüsü | |

**YAZARIN**

Soyadı : Keler
Adı : Raşit Hasan
Bölümü : Felsefe

**TEZİN ADI:** Values, Practices, and Philosophy of Science

**TEZİN TÜRÜ:** Yüksek Lisans ☐ Doktora ☐

| | |
|---|---|
| 1. Tezimin tamamından kaynak gösterilmek şartıyla fotokopi alınabilir. | |
| 2. Tezimin içindekiler sayfası, özet, indeks sayfalarından ve/veya bir bölümünden kaynak gösterilmek şartıyla fotokopi alınabilir. | |
| 3. Tezimden bir bir (1) yıl süreyle fotokopi alınamaz. | |

**TEZİN KÜTÜPHANEYE TESLİM TARİHİ:**