Here be dragons
Exploring the hinterland of science

Maarten Boudry

ISBN 978-90-7083-018-2

Maarten Boudry

Here be dragons

Exploring the hinterland of science

Proefschrift voorgedragen tot het bekomen van de graad van
Doctor in de Wijsbegeerte
Promotor: Prof. dr. Johan Braeckman

UNIVERSITEIT GENT

Supervisor	Prof. dr. Johan Braeckman Wijsbegeerte en moraalwetenschap
Dean	Prof. dr. Freddy Mortier
Rector	Prof. dr. Paul Van Cauwenberghe

Nederlandse vertaling:
Hic sunt dracones. Een filosofische verkenning van pseudowetenschap en randwetenschap

Cover: The image on the front cover is an excerpt of a map by the Flemish cartographer Abraham Ortelius, originally published in *Theatrum Orbis Terrarum* (1570).

ISBN: 978-90-7083-018-2

UNIVERSITEIT GENT

Faculty of Arts & Humanities

Maarten Boudry

Here be Dragons

Exploring the Hinterland of Science

Proefschrift voorgedragen tot het bekomen van de graad van
Doctor in de Wijsbegeerte

2011

Acknowledgements

This dissertation could not have been written without the invaluable help of a number of people (a philosopher cannot help but thinking of them as a set of individually necessary and jointly sufficient conditions). Different parts of this work have greatly benefited from stimulating discussions with many colleagues and friends, among whom Barbara Forrest, John Teehan, Herman Philipse, Helen De Cruz, Taner Edis, Nicholas Humphrey, Geerdt Magiels, Bart Klink, Glenn Branch, Larry Moran, Jerry Coyne, Michael Ruse, Steve Zara, Amber Griffioen, Johan De Smedt, Lien Van Speybroeck, and Evan Fales. Special thanks go to my co-authors and general partners-in-crime Stefaan Blancke, Filip Buekens, Massimo Pigliucci, Bert Leuridan and Johan Braeckman (see below), whose constructive criticism and inspiring ideas have never failed to stimulate me while writing and re-writing different chapters.

My colleagues from our research group The Moral Brain have not only made this research a thoroughly enjoying experience, but they have also reminded me that, as David Hume once wrote, "truth springs from argument amongst friends". In that respect, I am particularly grateful to my supervisor Johan Braeckman, intrepid explorer of scientific hinterlands, on whom I could always count for assistance, advise, and frankly, a good share of laughs as well. Thanks to Griet, Stefaan and Charlie for kindly proofreading parts of this dissertation, and to Gitte for helping me with the lay-out and cover.

I also owe gratitude to my parents, family and friends for their continuing support and encouragement, as well as for their inexplicable patience for abstruse philosophical matters. Speaking of which, my last word of thanks must go to Sarah, without whom this work would never have seen the light of day. She knows what it's like to work on an academic dissertation, and perhaps more tellingly, she know what it's like to live with one who does.

Table of Contents

Introduction

A successful pseudoscience is a great intellectual achievement. Its study is as instructive and worth undertaking as that of a genuine one. – Frank Cioffi (Cioffi 1998, p. 115)

Let us not, in the pride of our superior knowledge, turn with contempt from the follies of our predecessors. The study of the errors into which great minds have fallen in the pursuit of truth can never be uninstructive. – Charles Mackay (1974 [1841], p. 84)

Pseudosciences have stirred relatively little philosophical excitement when compared to the 'real thing'. In certain respects, this is understandable. Why should philosophers waste time with the fake and phony, with that which fails to live up to its pretensions? Isn't there enough real science out there to tickle our philosophical curiosity? Not only have contemporary philosophers been largely indifferent to pseudoscience, but sometimes they have been downright suspicious of the subject. It is widely acknowledged that the very project of distinguishing science from pseudoscience, i.e. the traditional demarcation project, has fallen on hard times. Following the influential critiques of the likes of Larry Laudan (1983), most philosophers today shy away from branding theories as pseudoscience, and enthusiasm for demarcationism has waned significantly over the past decades. Not only has Laudan proclaimed that all philosophical attempts to distinguish science from non-science have failed, but he has also complained that the very terms "pseudoscientific" or "unscientific", with their connotations of dismissal and disapproval, are "hollow phrases which do only emotive work for us" (1983, p. 125). According to Laudan, such parlance should be erased from our philosophical vocabulary altogether, because it is inspired by a naïve conception of the nature of science, assuming a simple dividing line between pseudoscience and the real thing.

To be sure, subsequent philosophical analyses of traditional demarcation solutions (e.g. in terms of falsifiability) have exposed several shortcomings. Laudan is right that the search for a silver bullet of demarcationism, i.e. a small set of necessary and sufficient conditions for neatly separating science from non-science, is a philosophical dead end. However, as we will see, even if the project of a sharp and clear-cut demarcation between science and pseudoscience is futile, and even if the term "pseudoscience" could certainly benefit from increased rigor, this does not mean that the difference is simply non-existent. A pragmatic ballpark definition is not difficult to come up with. If a theory deviates from scientific standards and epistemic virtues by a sufficiently wide margin, while it is touted as scientific by its advocates, we may reasonably brand it as 'pseudoscience' (Boudry 2010).

Even if philosophers have not come up with a single criterion to separate the wheat from the chaff, can't we all agree that some theories clearly belong to the one rather than the other? Indeed, somewhat paradoxically, a second major explanation for the philosophical disinterest in pseudoscience is the widespread sentiment that some ideas are so *obviously* wrong that we should not even bother arguing about them, let alone investigating them. But this casual attitude underestimates the ways in which "obviously wrong" ideas can be wasteful and dangerous in society at large (Pigliucci 2010). The number of people who put faith in bogus treatment and relinquish any form of scientific medicine is alarmingly high. Cults and sects based on pseudoscientific belief systems continue to attract followers. Huge amounts of intellectual resources are wasted in far-fetched conspiracy theories or in shoring up discredited theories like homeopathy and psychoanalysis, not to mention the never-ending quest for evidence of the paranormal. In other words, we ignore pseudoscience at our own peril.

But aren't there professional skeptics of pseudoscience out there for cleaning up the mess as the need arises? Shouldn't we just be thankful to them for providing this service and move on to more interesting and weighty philosophical matters? In a foreword to Michael Shermer's *Why people believe weird things*, the late Stephen Jay Gould noted that

> [S]kepticism or debunking often receives the bad rap reserved for activities — like garbage disposal — that absolutely must be done for a safe and sane life, but seem either unglamorous or unworthy of overt celebration. (Gould 1997, p. ix)

After reading the skeptical investigations of Larry Kusche and Joe Nickell into, respectively, the Bermuda Triangle mysteries and the miracle stories surrounding the shroud of Turin, philosopher John Earman (2000, p. 3) confesses that he finds such readings "often unrewarding" and even "downright tedious":

> [a]fter a few chapters the reader hankers after a silver bullet that will spare us further details by putting an end to all nonsense. (Earman 2000, p. 3)

But of course, as Earman recognizes, it is precisely this silver bullet which the likes of Laudan have disarmed and snatched from the philosophical arsenal. Before the reader is discouraged from reading any further, however, we want to mention a number of reasons for why a philosophical study of pseudoscience is rewarding in its own right, even apart from gaining insight into a pernicious phenomenon in society at large:

(1) More often than not, scientists and philosophers agree on which theories do not deserve to be called scientific, but their opinions diverge over *why* exactly this is so. Where exactly does a pseudoscientific endeavor go awry? How do we identify the false premises and fallacies in a pseudoscience? From a historical perspective, when did a theory or a project cease to be respectable, if indeed it ever was? Which pseudosciences suffer from intrinsic flaws, and which ones simply lack scientific credentials due to a contingent state of affairs (i.e. lack of empirical evidence)?

(2) In light of the previous point, we need to carefully consider what is the appropriate way to deal with pseudosciences in public debate and in society at large. What burden of proof do we impose on pseudoscientists? In other words, what kind of evidence or theoretical progress would convince us of the value of some pseudoscientific doctrine, or at least would incline us to take that doctrine seriously?

(3) What are the common fallacies and flaws of reasoning which pave the road to pseudoscience? By peering into the dark mirror of science, and by retracing the ways in which pseudoscience goes wrong, we gain better insight into the frailties and limitations of human reasoning. Successful pseudosciences are informative because they reflect the forms of irrationality people are particularly vulnerable to. What are the cognitive and psychological underpinnings of pseudoscience? What is the role of fallacies and reasoning heuristics?

(4) What makes for a successful pseudoscience, and how do seemingly weird belief systems propagate in the face of adverse evidence and criticism? What is the epistemic structure of pseudoscientific theory, and what kind of methodology do the supporters of such a theory apply?

(5) The study of pseudoscience and irrationality furthers our understanding of rationality itself. To which extent do the classical canons of rationality differ from or resemble those employed by pseudoscientists?

On medieval land maps, dragons and other mythological creatures were often depicted over unexplored or dangerous regions. On the philosophical map, the territory of *bona fide* sciences has been charted quite thoroughly. Although many philosophers have attempted to trace the boundaries separating science from non-science, by and large the hinterland of science itself remains *terra incognita*. The present works intends to demonstrate that a journey through these little known regions, and a study of the strange beasts dwelling there, can be an exciting philosophical adventure.

Science, supernaturalism and design

The first part of this dissertation, which was partly written in collaboration with my colleague Stefaan Blancke and my supervisor Johan Braeckman, provides a discussion of supernaturalism and design with respect to science. The family of doctrines that go under the banner of creationism, among which its most recent offspring, Intelligent Design theory, have been univocally rejected as pseudoscience by all major scientific organizations. Likewise, in the philosophical community, barring a few theistic dissenters (e.g. Paul Nelson, Alvin Plantinga), there is a broad consensus that Intelligent Design Creationism (IDC) is not science and hence does not deserve a place in the biology classroom. On closer inspection, however, this consensus disguises some important quarrels over the proper rationale for excluding IDC and other controversial theories from the corpus of modern science.

According to a widespread philosophical opinion, which has been adopted in the judicial ruling in the Kitzmiller v. Dover case on the teaching of IDC in biology classrooms, and which has been embraced by a large number of scientists and philosophers, the scientific enterprise is intrinsically limited to natural explanations, and it is not equipped to evaluate supernatural claims. In Chapter 1, we dissect this commitment of science to *methodological naturalism (MN)*, arguing that there is an often-neglected distinction between two forms of MN, each with its respective rationale and view on the proper role of MN in science. In contrast to the philosophical standard view (Intrinsic MN), according to which the commitment to naturalistic causes and explanations is an *a priori* and self-imposed limitation of science, we argue that MN is a provisory and well-tried rule of thumb that simply reflects the past experiences of scientists, viz. the consistent success of natural explanations and the dismal failure of supernatural ones. Evidently, the discussion is intimately linked to the traditional project of demarcating science from pseudoscience: is the exclusive commitment to natural explanations a necessary condition of science? In other words, is IDC pseudoscience or non-science because it violates Intrinsic MN, or do its problems lie elsewhere?

In this chapter, we explore what, if anything, would constitute compelling evidence for supernatural phenomena, and what this tells us about the shortcomings of IDC theory. Along the way, we touch upon issues such as the fallibility of science and its openness to new evidence, indispensable procedural assumptions in scientific practice, and the theoretical possibility of causal interaction with supernatural entities.

The philosophical analysis in Chapter 1 clears the ground for Chapter 2, in which we spell out the implications of this neglected distinction for the public image of science and for the vexed issue of the relationship between science and religion. We argue that ruling the supernatural out of science by philosophical fiat (Intrinsic MN) has been grist

to the mill of IDC, and has fostered misconceptions about science and about the empirical evidence for evolution. We draw upon examples from the history of science and we retrace the political strategies involved in these seemingly arcane philosophical issues. By discussing the literature of IDC advocates and their theological adversaries, theistic evolutionists, and by scrutinizing the theological arguments launched from both camps, we try to find out how the conception of MN bears on the (pseudo)scientific status of IDC theory and the conflict between science and religion.

This second chapter drives home the point that exposing the lack of scientific credentials of a theory is not merely a tedious task reserved for skeptics, but has actual import for the public understanding of science and the acceptance of a foundational theory like evolution. Lest we play in to the hands of creationists, we have to devise careful philosophical arguments for deciding what is and is not science.

In Chapter 3, which is a collaborative effort with biologist and philosopher Massimo Pigliucci, we follow a similar dialectic, this time with regard to the use of machine/information type metaphors in biology and science education. In popular scientific writings and biology textbooks, genes are often described as the "blueprint" for the construction of an organism, and cells are likened to biochemical "factories", complete with assembly lines, transportation systems, information carriers, etc. Tracing back these design metaphors in the mechanical philosophy of the 17th century and the emerging program of natural theology, we argue that such a conception is out-dated and makes little biological sense today. Importantly, the frequent employment of these metaphors fosters widespread misconceptions and design intuitions about biological systems, which IDC advocates have been quick enough to exploit, pushing the design analogy to the point of identity: biological systems look designed because they *are* designed.

As these chapters on design metaphors and methodological naturalism make clear, pseudoscientists may well be irredeemably wrong or misguided, but they can unwittingly force us to think more clearly about established scientific theories, and may inadvertently reveal problems that we had not noticed before. As Pigliucci himself put it in his book on creationism:

> Often evolutionary biologists dismiss creationist arguments out of hand because they are "obviously" wrong, without realizing that even a wrong argument can point the way to a legitimate question underlying the fabric of evolutionary theory. (Pigliucci 2002, p. 236)

In Chapter 4, we have taken up the subject of the biological design argument, around which the program of natural theology was centered from the end of the 17th century onwards, and which still forms the central tenet of the contemporary IDC movement. This chapter, which connects with the first chapters on methodological naturalism, bears on an important question in the philosophy of biology: was it Darwin who gave

the final deathblow to the design argument, by offering a genuine naturalistic explanation for the phenomena that inspired the natural theologians, or had David Hume already shattered the argument by sheer logical force? As a reference point, this chapter focuses on the recent reconstruction of the design argument by philosopher of science Elliott Sober in the framework of Bayesian probability theory. We discuss the intrinsic objections which Sober has leveled against the design argument with regard to its choice of auxiliary assumptions (about the attributes of the designer). Along the way, we explore some issues relating to the problem of accommodation versus prediction in the philosophy of science, the explanatory virtue of unification, and the practice of gerrymandering and *ad hoc* reasoning in pseudoscience. This paper lends further support to the thesis on the *a posteriori* and pragmatic nature of methodological naturalism in science, as developed in the first two chapters.

The structure of pseudoscience

In the second part of this doctoral thesis, we delve into the heart of darkness, dissecting the structure of pseudosciences and weird belief systems in general. In Chapter 5, we provide a theoretical framework of different immunizing strategies and epistemic defense mechanisms that are often encountered in pseudoscientific theories and a variety of other "weird" belief systems. Immunizing strategies are defined as arguments that are brought forward from *outside* a theory to forestall refutation and ward off criticism, whereas epistemic defense mechanisms are conceived as structural and theory-internal features that have the same effect of insulating the theory in question. The distinction is presented as a matter of gradients, and different examples of each category are provided. We discuss the use of face-saving auxiliaries in bona fide science and connect our discussion with Imre Lakatos' notion of progressive and degenerative research programmes. By acknowledging that many a scientific theory displays a "protective belt" of auxiliaries around its core hypothesis, we avoid the pitfalls of naïve falsificationism, while retaining some valuable insights about the virtues of empirical boldness and vulnerability to refutation.

In Chapter 6, we develop this theoretical analysis further and apply the same framework to the discussion about *ad hoc* reasoning in the philosophy of science, which can be construed as part of the traditional demarcation problem facing any sophisticated falsificationist. What, if anything, distinguishes the legitimate recourse to auxiliaries in science from illicit *ad hoc* moves? Instead of rehearsing the classic

examples of successful and more debatable introductions of face-saving auxiliaries in science, our strategy consists of focusing on examples of blatant *ad hoc* reasoning from the domain of pseudoscience. The motivation behind this approach is that, as we have more robust intuitions about the fishiness of such palpably *ad hoc* moves, they allow us to get a better grasp on what it is precisely that we find objectionable about *ad hoc* reasoning, and how we recognize a theoretical move as such. This discussion of adhocness connects with our assessment of the design hypothesis in Chapter 1 and particularly our critique of Elliott Sober in Chapter 4, while it illustrates again our general contention that a careful study of pseudoscience may throw light on a number of unresolved issues in the philosophy of science.

In Chapter 7 and Chapter 8, we develop two case studies exemplifying the use of immunizing strategies and epistemic defense mechanisms in the Intelligent Design literature. In the first case study, we dissect the equivocations in the IDC concept of *irreducible complexity* on different levels, documenting how this conceptual incoherence has allowed IDC advocates to switch back and forth between different versions and make a moving target out of their claims. The point of departure of this analysis is a problem reminiscent of Chapter 2: evolutionary theorists and philosophers critical of IDC all agree that the argument from irreducible complexity is a spurious attempt to demonstrate design in nature, but they seem to pursue two different, seemingly contradictory lines of criticism. In this case, as we explain, IDC advocates themselves are to blame for the confusion. Depending on how we reconstruct the elusive concept of irreducible complexity, a different line of criticism is in order. In the shorter Chapter 8, we take issue with the way Intelligent Design advocates have applied the No Free Lunch theorem, a mathematical result in optimization theory, to the process of evolution by natural selection and the modeling of such evolutionary processes in computer simulation. Once again, we encounter instances of bait-and-switch reasoning regarding the concept of *teleology*, and the notion of "tailoring around" biological fitness functions.

In Chapter 9, which was written in collaboration with philosopher of language Filip Buekens, we provide a second case study of an epistemic defense mechanism, this time focusing on the obscurity and hermeticism of Lacanian psychoanalysis, and particularly the different rationalizations that have been put forward for the incoherence of Jacques Lacan's pronouncements. Not only are these arguments either circular, unsound or simply inconsistent, but interestingly, they draw on the internal conceptual resources of Lacanian theory. Seeing an epistemic defense mechanism in action illustrates how pseudoscientific or pseudo-philosophical belief systems succeed in fending off critical arguments from outsiders, and how the whole project of questioning the value of Lacanian psychoanalysis, by pointing out contradictions and obfuscation in Lacan's pronouncements, may appear to miss the mark entirely from the Lacanian point of view.

Chapter 10 harks back to the original, Freudian version of psychoanalysis, and confronts it with a controversial program in the philosophy of science that goes under the name of Social Constructivism (SC). As it turns out, certain robust forms of SC unwittingly offer an uncanny description of the epistemic predicament in which Freudian psychoanalysts find themselves. In virtue of some of psychoanalysis' well-known methodological deficiencies and conceptual problems, we argue that this most exquisite of pseudosciences exemplifies what science *would* look like if it *were* to function as SC claims it does. A number of theoretical results about *ad hoc* reasoning and immunization from earlier chapters will resurface in this discussion of psychoanalysis. In addition, our analysis allows for the development of a new and independent argument against SC as an account of *bona fide* science, thus making the connection again with the literature on the demarcation project.

In Chapter 11, we draw on psychological research on the frailties of human reasoning to find out where the appeal and cultural success of certain weird belief systems derives from. After providing a summary of the psychological findings on belief perseverance, cognitive dissonance and rationalization, we assemble the theoretical findings from Chapter 5 on epistemic defense mechanisms and challenge the common assumption about the alleged *fragility* of weird belief systems. In the second part of this chapter, we insert both theoretical perspectives in an epidemiological framework on the construction and dissemination of belief systems. By outlining the epistemic rationale of pseudosciences and other weird belief systems, it will transpire from this epidemiological model that their pervasiveness and continuing popularity is a predictable result of the way our mind operates. Although our epistemological analysis is *informed* by research in cognitive psychology, it provides a level of explanation in its own right, revealing patterns and epistemic structures that are not visible on the lower level of cognition. This chapter also sheds new light on the question of the sincerity of believers and on the level of conscious deliberation that is implied in our talking of 'strategies' and 'defenses'.

The final chapter of this dissertation brings these insights on self-validating belief systems to bear on the psychology of petitionary prayer and the *modus operandi* of supernatural beings. In this case study, we take as a point of departure a series of interesting studies on the psychology of divine causation and petitionary prayer conducted by Justin Barrett. This chapter elaborates on the mechanisms of psychological self-correction and cultural selection that are suggested in the account of Chapter 11, and explores how they play out in the formation of beliefs about the efficacy of petitionary prayer. This chapter is intended as a contribution to the burgeoning field of the cognitive science of religion (CSR), which brings research in cognitive and evolutionary psychology to bear on the study of religion. It also connects with the problem of supernatural versus natural causation which we tackled in Chapter 1. In this sense, but hopefully in this sense only, it brings us right back where we started.

Part 1 – Science, Supernaturalism and Design

Chapter 1.
How Not to Attack Intelligent Design Creationism: Philosophical Misconceptions about Methodological Naturalism

Even postulating an unobserved Creator need be no more unscientific than postulating unobserved particles. What matters is the character of the proposals and the ways in which they are articulated and defended. – Philip Kitcher (1982, p. 125)

Abstract.[1] In recent controversies about Intelligent Design Creationism (IDC), the principle of *methodological naturalism (MN)* has played an important role. In this paper, an often neglected distinction is made between two different conceptions of MN, each with its respective rationale and with a different view on the proper role of MN in science. According to one popular conception, MN is a self-imposed or intrinsic limitation of science, which means that science is simply not equipped to deal with claims of the supernatural (Intrinsic MN or IMN). Alternatively, we will defend MN as a provisory and empirically grounded attitude of scientists, which is justified in virtue of the consistent success of naturalistic explanations and the lack of success of supernatural explanations in the history of science (Provisory MN or PMN). Science *does* have a bearing on supernatural hypotheses, and its verdict is uniformly negative. We will discuss five arguments that have been proposed in support of IMN: the argument from the

[1] This chapter has previously been published in *Foundations of Science* (Boudry, Blancke et al. 2010a). Some paragraphs have been added for the purpose of this dissertation.

definition of science, the argument from lawful regularity, the science stopper argument, the argument from procedural necessity, and the testability argument. We conclude that IMN, because of its philosophical flaws, proves to be an ill-advised strategy to counter the claims of IDC. Evolutionary scientists are on firmer ground if they discard supernatural explanations on purely evidential grounds, instead of ruling them out by philosophical fiat.

1.1 Introduction

In the recent debates between evolutionists and proponents of Intelligent Design Creationism (IDC), the principle of **Methodological Naturalism (MN)** has been an important battleground. In response to typical creationist accusations about science's alleged metaphysical bias towards naturalism and materialism, some philosophers and scientists have pointed out that science is naturalistic only on the level of its methodology (MN), but is neutral with respect to metaphysics. The principle MN has itself become the "focal point of intense criticism" (Miller 2009, p. 118) by antievolutionists, and there has also been some discussion among philosophers of science about the correct understanding of MN in relation to its metaphysical counterpart (Edis 2002; Forrest 2000; Koperski 2008; Miller 2009; Nelson 1996; Pennock 1999; Plantinga 2001b; Richter 2002; Ruse 2005; Shanks 2004; Smith 2001).

In fact, there is an important divergence of opinion on the rationale of MN and its proper role in science. We will argue that the most widespread view, which conceives of MN as an intrinsic or self-imposed limitation of science, is philosophically indefensible. On that account, it is also an ill-advised strategy to counter the claims of IDC and other forms of creationism.[2] In Chapter 2, we will see that opponents of evolution have exploited the philosophical flaws in this popular presentation of MN to accuse scientists of philosophical prejudice and dogmatism (Boudry 2009b). Alternatively, we will defend MN as a provisory attitude of science based on the successful track record of natural explanations and the miserable track record of supernatural explanations. Supernatural claims do not fall beyond the reach of science; they have simply failed.

[2] As the principle of MN has in the past been attacked by advocates and sympathizers of the IDC movement, we would like to state at the outset, and to avoid any possible confusion, that we do *not* in any way defend IDC. We completely agree with the majority opinion among philosophers and scientists that IDC has no scientific merits whatsoever, but we differ as to the proper *grounds* for rejecting IDC.

1.2 Naturalism in Science

In the past, creationists have often taken offence at what they saw as the 'dogma' of naturalism and materialism in science. They complained that the hypothesis of special creation is rejected in favor of evolution by natural selection simply because scientists dogmatically cling to metaphysical naturalism, i.e. the claim that nature is all there is. In this worldview, supernatural forces are dismissed out of hand, and there is only place for blind material forces and processes. For instance, already in 1971 Norman Macbeth wrote: "If a Watchmaker is thus carefully excluded at the beginning, we need not be surprised if no Watchmaker appears at the end. The dice have been loaded against him" (Macbeth 1974, p. 126). According to Duane Gish, the universal acceptance of evolutionary theory has nothing to do with scientific evidence but everything with metaphysical prejudice:

> The reason that most scientists accept the theory of evolution is that most scientists are unbelievers, and unbelieving, materialistic men are forced to accept a materialistic, naturalistic explanation for the origin of all living things. (Gish 1973, p. 24)

With the advent of IDC, this philosophical argument rose to prominence, as Phillip Johnson made it the central tenet of his influential *Darwin on trial* (Johnson 1993). In response to this accusation, many evolutionary scientists and philosophers of science have claimed that creationists misconstrue the nature of naturalism in science. They argue that science is committed to Methodological Naturalism (MN), but not to **Ontological/Metaphysical Naturalism (ON)**. For example Robert Pennock:

> Ontological Naturalism should be distinguished from the more common contemporary view, which is known as *methodological naturalism.* The methodological naturalist does not make a commitment directly to a picture of what exists in the world, but rather to a set of methods as a reliable way to find out about the world – typically the methods of the natural sciences, and perhaps extensions that are continuous with them – and indirectly to what those methods discover. (Pennock 1999, p. 191)

The principle of MN demands that scientists appeal exclusively to natural causes and mechanisms. We want to distinguish two conceptions of MN, along with their respective rationale and their different perspective on the proper role of MN in science.

1.3 Two versions of Methodological Naturalism

In the first version, MN is conceived of as an intrinsic and self-imposed limitation of science, as something that is part and parcel of the scientific enterprise by definition. We will term this view **Intrinsic Methodological Naturalism (IMN)**. In the Kitzmiller vs. Dover case on the teaching of IDC in biology lessons, Judge John E. Jones denied the status of science to IDC because it "fails to meet the essential ground rules that limit science to testable, natural explanations" (Jones 2005, p. 70). Based on the testimonies of Kenneth Miller, Robert Pennock and John Haught, Jones stated that "This rigorous attachment to 'natural' explanations is an essential attribute to science by definition and by convention" (Jones 2005, p. 66). Philosopher of science Michael Ruse, among others, agrees that science "by definition deals only with the natural" (Ruse 1982, p. 322; see also Maienschein 2007; Miller 2009; Scott 2004, 1998; Strahler 1992).[3] The position of IMN is also endorsed by the National Academy of Sciences in their official booklet Teaching about Evolution and the Nature of Science:

> Because science is limited to explaining the natural world by means of natural processes, it cannot use supernatural causation in its explanations. Similarly, science is precluded from making statements about supernatural forces because these are outside its provenance. (National Academy of Sciences 1998, p. 124)

Defenders of IMN claim that science has no bearing on questions of metaphysics and is not equipped to deal with claims about the supernatural. Therefore, science is *not* committed to metaphysical naturalism. As John Haught explained: "By its very nature, science is obliged to leave out any appeal to the supernatural, and so its explanations will always sound naturalistic and purely physicalist" (Haught 2004, p. 231; Miller 2009). Similarly, science cannot negate the existence of the supernatural either. In the words of Eugenie Scott, executive director of the National Center for Science Education "Science is a way of knowing that attempts to explain the natural world using natural causes. It is agnostic toward the supernatural – it neither confirms nor rejects it" (Scott 1999, p. 29; see also Scott 1998).

In contrast with this view, which we will criticize in the section below, we defend an alternative view of MN and of its legitimate function in scientific practice. According to what we call **Provisory or Pragmatic Methodological Naturalism (PMN)**, MN is a provisory and empirically grounded commitment to naturalistic causes and

[3] In his recent writings, Ruse takes a more pragmatic approach, arguing that scientists adopt MN because it has proven successful in the past. See for example Ruse (2006, pp. 47-51).

explanations, which *in principle* is revocable by extraordinary empirical evidence. According to this conception, MN did not drop from thin air, but is just the best methodological guideline that emerged from the history of science (see also Shanks 2004; Coyne 2009a; Edis 2006, pp. 16-17; Sarkar 2007), in particular the pattern of consistent success of naturalistic explanations. Appeals to the supernatural have consistently proven to be premature, and science has never made headway by pursuing them. The rationale for PMN thus excludes IMN: if supernatural explanations are rejected because they have *failed* in the past, this entails that, at least in some sense, they *might* have succeeded. The fact that they didn't is of high interest and shows that science *does* have a bearing on the question of the supernatural.

Although the focus in this paper is on different conceptions of MN, we note that Intrinsic MN should not be confused with a form of ON that is traditionally called philosophical naturalism (PN). According to this position, which has exerted a strong influence on the early scientific revolution, the notion of a supernatural explanation is simply incoherent.[4] The proponent of PN maintains that only physical causes can bring about physical effects (the thesis of 'causal closure'), and hence that the notion of a supernatural or non-physical cause is conceptually confused.

As PN rejects appeals to the supernatural *a priori*, it seems more akin to IMN than to PMN, which discards supernatural explanations merely on the basis of their miserable track record. However, PN is a much more stronger position than IMN. Whereas IMN claims that science is ill-equipped to deal with supernatural causes, leaving open the possibility that they exist, PN rules out their existence altogether.

1.4 Five arguments in support of Intrinsic Methodological Naturalism

1.4.1 ARGUMENT 1 – argument from the definition of science

The foremost argument in support of IMN is that MN is simply part of science *by definition*. Michael Ruse, for example, maintains that science "by definition deals only with the natural, the repeatable, that which is governed by law" (Ruse 1982, p. 322). In

[4] Problems with the definition of 'supernatural' will be considered below.

his memorandum opinion of the Kitzmiller v. Dover case, Judge John E. Jones concluded that the "rigorous attachment to 'natural' explanations is an essential attribute to science by definition and by convention" (Jones 2005, p. 66). However, there have been many attempts in the past to define science in terms of a small set of necessary and sufficient conditions, and none of them has achieved general consensus among scientists and philosophers of science (Laudan 1996a; Edis 2006). It is true that the notion of 'supernatural' is completely absent from the corpus of modern scientific knowledge. But does that mean that supernatural events, if any of these would ever occur in our universe, are necessarily beyond the reach of science?

The definition argument for IMN sits uncomfortably with the fact that reputable scientists and skeptics have investigated allegedly paranormal phenomena which, if corroborated through repeatable and careful experiments, would point to the existence of supernatural forces, or at least so they claim. In a famous study by Benson et al., for instance (Benson, Dusek et al. 2006), the therapeutic effect of intercessory prayer in cardiac by-pass patients was investigated through a methodologically sound RCT-trial. Although the study failed to demonstrate any effect, prima facie it is an honest attempt to establish supernatural intervention by scientific means. If intercessory prayer really did help patients to recover from illness, one would expect this to become visible through carefully conducted trials like these.

If defenders of IMN are correct that science cannot deal with the supernatural "by definition", does it mean that these experiments were pointless to begin with, or that scientists are not entitled to be skeptical about the therapeutic effect of intercessory prayer, because such purported phenomena necessarily lie beyond the epistemic reach of science?

The solution depends on one's definition of the term 'supernatural'. At least one philosophical approach vindicates IMN and argues that scientific experiments like the one mentioned are either logically absurd, or what they purport to demonstrate is not 'supernatural' after all. We argue that this approach is philosophically defensible but sterile, and we propose a more fruitful definition.

1.4.1.1 How unnatural is the supernatural?

The philosophical way to vindicate IMN straight away is to define 'supernatural' as any phenomenon that is inaccessible by scientific means *in principle*. Thus, as soon as an allegedly supernatural phenomenon becomes scientifically detectable, it ceases to be supernatural and must thenceforth be reconsidered as 'natural'.[5] 'Scientific' evidence

[5] Another way to uphold this analytic definition is to maintain that science simply ceases to be scientific as soon as supernatural explanations would become successful. Michael Ruse, for example, wrote that "even if

for the existence of supernatural forces is immediately precluded by analytic definition. In her discussion of supernaturalism, Barbara Forrest adopts this analytic approach (although her actual position is more complicated, see below).

> To become more than a logical possibility, supernaturalism must be confirmed with unequivocal empirical evidence, and such confirmation would only demonstrate that this newly verified aspect of reality had all along never been supernatural at all, but rather a natural phenomenon which just awaited an appropriate scientific test. (Forrest 2000, p. 25)

Robert Pennock, for his part, wrote that "if we could apply natural knowledge to understand supernatural powers, then, by definition, they would not be supernatural" (Pennock 1999, p. 290). However, if the scope of the term 'naturalistic' is simply extended to whichever new aspect of reality science might discover, it becomes almost trivial to uphold MN in scientific practice. How could the principle of MN possibly have any adversaries on Pennock's definition of natural and supernatural, or, alternatively, what exactly would it mean to *defend* MN? As one broadens the scope of the natural and narrows down the reference domain of the supernatural, one's commitment to natural explanations loses any practical import.

Consider the claims of IDC. Although in their public pronouncements defenders of IDC tend to be reluctant to identify their Intelligent Designer as the Judeo-Christian God, there is no doubt that their views are religiously motivated and indebted to the tradition of natural theology (see Chapter 4). But if God has left observable traces in our material universe, as IDC proponents claim He did, these are in principle open to scientific investigation, and thus God would be reduced to the realm of the 'natural', by a matter of philosophical definition. Pennock thinks it is ironic that, in the course of introducing God in science, IDC theorists actually *naturalize* God without seeming to realize. That may well be true according to Pennock's definition of 'supernatural', but by the same token IDC theory does not violate the strictures of IMN any longer, and Pennock's argument on the basis of IMN misses the mark. Pennock's conception of supernaturalism is so restricted that there is *nothing left* for him to reject as a naturalist.

Imagine that IDC theorists, contrary to the actual state of affairs, had provided us with clear and unmistakable evidence for intelligent design behind functional biological complexity (in the next section we will consider what could constitute such evidence). Even if the designer were to reside beyond the dimensions of the known material universe, causing major metaphysical upheaval, it would not fit the bill of a

---

Scientific Creationism were totally successful in making its case as science, it would not yield a *scientific* explanation or origins. Rather, at most, it could prove that science shows that there can be *no* scientific explanations of origins" (Ruse 1982, p. 322; see also Haught 2000, p. 201).

"supernatural entity" according to Pennock's logic, and hence our reply to the IDC proponents would have to be along the following lines: 'You see, now we have a *scientific* proof for Intelligent Design. By definition, that means that we are dealing with a *natural* phenomenon. Thus, I was right after all, supernatural causes and forces have no place in science.'

But why would the IDC theorist be bothered that some philosophers regard his hypothesis as 'natural' after all? If he had really succeeded in demonstrating the existence of an intelligent creator residing outside the known material universe, that terminological discussion would probably be the least of his worries. Therefore, when Pennock and Forrest adopt this analytical definition of the 'supernatural', they cannot longer challenge IDC by using IMN as a philosophical shield, because it misses the mark by their own definition. Instead, they will have to convince IDC theorists that the so-called Intelligent Designer would be 'natural' like anything else. This is a different route to the same conclusion we are defending in this article: that claims of IDC have to be confronted head on, and rejected on scientific grounds, instead of being excluded *by fiat* on shaky philosophical grounds.[6]

A being that would be truly 'supernatural' under the analytic definition would have to be completely isolated from our material universe, and not be able (or willing) to interact with it at all. But this hypothetical being is completely irrelevant for our concerns and is surely not the 'intelligent designer' that the IDC theorist (or any sensible theist, for that matter) believes in. The controversial claim of IDC theory is precisely that an Intelligent Designer meddled very directly in terrestrial affairs during the course of evolution: he constructed the flagellum of the bacterium *E. coli*, for one thing (an act with tangible consequences if there ever was one).[7]

---

[6] The discussion between Alvin Plantinga and Michael Ruse on this point is interesting. Plantinga objects to Ruse's appeal to the definition argument that "it is hard to see how anything like a reasonably serious dispute about what is and is not science could be settled just by appealing to a *definition*" (Plantinga 2001b, p. 345). In his reply, Ruse seems to admit that deciding the matter by means of simple analytic definitions would be unsatisfactory, and he goes on to deny that this was his intention. He agrees that this victory "would altogether be too easy to achieve" (Ruse 2005, p. 50). In trying to explain what he *did* mean by his definition argument, however, Ruse gets entangled in his own reasoning:

> What I am trying to do is to offer a lexical definition: that is to say, I am trying to characterise the use of the term "science." And my suggestion is simply that what we mean by the word "science" in general usage is something that does not make reference to God and so forth, but that is marked by methodological naturalism. I am not saying anything at all about whether or not God exists, or has any role in the world or anything like that. I am simply saying that science does not allow for this possibility, judged *qua* science. (Ruse 2005, p. 50)

But it is hard to see how a *lexical* definition of science carries any more epistemic weight than an *analytical* definition.

[7] Although the analytic definition of the supernatural misses the mark of IDC and is therefore uninteresting in this context, we have to note that it accords well with the psychological function of the vulgar notion of the 'supernatural', and the very reason for the widespread fascination with the topic. People are attracted to the mysterious, to that which seems to defy ordinary experience and scientific explanation, and call this

1.4.1.2 A tentative definition of the supernatural

In accordance with our reconstruction of MN as an empirically grounded and provisory methodological guideline of science (PMN), we propose to define 'supernatural' as referring to any phenomenon which has its basis in entities and processes that transcend the spatiotemporal realm of impersonal matter and energy described by modern science (for a similar approach, see Stenger 2008, pp. 14-16). In contrast with the foregoing analytical take on the issue, if any such supernatural force were to *intervene* in our material universe (and of course these are the cases of particular terrestrial interest) we insist on terming it 'supernatural' here.[8] As we will see, this definition is closer to the IDC's conception of supernatural agency, and it is more relevant to the discussion of MN.

Thus, if we possessed compelling empirical evidence for some intelligent entity residing beyond our spatiotemporal universe, but nonetheless capable of interfering with our material world (see below), we would have a demonstration of a 'supernatural' phenomenon. To be sure, as far as we can see there is absolutely no such evidence (but see Swinburne 2004, for an opposing point of view), and therefore scientists are well advised not to waste too much time pursuing supernatural explanations (PMN).[9] However, that does not mean that such evidence would be intrinsically impossible, or that it falls beyond the reach of science.

1.4.2 ARGUMENT 2 - Lawful regularity and the supernatural - Anything goes?

Biologist J. B. S. Haldane once said that, when he designed an experiment, he assumed that "no god, angel, or devil is going to interfere with its course" (Haldane 1934, p. vi). Many defenders of IMN are concerned that, if the idea of supernatural interference were allowed in science, experimental practice would become impossible. Biologist Richard Lewontin sets up a stark contrast between two "irreconcilable world views":

'supernatural' or 'paranormal'. Once a phenomenon has been given a proper and scientific explanation, it no longer satisfies the craving of many for the mysterious, and hence they look for something else.

[8] If one adopts the position of philosophical naturalism (PN), this causal influence will be ruled out. The problem is the one Descartes encountered when he proposed the pineal gland as the place where the human soul interacts with the body. However, we are willing to grant, for the sake of the argument, the *logical possibility* of the idea of supernatural interventions.

[9] Already in 1748 David Hume warned: "The many instances of forged miracles, and prophecies, and supernatural events, which, in all ages, have either been detected by contrary evidence, or which detect themselves by their absurdity, prove sufficiently the strong propensity of mankind to the extraordinary and the marvellous, and ought reasonably to beget a suspicion against all relations of this kind. (Hume 2000 [1748], p. 89)

> Either the world of phenomena is a consequence of the regular operation of repeatable causes and their repeatable effects, operating roughly along the lines of known physical law, or else at every instant all physical regularities may be ruptured and a totally unforeseeable set of events may occur. . . . We cannot live simultaneously in a world of natural causation and of miracles, for if one miracle can occur, there is no limit. (Lewontin 1983, p. xxvi)

Therefore, science simply *has* to adopt the principle of MN, lest the whole enterprise be compromised. As creationist Phillip Johnson put it bluntly: "there is no way to tell God when he has to stop (Johnson 2001, p. 65).

Consider Robert Pennock's elaboration of this argument in terms of natural laws. Pennock points out that "science does not have a special rule just to keep out divine interventions, but rather a general rule that it does not handle any supernatural agents or powers since these are taken by definition to be above natural laws" (Pennock 1999, p. 284).[10] For Pennock, this constitutes the most important reason for disallowing them in scientific practice:

> Lawful regularity is at the very heart of the naturalistic worldview and to say that some power is supernatural is, by definition, to say that it can violate natural laws. [...] Controlled, repeatable experimentation [...] would not be possible without the methodological assumption that supernatural entities do not intervene to negate lawful natural regularities. (Pennock 1999, p. 321)

In her discussion of metaphysical and methodological naturalism Barbara Forrest presents basically the same argument:

> Introducing supernatural explanations into science would destroy its explanatory force since it would be required to incorporate as an operational principle the premise that literally anything which is logically possible can become an actuality, despite any and all scientific laws; the stability of science would consequently be destroyed. (Forrest 2000, p. 10; for other examples, see Lewontin 1997; Scott 1998, p. 20; Strahler 1992, pp. 13-15)

We think there are two problems with this argument. First, Forrest and Pennock do not explain why *any* occurrence of supernatural intervention in the natural realm would necessarily frustrate *all* experimental work or automatically "destroy the stability of science". That anything logically possible can become an actuality is not as dramatic as

[10] He admits that hypotheses involving supernatural causation have uniformly proven unsuccessful in the history of science, but in the end he does not think this is the fundamental rationale for MN (see also Scott 2004). "Clearly, it is not just because such persistence has proven successful in the past that science encouraged this attitude" (Pennock 1999, p. 196).

it seems – in fact, it just restates the definition of logical possibility. Second, even if mysterious supernatural forces acting in the natural world would make scientific endeavors impossible, by itself this does not amount to a good argument for adopting IMN.

As for the first argument, suppose the RCT in *American Heart Journal* turned out to confirm the hypothesis of therapeutic efficacy of intercessory prayer. Moreover, suppose that further experimental work following this demonstration, which would arguably mark a complete revolution in science, established that this form of supernatural causation displays certain predictable regularities. For instance, it works only with prayers officially sanctioned by the Catholic Church, only if the ill person is baptized by a Catholic priest, etc. Though it may be ridiculous to speculate that anything of the sort would ever happen, as no alleged case of miraculous healing has even been authenticated scientifically[11], if it would, there is no obvious reason why the scientific enterprise would collapse at a single stroke.[12] The fact that some prayers actually do help people recover would admittedly cause a complete metaphysical revolution in science (imagine the enthusiasm of theologians), but if the range of action of this supernatural power turned out to be restricted, why would it endanger the rest of our scientific endeavors?

Pennock thinks this is because the idea of supernatural design is a "one-size-fits-all" (Pennock 2007, p. 319) explanation and will therefore leave the door wide open to all sorts of appeals to the supernatural. For example, it would make the practice of law open to "both suits and defenses on a range of possible divine and occult interventions" (Pennock 1999, p. 295). But while this suspicion is certainly justified in the case of the contemporary IDC movement and its political agenda (Forrest, Barbara C. and Gross, Paul R. 2007), we have to ask whether it is an *intrinsic* problem with supernatural explanations. Let us grant, for the sake of the argument, that defenders of IDC intend to appeal to the supernatural only if their own criteria – irreducible complexity, complex specified information – compel them to do so. Behe for example tries to reassure us that "hypotheses for the involvement of an intelligent agent in the development of life or other historical events have to be evaluated on a case-by-case basis" (Behe 2006, p. 242).

This seems reasonable, and skeptical investigators of the paranormal and supernatural would undoubtedly concur. However, as many scientists and philosophers

[11] In order to be beatified by the Roman Catholic Church, a person has to perform a miraculous healing from beyond the grave. At some point in the procedure, the alleged miracle has to be investigated by the Consulta Medica, a board of doctors appointed by the Vatican, to determine whether the recovery was sudden and permanent, and to rule out any scientific explanation for the healing. However, these putative miracles are not accepted by the scientific community at large.

[12] True enough, any such supernatural intervention would force us to revise the fundamental law of energy conservation, or at least allow some exceptions.

have amply demonstrated (Miller 2000; Pennock 1999; Shanks 2004), Behe's notion of irreducible complexity and Dembski's criterion of complex specified information are simply very bad filters for detecting design, because they single out biological phenomena that present no problem whatsoever for standard evolutionary explanations (see also Chapter 7). Thus, it is true that if we would allow the appeal to supernatural causation on so flimsy evidential grounds as the IDC movement wants us to do, in practice that would be an invitation for more spurious appeals to the supernatural. But this is a general logical point: if we allow a particular form of sloppy thinking on one occasion, we are left with no grounds for disallowing the same reasoning in other cases. Thus, the claims of IDC are invalidated by *specific* flaws of reasoning and by the simple lack of evidence, not because of some perceived *intrinsic problem* associated with supernatural explanations.

Our second and more fundamental objection is that reasoning from the perceived 'danger' of allowing supernatural explanations amounts to an *argumentum ad consequentiam*. Pennock and Forrest would probably admit that supernatural interventions in the natural world are at least logically possible, because if they *did* involve logical inconsistencies, that would rule them out ontologically as well, a conclusion Pennock clearly wants to avoid (otherwise he would not insist on the distinction between ontological naturalism and IMN). The point is that we cannot exclude the possibility of massive supernatural interference in the universe *a priori*. For example, if we were stuck in this universe with a whimsical and meddlesome creator, we would simply have no other option than to resign to the impossibility of reliable natural knowledge about the world. After all, when Haldane assumed that no God or devil interfered in his experimental practice, he did this not because of the perceived danger for the stability of science, but because of his own professional experience and the empirical success of this methodological guideline: "this assumption has been justified by such success as I have achieved in my professional career" (Haldane 1934, p. vi). Thus, if the devil *were* to thwart Haldane's experiments, that would just be too bad for Haldane, and too bad for science.

In 1775, the Royal Academy of Sciences in Paris issued a statement that it "will no longer accept or deal with proposals concerning perpetual motion". Numerous attempts had been made to construct a *perpetuum mobile*, and many had claimed success, but none of the devices could withstand critical scrutiny. Even in 1775, well before the full theoretical development of thermodynamics, the decision of the Royal Academy of Sciences was quite reasonable. To review all these complicated devices was very time-consuming, and the consistent failure of all proposals in the past suggested that the physical possibility of perpetuum mobiles was very unlikely. Nowadays, we know that they would violate the first law of thermodynamics (the conservation of energy), which is as rock solid as anything in modern science. However, even such a theoretical consideration does not warrant the conclusion that a device for creating energy *ex nihilo*

is absolutely impossible. It would simply require extraordinary evidence to convince the scientific community, because its existence would completely overthrow our fundamental laws of physics. The fact that serious scientists no longer deal with proposals for a perpetuum mobile does not mean that science *cannot* deal with them, or that scientists *have* to assume that perpetuum mobiles are impossible because these would "destroy the stability of science". This attitude of scientists is simply a good rule of thumb for not wasting too much time on highly improbable claims.

1.4.3 ARGUMENT 3 - Science Stopper

Apart from the previous objection, Robert Pennock offers an argument for IMN that goes back as far as Francis Bacon.[13] According to Pennock, the argument from design, and the appeal to supernatural explanations in general, are nothing but "science stoppers". If we would allow such arguments in science, "the scientist's task would become just too easy" (Pennock 1999, p. 292). With no prospect of a methodology to deal with supernatural phenomena and to proceed with scientific investigation, the supernatural is just a dead end. By adopting the IDC approach and making the design inference, "any motivation for further research would end" (Miller 2009, p. 130).

> [i]f one were to find some phenomenon that appeared inexplicable according to some current theory one might be tempted to attribute it to the direct intervention of God, but a methodological principle that rules out appeal to supernatural powers prods one to look further for a natural explanation (Pennock 1999, p. 293),

Pennock is right, of course, that in practice Behe's argument from irreducible complexity and Dembski's argument from complex specified information boils down to a facile science stopper. However, we think Pennock's formulation is unsatisfactory because it circumvents a more sound argument for dismissing these particular science stoppers, which again boils down to PMN. We claim that the appeal to supernatural explanations of IDC proponents is unscientific, not because it is the easy way out per se (although of course it is), but because it is the easy way out in the face of the very

[13] "For the handling of final causes mixed with the rest in physical inquiries, hath intercepted the severe and diligent inquiry of all real and physical causes, and given men the occasion to stay upon these satisfactory and specious causes, to the great arrest and prejudice of further discovery." Cited in Ruse (2004, p. 16)

reasonable *prospect* – based on our long-standing experience – that a little more effort will yield perfectly reasonable naturalistic explanation (Dennett 2003, p. 395).[14]

Consider Pennock's analogy that "[s]cience is godless in the same way as plumbing is godless" (Pennock 1999, p. 282). As far as we can see, however, plumbers abstain from supernatural explanations for stopped drains not because such an explanation is the easy way out *per se* (although presumably it is) but because the idea of a supernatural Obstructor did not prove to be a particularly fruitful hypothesis for plumbing work. We tend to overlook this simple pragmatic rationale simply because we live in a world where looking for natural explanations for stopped drains is so *dead obvious*.

Returning to the world of living systems, what reasons do we have, apart from our successful track record and the global materialistic picture emerging from science, to believe that we will eventually come up with naturalistic explanations for all biological phenomena? The more general point lurking here connects with the previous argument about the stability of science as well: what *a priori* guarantee do we have that nature is structured for the convenience of human scientific inquiry? As Alvin Plantinga writes: "Obviously we have no guarantee that God has done everything ... in such a way as to encourage further scientific inquiry, or for our convenience as scientists, or for the benefit of the National Science Foundation" (Plantinga 2001b, p. 357). Although we certainly do not share Plantinga's theistic premises, his point is equally valid for the non-theist. How can we be so sure that the universe is structured in a way that is favorable or rewarding to scientific inquiry, and that is comprehensible to the human mind? As we saw in the previous section, if God or the devil chose to thwart our every attempt at scientific inquiry, we would have no other option than to resign ourselves to that. In a world like that, science is simply useless as way of obtaining knowledge.

Our brains evolved primarily to deal with the adaptive problems our ancestors faced in Pleistocene environments. If one considers the opportunistic process of tinkering and modifying by which natural selection typically arrives at adaptive solutions, it would not have been terribly surprising if our brains were "cognitively closed" for the fundamental structures of the universe that surrounds us. For example, the phenomena described by quantum mechanics and general relativity theory are so difficult to grasp because our minds were simply not designed to cope with phenomena on a cosmological or microscopic scale. Only with a lot of mental effort and mathematical equipment have

[14] Darwin seems to have advocated this sober rationale for the pursuit of naturalistic explanations as well. In *The Expressions of the Emotions in Man and Animals*, he wrote: "No doubt as long as man and all other animals are viewed as independent creations, an effectual stop is put to our natural desire to investigate as far as possible the causes of Expression. By this doctrine, anything and everything can be equally well explained; and *it has proved as pernicious* with respect to Expression as to every other branch of natural history." [our italics] (Darwin 1965 (1872), p. 12) (quoted in Shanks 2004, p. 62)

scientists been able to overcome these cognitive limitations.[15] As Einstein once remarked: "The most incomprehensible thing about the world is that it is comprehensible".

Moreover, the *mere* fact that an explanation is easier in terms of intellectual effort does not mean that it is less likely. For example, Freudian dream interpretations, with their intricate associations and multiple levels of unconscious meaning, are certainly more demanding on an intellectual level than 'resorting' to the idea that dreams are just incoherent fragments of nocturnal brain activity. However, the Freudian theory of dreams is wrong beyond any reasonable doubt. Modern psychology discourages the search for hidden meaning in dreams, just like it discourages the search for a *perpetuum mobile*. In these cases, stopping further inquiry has been reasonable. Thus, Plantinga has a point when he writes: "The claim that God has directly created life [...] may be a science stopper; it does not follow that God *did not* directly create life" (Plantinga 2001b, p. 357). The mere fact that it is a "science stopper" cannot be a knockdown argument. However, when we consider that stopping the search for natural explanations in the face of biological complexity has *always been premature*, and that creationists have never grown tired of searching for new candidates of insurmountable 'gaps' when the old ones had worn out, we *do* have very good reasons to dismiss the current spate of challenges to evolutionary theory.

Finally, it is not even obvious that invoking a supernatural force in an explanation would automatically necessitate that one ceases to investigate the phenomenon altogether (Koperski 2008). One may be confident that, on the contrary, such supernatural phenomena would attract lots of attention from the scientific community at large.

1.4.4 ARGUMENT 4 – Procedural Necessity

In her discussion of MN, Barbara Forrest uses another argument in defense of IMN, although her position turns out be somewhat ambiguous. On the one hand, she does refer to the historical failure of supernatural explanations, and she defends a "tentative" denial of the supernatural: "the more science successfully explains, the less justification there is for the supernatural as an explanatory principle" (Forrest 2000, p. 11). This certainly suggests that she is on the PMN side of the debate. On the other hand, however, Forrest relies on the sterile definition of the supernatural we quoted above,

[15] Colin McGinn has claimed that the human mind is cognitively closed for some fundamental problems (see McGinn 1994).

which simply equates it with "that which is beyond the reach of science". Thus, she claims that a proof for the supernatural is *procedurally impossible*, because it is impossible "to prove the existence of something about which nothing can be known *through scientific investigation*" (Forrest 2000, p. 10). IMN thus follows naturally:

> Methodological Naturalism does exclude the supernatural as an explanatory principle because it is *unknowable by means of scientific inquiry*. (Forrest 2000, p. 14)[16]

But if the possibility of scientific evidence for the supernatural is already excluded from the very start, why would the unsurprising failure of supernatural explanations be of any interest to Forrest? Why would the metaphysical naturalist appeal to "the lack of explanatory success of supernaturalism" (Forrest 2000, p. 14), as Forrest does, if that success were *procedurally impossible*? Or, to approach the dilemma the other way around: if supernatural explanations are deemed to have *failed* in the past, in a non-trivial sense, it follows that they *might* have succeeded. It seems that Forrest wants to have her cake and eat it too. This leads her to a paradox, which she acknowledges but never resolves: "Paradoxically, supernatural claims are the kind of propositions for which empirical evidence is required, but impossible to obtain" (Forrest 2000, p. 16).

In Steven Schaferman's defense of naturalism, we find a similarly ambivalent position. On the one hand, he writes that naturalism is a "hypothesis that has been tested and repeatedly corroborated, and so has become reliable knowledge itself" (Schafersman 1997). Schafermans takes this to be evidence of metaphysical naturalism as well, which is not exactly the topic of this paper, but his reasoning is certainly in line with our outline of PMN (and we would concur with his tentative metaphysical conclusions). On the other hand, however, Schaferman thinks that evidence for the supernatural is a procedural impossibility in science:

> It is doubtful whether any empirical evidence can possibly exist that would prove, demonstrate, or even suggest the existence of the supernatural. Such evidence posited by philosophical supernaturalists would certainly be labeled incomplete, incoherent, illogical, meaningless, misunderstood, or misinterpreted by philosophical naturalists, and thus rejected as reliable evidence. (Schafersman 1997)

Thus, he writes that naturalism is a "methodological necessity" in the practice of science and even an "ontological necessity for understanding and justifying science by scientists."

We think this paradox can be resolved if Forrest and Schafersman plainly choose the P-horn of the MN-dilemma and abandon the uninteresting and sterile definition of the

[16] For another example of this argument, see Miller (2009, p. 127)

term 'supernatural', which, as we discussed, is not relevant for the claims of IDC anyway. The whole argument about the "procedural necessity" of MN can then be dropped.

From our perspective, the consistent failure of supernatural explanations is not only the rationale for PMN, but also provides the main argument for (provisionally) endorsing metaphysical naturalism. Forrest, however, maintains that IMN, grounded by the principle of procedural necessity, "provide[s] an epistemologically stable foundation for a metaphysics" (Forrest 2000, p. 14). We think this is peculiar, because the whole point of IMN, according to its defenders, is to *separate* methodology from metaphysics (Pennock 1999, p. 191). More specifically, as we will see in Chapter 2, defenders of IMN want to assure religious believers that science is *only* naturalistic as far as its methodology is concerned, and that it has no bearing on questions of metaphysics. Forrest's claim that IMN provides a foundation for metaphysical naturalism, conjoined with the contention that IMN is an intrinsic part of science, lends ammunition to the false creationist charge that scientists are *dogmatic* metaphysical naturalists (Boudry, Blancke et al.).[17]

1.4.5 ARGUMENT 5 - Methodological Naturalism and Testability

A last and important concern about supernatural explanations, which for some philosophers constitutes the main reason to dismiss them *a priori*, is that they are *intrinsically untestable*.

> [Science] rejects the possibility of supernatural explanations not as a matter of principle, but of *methodology*: What kind of research would one do, what kind of methodology would one use, if the premise were that God can do whatever He pleases whenever He wishes to do it? (Pigliucci 2002, p. 29)
>
> If there is an omnipotent force in the universe, it would by definition be impossible to hold constant (to control) its effects. [...] The reason that the ultimate statement of creationism cannot be tested is simple: the actions of an omnipotent creator are compatible with any and all observations of the natural world. (Scott 2004, pp. 19-20)

[17] Apart from that, we completely agree with Forrest's argumentation for metaphysical naturalism as "founded upon the methods and findings, respectively, of modern science" and "the only reasonable metaphysical conclusion—if by reasonable one means both empirically grounded and logically coherent" (Forrest 2000, pp. 8,9).

> A supernatural agent is unconstrained by natural laws or the properties and capabilities of natural entities and forces – it can act in any way and accomplish any conceivable end. (Miller 2009, p. 128)

We submit that Scott and Pigliucci are attacking *specific* kinds of supernatural claims to derive unwarranted conclusions about supernatural explanations in general. In fact, their unfalsifiability objection applies only to an omnipotent God who either deliberately conceals his own existence, or who has chosen to create exactly the world one would expect if there were only blind and material forces at work. Take for example the Omphalos hypothesis of Philip Gosse (1857), according to which God deliberately planted forged evidence for an old earth to test our faith, including dinosaur fossils and far-off stars. As many critics noted, the idea is as gratuitous as the idea that God created the world 5 minutes ago, and there is no conceivable way to refute it. However, we should note that the same holds for any number of supernatural *and* natural conspiracy hypotheses about the world, for example the claim that there are elves in my backyard who disappear whenever someone is watching them or trying to capture them on videotape. If a hypothesis is designed to be impervious to falsification, scientists are justified in dismissing it, and in that sense we completely agree with Scott and Pigliucci.

Critics have also pointed out that proponents of supernatural claims, notably IDC theorists, often make use of evasive maneuvers that render their theories immune to empirical falsification (see Chapter 5). For example, in response to the argument from imperfection and bad design, Michael Behe has simply replied that we cannot gather any scientific information about the character and intentions of the Designer, that His reasons are unfathomable and that any speculation about them is pure metaphysics. Pennock has rightly dismissed this immunizing strategy:

> [I]t provides the design argument with a virtually impenetrable shield [...] Behe has successfully insulated the design argument against the imperfection argument. Equipped with such bumpers it can now withstand any impact. (Pennock 1999, p. 249)

As we will see in Chapter 4, Darwin observed that many supernatural 'explanations' of the biological and physical world are merely "re-stating the fact in dignified language" (Darwin 1998 (1859), p. 151). They do not generate novel predictions, because they were "explicitly designed to yield the already known facts [...] *and nothing more*" (Worrall 2004, p. 68).

However, we have to be careful not to misconstrue the immunizing strategies and *ad hoc* amendments of creationists as *intrinsic problems* with supernatural claims. It is true that IDC proponents are guilty of immunization strategies, but as far as we can see, this unwillingness to take empirical risks is just an indication of the dismal state of their research programme. As we will see in Chapter 5, resorting to immunization strategies

is a typical feature of pseudo-science, supernatural or otherwise (Boudry and Braeckman 2010).

Thus, if only they chose to do so, IDC proponents could fashion a supernatural designer with specific attributes and intentions in such a way that the design hypothesis would yield unexpected predictions and is *not* "compatible with any and all observations of the natural world", as Scott claims (Scott 2004, p. 20; Richter 2002, p. 21). For example, if one supposes that the Designer is benevolent and has created the universe with good purpose, as almost any theist does, one is confronted with the problem of evil and suffering in the world (Hume 2007 [1779]; Kitcher 2007, p. 130). As Reed Richter pointed out, in response to Scott's defense of IMN, "'[s]upernatural' does not automatically imply arbitrary, capricious action as Scott implies".

Figure 1 Scientists witnessing extraordinary evidence for supernaturalism. Or do they? (based on a cartoon by Sidney Harris)

Moreover, even if we are talking about an omniscient and omnipotent God, the mere fact that He *can* make his own existence completely undetectable does not mean that He chooses to do so. An omnipotent God could as well prefer to reveal himself to the world and leave unmistakable traces of His creative activity, as the natural theologians from the 19th century thought He did. As an interesting analogy, consider the story *Non Serviam* by Stanislaw Lem, in which computer scientists have managed to create conscious life forms in an artificial mathematical environment whose properties are fundamentally different from the three space dimensions and single time dimension of their creator's world (Lem 1999, pp. 167-196). While the programmer is observing and monitoring the actions and behavior of these 'personoids', he discovers that the creatures start to speculate about their own origins and the possible existence of a creator. Throughout the story, their creator remains completely imperceptible for the personoids themselves, but this is only because the programmer has decided not to reveal his existence to them, and to strictly adhere to the principle of non-intervention (which he does for ethical reasons). If only he chose to do so, however, the programmer could easily reveal his own existence to his creatures. From the point of view of the personoids, the action of the programmer would be completely "miraculous", i.e. a breach in the fabric of their world, completely unexplainable in terms of their "physics", but with detectable consequences nonetheless.

There is no rational reason at all to believe that we are in the predicament of the personoids in *Non Serviam*, but the story makes clear that supernatural intervention could be quite straightforward and scientifically detectable (Richter 2002), even if we would have no full 'access' to this supernatural realm, on account of our limitations as three-dimensional beings composed of matter and energy. In line with Lem's story, Taner Edis has developed a hypothetical metaphysical picture based on the analogy of a computer program, which nicely illustrates that a supernatural designer residing beyond the material universe could still be able to "acquire information about the world and act on it." (Edis 2002, p. 41). Victor Stenger even made up a list of hypothetical observations that would favor the God hypothesis (Stenger 2008, pp. 231-234). We mentioned the alleged therapeutic effect of intercessory prayer as one possible source of scientific evidence for the supernatural. In general, if one wants to demonstrate the supernatural, one has to "seek ways in which the natural order is disrupted, indicating a reality beyond the material world" (Edis 2002, p. 43). Repeated miracles or psychic wonders, established through carefully controlled experiments, would "put an unbearable strain on a naturalistic view of the world" (Edis 2002, p. 188). And if we could detect specific patterns in the incidence of these miracles, they might provide us with clues about the nature of the supernatural cause behind them, or even the intentions of the supernatural being responsible for them.

We conclude that the argument from 'intrinsic unfalsifiability' misconstrues typical immunization strategies and *ad hoc* maneuvers of creationists as *general and intrinsic problems* with supernatural explanations.

1.5 Discussion & Conclusion

In this paper, we reviewed five arguments in favor of the conception of MN as an intrinsic property of science (IMN), and we found them all wanting: the argument from the definition of science, the argument from lawful regularity, the science stopper argument, the argument from procedural necessity, and the testability argument. Instead, we defended MN as a provisory and empirically grounded commitment of scientists to naturalistic causes and explanations, which is in principle revocable by overwhelming and unmistakable empirical evidence (PMN). Evolutionary scientists are on firmer ground if they discard supernatural explanations on purely evidential grounds, and not by philosophical fiat.

The numerous attempts to establish the existence of supernatural and paranormal phenomena *might* have succeeded (Humphrey 1996, pp. 23-28; Edis 1998). As we argued in the last section, unless the alleged supernatural Creator is involved in a cosmic conspiracy that makes his existence completely undetectable to us, it would not be terribly difficult to look out for scientific evidence for his presence.

In a letter to biologist Asa Gray, Darwin wondered what would convince *him* of design (see also Coyne 2009a):[18]

> If I saw an angel come down to teach us good, and I was convinced from others seeing him that I was not mad, I should believe in design. If I could be convinced thoroughly that life and mind was in an unknown way a function of other imponderable force, I should be convinced. If man was made of brass or iron and no way connected with any other organism which had ever lived, I should perhaps be convinced. But this is childish writing. (Darwin 2000, pp. 169-170)

Although the idea of one unmistakable 'smoking gun' for design is probably misguided, one can easily imagine a collection of observations that would lend plausibility to the

[18] A similar argument is brought forward by Cleanthes in Hume's *Dialogues concerning natural religion* (Hume 2007 [1779])

hypothesis of a supernatural origin of life (see also Coyne 2009a; Richter 2002). Suppose that species just popped into the fossil record without any discernible traces of evolutionary descent and without demonstrable relationship to other species. Suppose that all available dating methods concurred on a 6000 year old earth and universe, and that all attempts to explain living systems by any combination of natural mechanisms consistently failed. To top it all, suppose that the letters of the book of Genesis were discovered to be encoded in human DNA (for a discussion of deliberate signatures from Designers, see Dennett 1996, pp. 316-318). This may seem like a preposterous thing to imagine, but it does not involve any logical contradictions, and it is difficult to deny that it would constitute compelling evidence for the hypothesis of supernatural interference in the universe.[19]

One final argument for resisting this conclusion is based on Arthur C. Clarke's third law, which states that "any sufficiently advanced technology is indistinguishable from magic". The idea is that we are never in a position to exclude advanced natural technology (for example, from extraterrestrial beings) in accounting for an apparent miracle, and hence that it is impossible to ever establish the occurrence of a supernatural event. First, it remains to be demonstrated that there is no logically possible event the occurrence of which would entitle us to conclude that a fundamental law of nature had been violated. Second, the argument on the basis of Clarke's third law is in fact a weak logical point, not unlike Duhem's underdetermination thesis. Although it is framed in the context of miracles, it is not restricted to the question of supernatural causation. Take any scientific explanation for a natural event. No matter how plausible, parsimonious and well-established our best explanation, we can never logically rule out a far-fetched conspiracy theory that is observationally equivalent to the explanation we accept. For example, any sufficiently powerful Cartesian demon may have orchestrated the events reported in a local news bulletin. Likewise, even if we can never logically rule out advanced extraterrestrial technology, that does not mean that the balance of probabilities will always tilt in favor of the space aliens.

The argument reminds one of the village skeptic in bad horror movies, a cliché character who typically remains unmoved even after having witnessed the most outlandish supernatural events (the point presumably being that skeptics are such stubborn and closed-minded people). But we don't think that is how a rational person would react in the face of massive breaches in the laws of the universe, Clarke's third law notwithstanding.

---

[19] Philosopher of science Elliott Sober has recently argued that the design argument is intrinsically defective (Sober 2008, pp. 109-188). In the end we think his a priori objections are unavailing, because his own reconstruction of the argument of design in terms of "likelihood" does not capture the argument satisfactorily. See Chapter 4.

Figure 2 Arthur C. Clarke's Third Law, "Any sufficiently advanced technology is indistinguishable from magic", as interpreted by Calvin & Hobbes.

We conclude that naturalists do not need to resort to the claim that scientific evidence for anything supernatural is logically or conceptually *impossible*, and in any case it is imprudent to do so. As Sahotra Sarkar noted, philosophical modesty teaches us that all scientific knowledge is fallible and that we cannot rule out any metaphysical picture with absolute certainty. Metaphysical transformations, although not quite so radical, have happened before, for example with Newton's theory of gravity (the notion of 'action at a distance') and particularly with the advent of quantum mechanics (the demise of classical determinism). Compelling empirical evidence, however unlikely, would in the end trump methodological scruples (Sarkar 2007; see also Koperski 2008).

In the next Chapter, we argue that the principle of IMN is also an ill-advised attempt to reconcile science and religion. By excluding the supernatural from science by philosophical fiat, IMN has been grist to the mill of anti-evolutionists intent on accusing scientists of philosophical prejudice and dogmatism. To this end, they have exploited some of the specific philosophical weaknesses discussed in this paper. In our view, the conception of PMN salvages these philosophical problems and provides a more accurate picture of the proper role and rationale of science's naturalistic methodology.

Chapter 2.
Grist to the Mill of Intelligent Design Creationism: The Failed Strategy of Ruling the Supernatural out of Science *a priori*

Extinguished theologians lie about the cradle of every science as the strangled snakes beside that of Hercules. – Thomas Huxley

Abstract.[20] According to a widespread philosophical opinion, the methodology of science is intrinsically naturalistic. It is simply not equipped to deal with supernatural claims, so it has no authority on questions of metaphysics. This (self-imposed) limitation on the epistemic reach of science is often used as a way to divorce evolutionary science from metaphysical naturalism or atheism. We argue that ruling the supernatural out of science for intrinsic reasons is not only philosophically untenable, but has actually been grist to the mill of anti-evolutionism. The philosophical weakness of this conception of naturalism in science has been eagerly exploited by proponents of Intelligent Design Creationism to bolster their claims about scientists' alleged naturalistic bias and dogmatism. We argue that it fosters a misleading image of evolutionary theory and that it sits uncomfortable with the foremost arguments for evolution by natural selection.

[20] This chapter has been provisionally accepted in *History & Philosophy of the Life Sciences*. Parts of it have been presented at the conference *150 Years After Origin: Biological, Historical and Philosophical Perspectives* at Victoria College, University of Toronto, November 2009 (Boudry 2009b).

Thus, we conclude that the attempt to reconcile evolutionary science and religion on the basis of this philosophical doctrine is doomed to fail.

2.1 Introduction

For over a long time, creationists and intelligent design proponents have complained that modern science, and evolutionary theory in particular, is biased towards materialism and naturalism, and that it rules out any supernatural forces *by fiat* (Gish 1973; Macbeth 1974; Johnson 1993; Nelson 1996; Behe 2006). In response to these charges, a substantial number of philosophers and scientists have recently argued that science is only committed to something they call *methodological naturalism*. As we saw in Chapter 1, however, there has been some philosophical discussion about the correct understanding of methodological naturalism and its proper role in science (Nelson 1996; Pennock 1996; Forrest 2000; Behe 2006). In section 1.3, we have made an often neglected distinction between two conceptions of methodological naturalism, which involve two quite different views on the limits of science and the proper role of a naturalistic methodology (Boudry, Blancke et al. 2010a).

To briefly rehearse this distinction, a widespread philosophical opinion conceives of methodological naturalism as an intrinsic and self-imposed limitation of science, as something that is part and parcel of the scientific enterprise *by definition.* According to this view (Intrinsic MN or IMN), science is simply not equipped to deal with the supernatural and therefore has no authority on the issue (Pennock 1999; Scott 1998; Haught 2004; Ruse 2005; Jones 2005; Fales 2009). We reviewed five arguments for this conception of IMN but found none of them convincing. Instead, we defended methodological naturalism as a provisory and empirically grounded commitment of scientists to naturalistic causes and explanations, which is *in principle* revocable in the light of extraordinary empirical evidence (Provisory or Pragmatic MN – PMN). In this view, MN derives its justification from the impressive dividends of naturalistic explanations and the consistent failure of supernatural explanations in the history of science (Edis 2002; Shanks 2004; Coyne 2009a; Fishman 2009).

In this paper we analyze the wider implications of this important distinction between IMN and PMN for the status of evolutionary theory and its relationship with religion. We claim that ruling the supernatural out of science by definition or for intrinsic reasons proves a counterproductive strategy against Intelligent Design Creationism (IDC), and, for that matter, against any theory involving supernatural claims. As we will

show, IMN is actually grist to the anti-evolutionist mill on several accounts, and the attempts to reconcile religion and science on its basis are doomed to fail.

2.2 The theological friends of IMN

According to PMN, a reasonable scientist today will no longer pursue supernatural explanations, because these have proven to be a waste of time and effort. But that is not to say that they *cannot possibly* be true. All scientific knowledge is fallible, and in principle supernatural explanations might be vindicated one day, although the prospects are rather dim, to say the least. Defenders of IMN, however, want to exclude supernatural explanations from science for intrinsic reasons. They think that the commitment of scientists to natural explanations is unalterable in the light of future scientific developments, and the idea of a 'supernatural explanation' in science is nothing but an oxymoron.

On the face of it, therefore, PMN seems to take the claims of supernaturalism more seriously than does IMN. This does not mean, however, that IMN is typically advocated by adversaries of religion and supernaturalism, quite to the contrary. Precisely because IMN shuts the door for the supernatural completely, it is often used, in the words of one of its proponents, as a way to "divorce [evolutionary science] from supposedly atheistic implication" (Ruse 2005, p. 45).

The term 'methodological naturalism' itself was coined in 1983 by evangelical Christian and philosopher Paul deVries, who used it to make room for "other sources of truth" besides science.

> If we are free to let the natural sciences be limited to their perspectives under the guidance of methodological naturalism, then other sources of truth will become more defensible. However, to insist that God-talk be included in the natural sciences is to submit unwisely to the modern myth of scientism: the myth that all truth is scientific. (deVries 1986, p. 396)

Not surprisingly, IMN is typically embraced by philosophers sympathetic to religion, by theistic evolutionists and religious liberals intent on safeguarding a special epistemic domain for religious faith (Haught 2000), but also by 'accommodationist' atheists who simply wish to temper the heated opposition between religion and science (Ruse 2001, 2005). Although both theists and atheists have brought forward philosophical arguments for IMN that command our attention, it is also apparent from their writings

that their position is partly inspired by the desire to "protect the religious sensibilities of theistic evolutionists" (Schafersman 1997). In a way similar to Stephen Jay Gould's principle of Non-Overlapping Magisteria (NOMA) (Gould 1999), in the eyes of many IMN embodies the modern *modus vivendi* between science and religion. However, not every theist is content with this polite stand-off between science and religion...

2.3 The theological enemies of IMN

With the principle of IMN coming to the fore of the debate, proponents of IDC have certainly not backed down on their claim that evolutionary biologists are guilty of philosophical and naturalistic prejudice. Phillip Johnson even made it the central tenet of his influential *Darwin on trial* (Johnson 1993):

> For all the controversies over these issues, however, there is a basic philosophical point on which the evolutionary biologists all agree. ... The theory in question is a theory of *naturalistic* evolution, which means that it absolutely rules out any miraculous or supernatural intervention at any point. Everything is conclusively presumed to have happened through purely material mechanisms that are in principle accessible to scientific investigation, whether they have yet been discovered or not. (Johnson 2001, p. 61)

IDC proponents have consistently treated IMN as a token of the same old metaphysical prejudice, this time in a thin methodological disguise: for on what grounds, other than metaphysical bias, could one dismiss one class of explanation in favor of another? Robert Pennock among others has claimed that Johnson simply fails to appreciate the difference between methodological and metaphysical naturalism (Pennock 1999, p. 192; Pennock 1996), but we will argue that the situation is more complicated. In fact, Johnson's remarks illustrate that IMN, clearly being the target of his attack here, is actually grist to the IDC mill. The attempt to reconcile science and religion on this philosophical basis is a strategic failure.

2.4 Four reasons not to adopt IMN

2.4.1 A counterproductive strategy

Advocates of IMN attempt to give the naturalistic methodology of science a solid philosophical underpinning. In doing so, however, they have divorced the methodology of science from the successful track record of naturalistic explanations. This makes it look as if science has never bothered to consider supernatural explanations, and had discarded them already at the outset. IDC proponents, always eager to cast evolutionists in the role of dogmatists with naturalistic blinders (Pennock 1996), while picturing themselves as nothing but neutral and open-minded observers, have repeatedly exploited this philosophical weakness of IMN.

Philip Johnson, as one would expect from a 'good' lawyer, has seized upon this weak spot in the defense of his opponent and turned it to his advantage: if science is supposed to be neutral with respect to metaphysics, as defenders of IMN claim, why is the hypothesis of supernatural design already "disqualified at the outset" (Johnson 2001, p. 67; see also Plantinga 1996; Dembski 1999, pp. 97-121)? Elsewhere, Johnson has complained that "[b]y the use of labels, objections to naturalistic evolution can be dismissed without a fair hearing" (Johnson 1993, p. 7; see also Dembski 1999). Critical and open-minded scientists, according to Michael Behe, have to follow the evidence *wherever it leads*, instead of ruling out some options for philosophical reasons (Behe 2006, p. 243; see also Nelson 1996; Nelson 1998; Dembski 2004, pp. 168-172; Bledsoe 2006, pp. 255-256). Anti-evolutionists have repeated their complaints about scientists' naturalistic bias over and over, almost invariably choosing IMN as their target.

Consider how Alvin Plantinga has spelled out the implications of IMN from a historical perspective:

> Well, suppose we adopt this attitude [IMN]. Then perhaps it looks as if by far the most probable of all the properly scientific hypotheses is that of evolution by common ancestry: it is hard to think of any other real possibility. [...] So it could be that the best hypothesis was evolution by common descent – i.e. of all the hypotheses that conform to methodological naturalism, it is the best. But of course what we really want to know is not which hypothesis is the best from some artificially adopted standpoint of naturalism, but what the best hypothesis is *overall*. [...] (Plantinga 2001c, pp. 137-138)

Plantinga's argument illustrates how IMN backfires on science, because it suggests that the philosophical dice have always been loaded against the supernatural:

> The believer in God, unlike her naturalistic counterpart, is free to look at the evidence for the Grand Evolutionary Scheme, and follow it wherever it leads, rejecting that scheme if the evidence is insufficient. (Plantinga 2001c, p. 138; see also Dembski 2004, pp. 170-171)

In the eyes of IDC advocates, it is precisely this rigid methodological exclusion that makes scientists, and evolutionary biologists in particular, myopic to what they perceive as the self-evident fact of supernatural intervention. If it were not for IMN, so the argument goes, intelligent design would long have been vindicated.

Although this may seem like an arcane philosophical matter, IDC proponents never fail to point out to the public that it must be a very bad theory that needs to be shored up by shaky philosophical arguments. For creationist Paul Nelson, IMN is a desperate move to keep theology out of science (Nelson 1998; Bledsoe 2006). For sociologist and ID-sympathizer Steve Fuller, it is "as if contemporary science was so indefensible on its own merits that it required a philosophical fig leaf for protective cover" (Fuller 2007, p. 117).

Many defenders of IMN insist that IDC advocates simply fail to grasp the difference between methodological and metaphysical naturalism (Scott 1998; Pennock 1999; Miller 2009). But the confusion is partly theirs. As we argued in Chapter 1, a complete disregard for potential supernatural causes makes sense only if we possess of airtight *a priori* reasons that the supernatural does not exist (a view to which defenders of IMN don't *want* to be committed), or that if it does, it never interferes with our material universe. This point has not escaped the attention of the more sophisticated creationists (Dilley 2010; Nagel 2008). In the absence of a sound rationale for disqualifying the supernatural, the dictum of IMN to proceed "as if" only natural causes are operative looks quite arbitrary.

Imagine what would happen if supernatural forces were *really* operative in our universe. In such a world, IMN would be a very bad methodological device indeed, because it would exclude a real and tangible factor governing the universe from scientific consideration. As Richard Dawkins wrote (see also Edis 2002, 1998):

> A universe with a supernatural presence would be a fundamentally and qualitatively different kind of universe from one without. The difference is, inescapably, a scientific difference. (Dawkins 1997, p. 399)

This is the reason why, despite the disclaimers of Scott and Pennock, IDC theorists who read about the principle of IMN persist – albeit falsely – that scientists must be dogmatic metaphysical naturalists (Johnson 1995; Dembski 1999; Dilley 2010). In the conception of PMN we defend, science provides support for, but does not *collapse* into metaphysical naturalism.

The writings of ID creationists show that this image of 'unfair exclusion' is actually to their strategic advantage. Consider Richard Lewontin's often-quoted statement about materialism in science:

> It is not that the methods and institutions of science somehow compel us to accept a material explanation of the world, but, on the contrary, that we are forced by our *a priori* adherence to material causes to create an apparatus of investigation and a set of concepts that produce material explanations [...]. Moreover, that materialism is absolute, for we cannot allow a Divine Foot in the door. (Lewontin 1997, p. 28)

Phillip Johnson lauds this paragraph as "the most insightful statement of what is at issue in the creation/evolution controversy that I have ever read from a senior figure in the scientific establishment" (Johnson 1997, p. 23). Michael Behe quotes molecular biologist Richard Dickerson's discussion of IMN in full, in which Dickerson argues that science is "a game with one overriding and defining rule", namely that of IMN. In the discussion that follows, Behe has IMN where he wants to have it: "the clear implications is that [the supernatural] should not be invoked *whether it is true or not*" (Behe 2006, p. 239). Behe opposes Dickerson's conclusion: "Science is not a game, and scientists should follow the physical evidence wherever it leads, with no artificial restrictions" (Behe 2006, p. 243).

Defining the supernatural out of science is thus counterproductive and weakens the scientific case against design (Kurtz 1999, p. 28; Stenger 2008). It lends ammunition to an accusation that was already voiced by young-earth creationists in the seventies (Gish 1973, p. 24; Macbeth 1974, p. 126), namely that evolution by natural selection appears to win the scientific debate only because supernatural designers were already carefully excluded from the outset. Borrowing from IDC's own metaphoric imagery, IMN may very well be the philosophical crack into which IDC theorists are now trying to drive their wedge.

Before moving on, we should note that some advocates of IMN have made an analogous claim to the opposite effect, namely that rejecting the supernatural on *empirical* grounds is counterproductive. For example, theologian and theistic evolutionist John Haught maintains that those who place the God hypothesis within the reach of science deprive themselves of the best philosophical argument against IDC:

> [I]n wedging ultimate explanation into what should be a purely scientific channel of inquiry [evolutionary materialists] are doing exactly what their ID adversaries do: they are conflating science with a worldview. [...] By promoting their own peculiar alliance of science and philosophical belief they leave themselves with no methodological high ground to stand when they complain about ID's mixing of biology with theology. (Haught 2000, p. 207)

First, Haught's argument begs the question: of course, if one presupposes that science is not equipped to deal with supernatural claims, it follows naturally that those who think that evolution disfavors the God hypothesis are conflating science and metaphysics. But this is precisely the issue at hand. If IMN is philosophically untenable, as we claim, and if scientific knowledge *does* bear on the God hypothesis, materialism is no longer just a "philosophical belief", but the most rational conclusion emerging from the scientific data. Second, even while he presents the views of his opponents, Haught curiously holds on to the claim that the only way to attack IDC is to claim "methodological high ground", whereas this is precisely what "scientific materialists" (in our terms, defenders of PMN) are trying to avoid. As Haught should know, evolutionary materialists concur with IDC proponents in their view that a supernatural designer, if any such entity exists, is *in principle* within the reach of science. *They* are not the ones to complain about any conflation of science and theology on the part of IDC, and indeed, it would be inconsistent for them to do so.

2.4.2 Making sense of the history of science

Where did the idea originate that science cannot deal with supernatural causes? Attempts to explain mysterious phenomena that were previously attributed to deities in terms of natural causes date back to early Greek philosophy, and came to typify the activity of Christian natural philosophers from the Middle Ages onward (Numbers 2003, p. 266). Early pioneers of the scientific revolution like Galileo Galilei were the first to apply a naturalistic methodology to the study of the visible world with great success. This development culminated in Newton's vision of a clockwork universe, in which natural phenomena were described solely in terms of simple interactions and movements of matter. However, scientific pioneers such as Newton were certainly not metaphysical naturalists (Schafersman 1997), as they left the door open for a supernatural presence in the natural world. Famously, Newton resorted to supernatural intervention to explain the stability of the solar system.

Although natural philosophers and scientists following Newton professed their preference for naturalistic explanations above divine miracles and interventions, many were convinced that the living world and especially the human mind would forever defy a naturalistic understanding. Over time, however, the track record of naturalism became ever more impressive, and even phenomena that were previously deemed to require supernatural causes yielded to a naturalistic description. For example, this happened with Newton's problem of the stability of the solar systems, the origin of species and the problem of biological adaptations, the human moral sense, the phenomenon of mystic experiences and so forth (Tyson 2005). Especially in the wake of Darwin's theory, which delivered a promissory note for a complete naturalistic picture

of the living world (Bowler 2007), many scientists removed the supernatural from their explanatory resources altogether.

Ironically, this progressive naturalization has in retrospect fostered the ill-founded opinion that science is simply not equipped to evaluate supernatural claims in any case, and that the very notion of a supernatural explanation is an oxymoron. In their polite reluctance to make theists face the embarrassment, defenders of IMN now pretend that there really was no dispute to begin with, because science simply *cannot* deal with supernatural causes *in principle*. IMN suggests that natural explanations inevitably *had* to come out at the end of the day, and that things could not have been otherwise. But even after the advent of Darwin's theory, there was still substantive discussion – justifiable or not – about supernatural guidance in nature. Many scientists did accept the fact of the common ancestry of species, but they rejected Darwin's main mechanism, natural selection (Bowler 2007). Although the idea of special creation of individual species moved decidedly out of fashion, scientists still toyed with the idea that some sort of teleological driving force, or a form of supernatural guidance, was indispensable to account for evolutionary progress and biological complexity (Bowler 2007, p. 108). Eventually, with the development of the Modern Synthesis in the 1930s and the revolution of molecular biology, even these teleological and supernatural explanations were gradually abandoned.

From the perspective of IMN, chasing God from science is a matter of enhanced philosophical understanding of the nature of science. For example, Michael Ruse maintains that theories and hypotheses in the history of biology have been scientific only insofar as they began to adhere to the strict prescript of MN.

> [E]volutionism grew up from being a pseudoscience, through being a popular science, to being what I term a mature or "professional" science. At various stages along this process, one sees a transformation as evolution does become more subject to the strict dictates of methodological naturalism. (Ruse 2005, p. 48)

But Ruse's accounts sets the cart before the horse. It is not very different from saying that, at the turn of the 19th century, physicists became more and more subject to the "strict dictates of atomism", as if atomic theory was itself not a result of contingent scientific discoveries. As Taner Edis wrote; "[t]here is no Scientific Method residing in a realm separate from the results of science; our methods are our results" (Edis 2002, p. 258). To suggest that the naturalization of evolutionary biology is simply the result of some timeless philosophical insight having dawned on biologists is to obscure the evidential reasons for these developments. Therefore, calling the efforts of early biologists "unscientific" with hindsight is inappropriately anachronistic. As Taner Edis wrote:

> Nineteenth-century biologists did not come to think special creation was a hypothesis they were not allowed to entertain. They rejected it, deciding evolution explained life better. And intelligent design is still, on the face of it, a straightforward fact claim. (Edis 2002, p. 58)

The pioneers of the life sciences could very well have bumped into phenomena that defied their every attempt at naturalistic explanation (but they didn't). In the world *we* happen to live in, science is capable of offering a comprehensible natural explanation for many phenomena that were previously deemed 'mysterious'. But apparently this perspective can easily distort our view on what is logically and metaphysically possible. We are so accustomed to the absence of any credible evidence for the supernatural that we are tempted to conclude that such evidence has to be *impossible*. But as Koperski noted regarding science's naturalistic methodology, it is wrong to think that "once a concept achieves the status of shaping principle it becomes an immutable axiom for all future science" (Koperski 2008, pp. 437-438).

2.4.3 The retreat of the theologians

Evolutionary science has been increasingly successful in finding impersonal and blind material explanations for phenomena that were previously held to be inexplicable in anything other than supernatural terms. Not surprisingly, evolutionary theory has attracted the attention of many a worrisome theologian. Theological doctrines have been revised in the light of new scientific developments, and metaphysical theories have over time been converted into vague metaphors (Bowler 2007). Some theologians, among whom defenders of IMN, have been quick to argue that the whole project of finding God in nature was misguided in any case, and they have tried to erect philosophical walls that safeguard a place for God (Edis 2002, pp. 51-58; see also Dennett 1996; Stenger 2008). As Taner Edis wrote, however:

> Those scientists and theologians who looked to nature to find God had good reasons to do so. It was always possible that natural science would confirm the glory of God. It just did not happen. There is no averting disaster by retroactively calling the whole enterprise a religious mistake. (Edis 2002, p. 108)

IDCers have rightly sensed that this enormous success of naturalism makes the idea of a supernatural creator alarmingly implausible. By contrast, advocates of IMN, atheists and theists alike, have tried to soft-pedal these implications. For example, Robert Pennock, in his otherwise very informative book on IDC, reassures his readers that "[s]cience is godless in the same way as plumbing is godless" (Pennock 1999, p. 282). But Pennock's analogy is highly misleading. If the explanatory domain of modern science

did merely consist of stopped drains and water pipes (or something equally modest), of course one could still comfortably resort to God when it comes to weightier matters of explanation. But as Pennock surely knows, modern science has extended its explanatory reach far beyond, including many domain that were traditionally reserved for God: the origin of life, the beginning of the universe, the human mind, the edges of the observable world etc. That God turned out to be superfluous in all *these* domains is of far greater importance than that He is of no use to plumbers.[21] Ironically, Pennock tries to present this conspicuous absence of God from modern science, which is the logical result of His superfluity on every level of scientific explanation, as an indication of science's metaphysical neutrality.

> Nowhere in evolutionary theory does it say that God does not exist, for the simple reason that, like cell theory and relativity theory and every other scientific theory, it says nothing at all about God. But to say nothing about God is not to say that God is nothing. (Pennock 1999, p. 333)

But nowhere in biology does it say that Bigfoot does not exist either. Surely Pennock does not want to conclude that biologists are completely neutral on the question of Bigfoot's existence?

As another example, consider Christian philosopher Howard Van Till's claim that science is "religiously inconclusive": "[m]odern scientific theories concerning the properties, behaviour and formative history of the physical universe are logically independent of both theism and naturalism, favoring neither one nor the other" (see also Van Till 2001, p. 153; Haught 2003, p. 776). But, as Alvin Plantinga has correctly pointed out, the claim that science cannot logically exclude theism is "a statement weak *in excelsis*" (Plantinga 2001a, p. 202). Logical possibility is a very weak criterion for belief, because there is precious little that science *can* logically exclude (even Bigfoot). The argument from logical consistency is a red herring that diverts attention away from the fact that evolutionary science has dramatically undermined a whole class of positive arguments for the God hypothesis.[22]

---

[21] Massimo Pigliucci, even though his position is close to IMN, has correctly pointed out that it is perfectly rational to spell out the "philosophical" implications of modern evolutionary theory:

> Although asserting that evolution is purposeless is indeed a philosophical conclusion, it follows from everything we know and is consistent with the assumption in every science (not just evolutionary biology) that we can explain nature without recourse to the supernatural (gravitation seems as undirected and purposeless as evolution). (Pigliucci 2002, p. 258)

[22] Defenders of IMN often point out that many prominent scientists were religious believers and scientific naturalists nonetheless. Gregor Mendel, for example, was a metaphysical supernaturalist, but he developed his knowledge of the laws of heredity using natural causes only. First, simply because Mendel used only naturalistic explanations in his particular domain does not commit him to the idea that science in general is only allowed to use naturalistic explanations. Maybe he found naturalistic ones simply the most successful or fruitful explanations in the science of heredity? Second, and more importantly, the case of a scientist

In this context, defenders of IMN often set up a straw man for the alternative view, namely the argument that science has *conclusively proven* that atheism is correct. Evolutionary biologist and Roman Catholic Kenneth Miller sets up this false dilemma:

> [T]he conflict depends [...] on an unspoken assumption. That assumption is, if the origins of living organisms can be explained in purely materialistic terms, then the existence of God – at least any God worthy of the name – is disproved. (Miller 2000, p. 190; see also Scott 1998; Sober 2010)

But who has defended that "unspoken assumption"? Even Daniel Dennett, whom Miller explicitly lists among the strident atheists allegedly sharing this assumption, is careful enough to argue that "[u]ndermining the best argument anybody ever thought of for the existence of God is not, of course, proving the nonexistence of God" (Dennett 2007, p. 139). Of course it is impossible to definitely *prove* that God does not exist, but it does not follow that scientific findings have no bearing whatsoever on the plausibility of the God hypothesis. We cannot disprove Russell's teapot either, but that doesn't make its existence a bit more plausible.

Defenders of IMN think that the mere *logical consistency* of science with the God hypothesis closes the case, but they ignore other important ways in which evolutionary science can bear on the God hypothesis. Although this strategy may be well-intended as a means to protect religious sensibilities, and may look like a convenient solution in the context of the separation of Church and State in the US, it does not hold up to philosophical scrutiny and is arguably a little sanctimonious.

2.4.4 Good fences make good neighbors?

Defenders of IMN hold that the epistemic authority of science is limited to the natural realm. In the words of geologist Keith B. Miller:

> Science is a methodology that provides a limited, but very fruitful, way of knowing about the natural world. This method works only if science confines itself to investigation of *natural* entities and forces. (Miller 2009, p. 117)

This self-imposed restriction to the "natural domain", which is implicit in the view of IMN, invites one to infer that, for sure, there exist other domains beyond the natural realm, but that they just happen to fall outside the scope of science (McMullin 2001, p.

---

professing to be a religious believer only proves that believing in God and defending science are psychologically compatible for some people, not that it makes much philosophical sense (Coyne 2009a).

168). At best, theistic defenders of IMN view science and religion on a par, as two valuable and non-overlapping sources of knowledge. However, from the claim that science is strictly involved with the natural domain, it is a small step to the conclusion that science is "limited" and only religion can offer us deep knowledge. For example, Reformed Christian Howard Van Till is a strong defender of science and IMN, but he does not buy any of its naturalistic conclusions:

> As I see it, granting the limited competence of natural science is not a concession to naturalism; rather, it is simply a recognition that we have empirical access only to creaturely phenomena. [...] science [provides] an incomplete picture of reality because of its inability to probe beyond the creaturely realm. (Van Till 2001, p. 161)

For his part, theologian John Haught has embraced IMN in almost lyrical terms, as it resonates with his conviction that theology offers us 'deeper' knowledge than science can attain:

> Theology is now freed from moonlighting in the explanatory domain that science now occupies, so that it may now gravitate toward its more natural setting - at levels of depth to which science cannot reach. (Haught 2004, p. 236)

But as we have shown in Chapter 1 (Boudry, Blancke et al. 2010a), if supernatural forces were to intervene in the natural world, as IDC proponents and other theists maintain, that would inevitably involve empirically detectable changes, and these are in principle open to scientific investigation.

If religion really were to constitute an equally valid source of knowledge, as defenders of IMN suggest, it would be *all the more* unreasonable to refuse taking into consideration its "deeper knowledge" for such questions as the origin of life. As Alvin Plantinga wrote, "surely the rational thing is to use *all* that you know in trying to understand a given phenomenon" (Plantinga 2001b, p. 341). But this is exactly what defenders of IMN try to avoid.

2.4.5 The empirical case against supernatural design

In *On the Origin of Species* Charles Darwin took the reigning paradigm of natural theology very seriously. Rather than dismissing special creation out of hand, he repeatedly contrasted the hypothesis with his own account of evolution by natural selection. In particular, Darwin devotes considerable time to anomalous phenomena which he argues are "inexplicable on the ordinary view of the independent creation of each species", and which support his "one long argument" for evolution through natural selection (Darwin

1998 (1859)): homologies, imperfect and rudimentary organs geographical distributions, embryology etc. For instance, Darwin writes:

> On the ordinary view of each species having been independently created, why should that part of the structure, which differs from the same start in other independently-created species of the same genus, be more variable than those parts which are closely alike in the several species? I do not see any explanation can be given. But on the view of species being only strongly marked and fixed varieties, we might surely expect to find them still often continuing to vary in those parts of their structure which have varied within a moderately recent period, and which have thus come to differ. (Darwin 1964, p. 155)

And on the graduated variety in nature, he wrote:

> [N]ature is prodigal in variety, but niggard in innovation. Why, on the theory of Creation, should this be so? Why should all the parts and organs of many independent beings, each supposed to have been separately created for its proper place in nature, be so invariably linked together by graduated steps? Why should not Nature have taken a leap from structure to structure? On the theory of natural selection, we can clearly understand why she should not; for natural selection, we can clearly understand why she should not; for natural selection can act only by taking advantage of slight successive variations; she can never take a leap, but must advance by the shortest and slowest steps. (Darwin 1964, p. 194)

Even now, many scientists consider the imperfections and oddities of nature more compelling arguments for evolution than the examples of 'perfect' adaptation, because the latter just mimic the actions of an alleged intelligent creator (Gould 1980). The point is that such empirical objections make sense only if one takes the theory of special creation seriously as an alternative scientific explanation. IMN inadvertently sabotages this empirical case against design, rendering Darwin's arguments against design superfluous. Worse still, defenders of IMN commit themselves to the peculiar view that a considerable part of *On the origin of species* is 'unscientific', because it discusses and evaluates supernatural explanations (even if in a purely negative way). After all, if supernatural explanations should be ruled out from the outset for intrinsic reasons, it makes little sense to argue that the empirical evidence speaks 'against' them. If supernaturalism fails on empirical grounds in the life sciences, this entails that, at least in some sense, it *might* have succeeded, something which is only allowed by PMN. The problem reminds one of an old Jewish joke: someone borrowed a copper kettle from B and after he had returned it, he was sued by B because the kettle now had a big hole it. His defense was: "First, I never borrowed a kettle from B at all; secondly, the kettle

already had a hole in it when he gave it to me him and thirdly, I gave the kettle back undamaged".[23] By setting up an artificial dividing line between science and supernatural claims, IMN has shot itself in the foot.

In Chapter 1, we argued that, not only in the life sciences, but also in other domains of inquiry, skeptics and parapsychologists have investigated extraordinary claims which, if corroborated, would substantiate the existence of immaterial and supernatural entities (e.g. ghosts, extra-sensory perception, the healing power of prayer; see Humphrey 1996). IDCers like Paul Nelson have been quick to point out that the empirical arguments against supernatural design sit uncomfortably with the widely advocated principle of IMN in science (Nelson 1996, 1998; Dembski 1999; Woetzel 2005; Dilley 2010), and on that particular point they are quite right. As Thomas Nagel put it in his own *rapprochement* with IDC theory:

> The conceivability of the design alternative is part of the background for understanding evolutionary theory. To make the assumption of its falsehood a condition of scientific rationality seems almost incoherent. (Nagel 2008, p. 201)

Indeed, the problem has been apparent in the ambivalence of defenders of IMN themselves concerning the epistemic reach of science. For example, Robert Pennock acknowledges the historical failure of supernatural explanations in the life sciences, but he thinks this cannot be the "main reason" for rejecting design explanations, and he then gives several *intrinsic* reasons for ruling the supernatural out of science (Pennock 1999). Or consider Barbara Forrest's discussion of scientific naturalism. On the one hand, she offers a definition of the 'supernatural' that renders supernatural explanations in science procedurally impossible or even constituting a *contradictio in terminis*. On the other hand, she proposes a "tentative rejection" of the supernatural "in light of the heretofore consistent lack of confirmation of it" (Forrest 2000, p. 23), a claim with which we can heartily agree. As we saw in section 1.3, however, the two conceptions of methodological naturalism are mutually exclusive and should not be conflated: either one defends PMN, implying that supernatural explanations *might* have succeeded, or one chooses IMN, which is to rule them out of science *a priori*.

[23] The example is given in Freud's *The Interpretation of Dreams*.

2.5 Functional Integrity and God as a Creature

Regarding the question of how to attack IDC, we have sided with those who confront the claims of IDC head on, and against the accommodationist policy of IMN. For the most part, this is an internal dispute among scientists and philosophers who are in any case firm advocates of evolutionary theory. Some defenders of IMN are metaphysical naturalists and atheists all the same, but they simply feel that this is a purely philosophical discussion which should be separated from scientific issues (Pigliucci 2010).

However, it is interesting to have a look at a parallel debate which has been raging between those who concur on a completely different metaphysical outlook, that of theism, but who quarrel over the proper way to support their case and the correct way to think about the relation between religion and science. Both camps disagree over this fundamental question: is there any sound theological argument to accept that supernatural claims fall beyond the epistemic purview of science, and hence to endorse the strictures of IMN?

From our perspective, only a world in which God does not intervene directly is a world in which IMN makes sense. If God really performed miracles, he would be on the scientific radar. But liberal theologians think that a worthy deity cannot but have a non-interventionist policy. Several theological justification for that assumption have been put forward, among which that of Diogenes Allen:

> God can never properly be used in scientific accounts, which are formulated in terms of the relations between the members of the universe, because that would reduce God to the status of a creature. According to a Christian conception of God as creator of a universe that is rational through and through, there are no missing relations between the members of nature. If in our study of nature, we run into what seems to be an instance of a connection missing between members of nature, the Christian doctrine of creation implies that we should keep looking for one. (Allen 1989, p. 45)[24]

One the one hand, miraculous interventions would reduce God to the sorry state of a creature. On the other hand, the Christian God would not allow for gaps in the natural economy of his creation. The latter assumption was coined the thesis of "functional integrity" by Howard Van Till:

[24] Cited in Plantinga (2001b, p. 347).

> In such a Creation there would be no need for God to perform acts of 'special creation' in time because it has no gaps in its developmental economy that would necessitate bridging by extraordinary divine interventions of the sort often postulated by Special Creationism. (Van Till 1996, p. 21)

This theological position safeguards religion from direct confrontations with science, provided that scientists themselves are prepared to honor the same philosophical boundary with non-theistic rationalizations of their own. On the other side of the theological debate, however, we find IDC proponents such as Alvin Plantinga and William Dembski, who dismiss these arguments as half-hearted rationalizations of an untenable retreat of theology.

The ill-conceived arguments of IDC advocates against evolution being what they are, it is difficult not to sympathize with them on this point. As Johnson succinctly put it in an exchange with Van Till about IMN:

> Effectively, that means that God must be exiled to that shadowy realm before the Big Bang, and He must promise to do nothing thereafter that might cause trouble between theists and the scientific naturalists. (Van Till and Johnson 1993)

According to Dembski, the thesis of the self-sufficiency of nature "artificially constricts the range of things God may ordain" (Dembski 1999, p. 64; see also Plantinga 2001b). Liberal theologians, however, fear that the idea of direct interventions by God in the course of nature is equivalent to the God of the gaps theology. This theological view postulates divine action to account for those phenomena which science has left unexplained. Many theologians find the idea unacceptable, because it would restrict God's action only to particular aspects of reality, which science might be able to explain later on. Plantinga, however, among other defenders of IDC, has pointed out that believing in divine intervention does not commit one to the 'God of the gaps' theology at all, a position which he himself forcefully dismisses as an "anemic and watered-down semideism" (Plantinga 2001b, p. 350). One can still perfectly maintain, as Plantinga does, that God constantly and directly sustains the whole of his creation, natural laws included.

Why, then, is the God of the gaps theology used as a red herring in the context of IMN, and why do theologians like McMullin, Van Till and Allen accept this retreat of God into the "shadowy realm before the Big Bang"? We think it suggests a different explanation: theological defenders of IMN seem to be aware, *unlike* Plantinga and IDC proponents, that appealing to supernatural explanations in the face of unresolved scientific problems has *always been premature*, and that such problems have consistently yielded to a naturalistic framework. From their perspective, pursuing arguments like 'irreducible complexity' in biology is a guaranteed dead-end for theology (see Chapter 7). Because they do not share Plantinga's misconceptions about evolutionary theory, they *do* realize that the scientific evidence for evolution by natural selection is

overwhelming. If science has failed to unearth any evidence for a supernatural Creator of the universe, what better solution than to pretend that it simply has no bearing on the supernatural at all? In this light, it is not surprising that some theologians have tried to reach a mutual understanding with those evolutionary scientists who, for their part, have no intention to tread on the domain of metaphysics. Faced with a pending defeat, liberal theologians have simply opted for a draw.

On the other side of the debate, Plantinga, Johnson and Dembski are keenly aware that, if the theory of evolution by natural selection is true, this constitutes an enormous threat to religious belief. In fact, it reduces the theistic God to little more than a logical possibility, which is a very weak basis for belief. In the words of Dembski:

> Atheists, materialists and naturalists had been offering promissory notes that natural laws were sufficient to explain life. It was Darwin's theory, however, that put paid to these promissory notes. [...] By giving a plausible picture of how mechanization could take command and make life submit to mechanistic explanation, [Darwin] cleared the ground for the triumphant march of mechanistic explanations in biology. (Dembski 1999, pp. 83-84)

The only solution from their perspective is simply to resist the claim that naturalistic evolution tells the whole story. As they are confident that the existence of God can be scientifically demonstrated, they will have none of the rationalizations offered by Allen, Van Till and others (Van Till and Johnson 1993; Dembski 1999; Plantinga 2001a). In a certain sense, they can't be blamed.

2.6 Confusion about methodological naturalism

As is clear from our discussion, we think the real issue in the debate about methodological naturalism with respect to IDC is not the confusion between metaphysical and methodological naturalism, as Eugenie Scott and others like to think, but the distinction between IMN and PMN. We think that the current divergence of opinion on the subject of methodological naturalism deserves more attention. IDC theorists often present the Intrinsic version of methodological naturalism as the consensus opinion among scientists, because they obviously perceive it as an easy target (see for example Johnson 2001, p. 61; Dembski 1999, pp. 117-119, 2004, pp. 170-171). This is not altogether unsurprising, as some scientific and philosophical advocates of IMN

themselves write as if IMN were universally accepted among scientists and philosophers (Edis 1998; Sarkar 2009).

Unfortunately, even some writers whose position is in line with what we call PMN fail to notice the popular appeal of the alternative view. For example, in his excellent critique of IDC, Niall Shanks has no patience with the suggestion that science is by definition restricted to natural causes and explanations, which he labels as a "smoke-and-mirrors strategy" (Shanks 2004, p. 139) of IDC advocates. But this is to underestimate the confusion on the issue among defenders of evolutionary theory. Shanks is right to dismiss Dembski's complaint that "methodological naturalism is the functional equivalent of a full-blown metaphysical naturalism" (Dembski 1999, p. 119), because in Shank's presentation it amounts to no such thing. However, Shanks seems to be unaware that the popular IMN version that Dembski attacks (but wrongly assumes to be the scientific consensus view), indeed *only makes sense* if we have prior reason to accept metaphysical naturalism (see for example Dembski 2004, p. 191; Nagel 2008, pp. 193-194).

Thus, Shanks writes that "the methodological naturalist will not simply rule hypotheses about supernatural causes out of court" (Shanks 2004, p. 141), whereas this is exactly what authors like Eugenie Scott, John Haught and Robert Pennock do. In a review of Shanks' book, IDC sympathizer Del Ratzsch unsurprisingly accuses Shanks of misrepresenting even the views of his evolutionist allies, and he confronts him with a catalogue of favorite quotes from IMN advocates (Ratzsch 2005, pp. 39-48).

In a reply to Paul Nelson's critique of methodological naturalism (Nelson 1996), philosopher of biology Kelly C. Smith rightly points out that science "is not in the business of ruling things impossible" (Smith 2001, p. 713), and he emphasizes that whenever supernatural explanations were invoked in the history of science, they never survived critical scrutiny for very long. However, Smith's article leaves the reader with the impression that there is no discussion among philosophers of science about this approach, whereas many of his colleagues hold views incompatible with his own. For example, would Eugenie Scott agree that in principle science is always open to the possibility of supernatural explanations?

2.7 Discussion

At some point in David Hume's *Dialogues concerning natural religion* (Hume 2007 [1779]), Philo and Demea take sides together against Cleanthes' *a posteriori* arguments for the

existence of God. They maintain that the human mind is simply too limited to be able to grasp the unfathomable nature of God, that His existence is self-evident and can be known *a priori*. Arguably, Philo (or Hume) is also just being careful to avoid that his skeptical arguments against the design argument collapse into outright atheism. Demea for his part does not want to make God's existence dependent on something as mundane as an *a posteriori* argument. Fallible and imperfect as these are, they would make the case for religious faith too vulnerable to atheist attacks.

In a similar manner, the philosophical principle of IMN has tried to consolidate a truce between (evolutionary) science and religion. In the context of the ongoing efforts by IDC advocates to sneak their pseudoscience in the classroom, this may look like a convenient and well-intended solution to uphold the separation between church and state (Scott 2004; Jones 2005). In Chapter 1, however, we concluded that the attempt to rule IDC out of science with intrinsic arguments does not hold up to philosophical scrutiny. As Young and Edis succinctly wrote, "scientists reject claims of intelligent design because of their failure, not because intelligent design is indelibly stamped with a philosophical scarlet letter" (Young and Edis 2006, p. xii).

In this paper, we developed our argument further and demonstrated that IMN also fails from a strategic perspective: not only does it fail to counter the classical creationist accusation of naturalistic bias in science and evolutionary biology in particular, but it actually makes matters worse. The philosophical weakness of IMN has been eagerly exploited by IDC theorists to repeat their old accusations, to cast themselves in the role of open-minded truth-lovers, and to point out the discrepancy between the letter of IMN and actual scientific arguments against supernatural design. With IDC proponents becoming ever more prominent and philosophically sophisticated, IMN will be subjected to more and more philosophical pressure. For that reason, excluding supernatural explanations from science with muddled philosophical arguments is certainly a bad idea. Moreover, IMN compromises the epistemic status of science and suggests that it is just *one* way of knowing – even a "limited" one - among others. Finally, it soft-pedals the metaphysical reverberations of evolutionary theory, and fails to appreciate the unremitting process of naturalization in the history of science.

These problems show that maybe IMN does not fail so much *as* a strategy, but because it was set up as a political strategy in the first place.[25] The main motivation

[25] Although we maintain that IMN is philosophically and historically unsound, we have to acknowledge the possibility that it has proven fruitful in bringing about the success of the scientific enterprise. Since IMN provides a convenient way to reconcile science and religion, it helps in uniting many people from different backgrounds and with different worldviews in the collaborative enterprise that science is. Arguably, this would have been more difficult if the naturalistic program of science and evolutionary theory in particular had been perceived as an immediate threat to religion. Up to the present day, scientific organizations like the NCSE and the NAS have won the continuing support of progressive theologians and religious scientists for

behind IMN seems to be a desire to reassure the faithful and retain the support of theistic evolutionists and religious liberals in the battle against anti-evolutionist forces. Understandable as this may be from a political perspective, the purported reconcilement between science and religion on the basis of IMN happens at the expense of philosophical and scientific integrity, and is therefore misguided.

---

evolution by trying to "divorce [evolutionary science] from supposedly atheistic implication" (Ruse 2005, p. 45)

Chapter 3.
Blueprint and Machine Metaphors in Biology and Science Education

To the few this is as clear as daylight, and beautifully suggestive, but to many it is evidently a stumbling-block.
- Alfred Russel Wallace to Charles Darwin, on the use of the metaphor of natural selection

The price of metaphor is eternal vigilance. – A. Rosenblueth and N. Wiener

Abstract.[26] Genes are often described by biologists using metaphors derived from computational science: they are thought of as carriers of information, as being the equivalent of "blueprints" for the construction of organisms. Likewise, cells are often characterized as "factories" and organisms themselves become analogous to machines. Accordingly, when the human genome project was initially announced, the promise was that we would soon know how a human being is made, just as we know how to make airplanes and buildings. Importantly, modern proponents of Intelligent Design, the latest version of creationism, have exploited biologists' use of the language of information and blueprints to make their spurious case, based on pseudoscientific concepts such as "irreducible complexity" and on flawed analogies between living cells and mechanical factories. However, the living organism = machine analogy was

[26] This chapter has been adapted from Pigliucci & Boudry (in press). One technical section on inheritance and gene-centrism has been left out for the purpose of this doctoral thesis.

criticized already by David Hume in his Dialogues Concerning Natural Religion. In line with Hume's criticism, over the past several years a more nuanced and accurate understanding of what genes are and how they operate has emerged, ironically in part from the work of computational scientists who take biology, and in particular developmental biology, more seriously than some biologists seem to do. In this article we connect Hume's original criticism of the living organism = machine analogy with the modern Intelligent Design Creationism (IDC) movement, and illustrate how the use of misleading and outdated metaphors in science can play into the hands of pseudoscientists. Thus, we argue that dropping the blueprint and similar metaphors will improve both the science of biology and its understanding by the general public.

3.1 Introduction: the machine metaphor in biological science and education

Scientific thinking and education are rife with the use of metaphors. Well-known examples include physics (heavy objects on a blanket as an image of space-time curvature), chemistry (atoms as miniature solar systems), ecology (the planet Earth as a homeostatic organism), and many others (Condit, Bates et al. 2002; Brown 2003). In biology and biological education in particular, metaphors are pervasive at almost every level of description and explanation. Brown (2003, p. 159) has even argued that "biology today reveals more forcefully than any other area of science the essential role of metaphor in scientific reasoning and communication." Perhaps this pervasiveness of metaphors is an inevitable result of the way human beings think (Lakoff and Johnson 1980; De Cruz and De Smedt 2010), but it has consequences for both science education and scientific research, and not necessarily for the better. For instance, Fox Keller (1995) made that case concerning information-type metaphors in the broad field of genetics (see also Nelkin 2001), while Martin (1994) has done the same for the more specific area of immunology and HIV-AIDS research.

The simple empirical observation that science talk depends heavily on analogical thinking mandates that we examine the consequences of deploying certain metaphors on how students and the public at large end up understanding science, and that we be attentive to the way pseudoscientists often seize upon these metaphors to foster misunderstandings. Specifically, if we want to keep Intelligent Design out of the classroom, not only do we have to exclude the 'theory' from the biology curriculum, but we also have to be wary of using scientific metaphors that bolster design-like

misconceptions about living systems. We argue that the machine-information metaphor in biology not only misleads students and the public at large, but cannot but direct even the thinking of the scientists involved, and therefore affect the sort of questions they decide to pursue and how they approach them. For instance, there undoubtedly were very good reasons to pursue the Human Genome Project during the 1990s and the early part of the 21st century, but obtaining the "blueprint" for how human beings are made (and therefore cure genetic-based diseases like cancer) was certainly not one of those good reasons. Yet, it was the blueprint rhetoric — based on the actual informatics-directed thinking of the scientists involved — that helped sell the 4.2 billion effort (in 2009 dollars) to the American public[27], not to mention redirect the research time and intellectual efforts of tens of thousands of scientists and graduate students the world over.

Similarly, the never ending debate in science education between creationists and evolutionary biologists hinges on the persistent (mis-)understanding of biological organisms as "machines," an understanding that is perpetuated by biologists themselves in textbooks and lectures. The organism-as-machine metaphor goes much further back in time than the decidedly modern genome-as-blueprint one, to René Descartes and other mechanistic philosophers. Interestingly, as we shall see, attempts to point out its misleading effects are also quite old, beginning with David Hume's devastating critique of the idea of intelligent design.

In this paper we discuss the effect of what we call the machine-information metaphor (or family of metaphors, to be more precise) in the areas of science education and scientific research. In the first case, we examine the 'debate' about Intelligent Design and argue that some defenders of evolution themselves (unwittingly) keep fueling the widespread misunderstanding among the general public, because they accept the fundamental soundness of their opponents' characterization of organisms as machine-like. In the second case, we argue that thinking of genomes as blueprints has not only led the general public astray but may have also delayed the incorporation of developmental biology into evolutionary theory, and is still delaying the expansion of the general concept of heritable information to phenomena that rightly should fall within its purview, like epigenetic, behavioral, cultural and even environmental transgenerational effects. We conclude with a brief look at a few alternative metaphors and suggest that biological research and teaching could and should actually be done without much use of grand metaphors, although the widespread tendency to employ the latter may be used as a critical thinking tool in both general and graduate level education to produce more science-savvy citizens and scientists

[27] See http://www.genome.gov/11006943

3.2 Machine metaphors, Intelligent Design and science education

When delving into unknown territory, scientists have often naturally relied on their experiences in more familiar domains to make sense of what they encounter (De Cruz and De Smedt 2010). In the early days of the scientific revolution, mechanical metaphors proved to be a powerful instrument to get a grip on new discoveries about the living world and the universe at large. According to Niall Shanks, we can trace back the emergence of machine metaphors at least to the Middle Ages, when new achievements of technology had a profound cultural influence and captured the collective imagination (Shanks 2004, pp. 25-27). Against this background of technological innovation, it is not surprising that the pioneers of anatomy and physiology relied on the metaphor of the animal body as a complicated piece of machinery to make sense of their discoveries. The mechanical language provided a richness of meaning and allowed them to structure the new phenomena in terms of familiar experiences (Lakoff and Johnson 1980). For example, the image of the human heart as a pump with intricate mechanical components played an important role in William Harvey's discoveries about blood circulation.

In the course of the 17th century, a new philosophy of nature became prominent that developed a conception of the universe in purely mechanical terms. According to this mechanical philosophy, which was developed by thinkers like René Descartes, Pierre Gassendi and Robert Boyle, the phenomena of nature can be understood purely in terms of mechanical interactions of inert matter (Ashworth 2003). This mechanization of nature proved an important driving force behind the Scientific Revolution, and at the end of the 17th century culminated in Newton's theory of motion. Newton's description of planetary orbits following the fixed laws of gravity conveyed an image of a clockwork universe set in motion by an intelligent First Cause. In fact, that was exactly how Newton conceived the universe and its relation to the Creator. For Newton and many of his contemporaries, the importance of the mechanical conception of nature was greater than the mere term 'metaphor' would suggest, as the development of mechanistic philosophy was itself largely inspired by religious motivations (Ashworth 2003). As Shanks wrote in his account of the history of the design argument, "the very employment of machine metaphors invited theological speculation" (Shanks 2004, p. 32).

In the second part of the 17th century, the mechanical pictures of living organisms and of the cosmos at large converged into an intellectual tradition where theology and science were intimately intertwined: natural theology. The most famous representative of this tradition was William Paley, whose work Natural Theology, of Evidence of

Existence and Attributes of the Deity, Collected from the Appearances of Nature (1802) made a deep impression on the young Charles Darwin. As the title of the book makes clear, Paley and the natural theologians conceived of Nature as a complicated machinery of intricate wheels within wheels, in which every organism has its proper place and is adapted to its environment. According to Paley, the contrivance and usefulness of parts exhibited by living organisms attests to the intelligence and providence of a benevolent Creator. This so-called 'design argument' already had a long intellectual pedigree, dating back to Plato, Cicero and Thomas Aquinas, but its most famous formulation is found in the first chapter of Natural Theology, in which Paley relies on the analogy between living organisms and a pocket watch to support his design inference.[28]

> In crossing a heath, suppose I pitched my foot against a stone, and were asked how the stone came to be there: I might possibly answer, that for any thing I know to the contrary, it had lain there for ever: nor would it perhaps be very easy to show the absurdity of this answer. But suppose I had found a watch upon the ground, and it should be inquired how the watch happened to be in that place; I should hardly think of the answer which I had before given, that for any thing I knew, the watch might have always been there. Yet why should not this answer serve for the watch, as well as for the stone? Why is it not as admissible in the second case as in the first? For this reason, and for no other, viz., that when we come to inspect the watch, we perceive (what we could not discover in the stone) that its several parts are framed and put together for a purpose . . . This mechanism being observed . . . the inference, we think, is inevitable, that the watch must have had a maker; that there must have existed, at some time, and at some place of other, an artificer or artificers, who formed it for the purpose which we find it actually to answer; who comprehended its construction, and designed its use. (Paley 1802, p. 5)

The idea is that without having witnessed the creation of the watch, without even knowing anything about the identity of the designer, the purposeful arrangement of parts forces the conclusion of intelligent design on the observer. Of course, there is at least one fundamental dissimilarity: human artifacts don't propagate, whereas living organisms do. However, Paley did not think this significantly endangers the analogy. Instead, he argued that it actually strengthens the design inference for the case of living organisms. After all, continues Paley, suppose that the watch we found in the heath did not only indicate the time, but was also capable of producing another watch with the

[28] Paley was not the first to pursue the analogy with a pocket watch. In fact, Paley borrowed the famous paragraph in the first chapter of his work from the book Regt gebruik der werelt beschouwingen (1715) by the Dutch physician Bernard Nieuwentijt, who was himself probably influenced by thinkers like William Derham, John Ray and Robert Boyle.

same features. Wouldn't this even more "increase [our] admiration of contrivance"? (Paley 1802, p. 9) Living organisms, according to Paley, surpass the ingenuity and complexity of human artifacts "in a degree which exceeds all computation" (Paley 1802, p. 13), and are worthy of a divine Creator alone. Even Kant, in his Critique of Judgment, clearly struggled with the apparently purposeful design of living organisms. Although he was wary of teleological accounts, always preferring "efficient" mechanical causes for explaining the world, he acknowledged that living systems necessarily had to be explained "as if" they were teleological in nature. Thus, he maintained that "it is absurd to hope that another Newton will arise in the future who shall make comprehensible by us the production of a blade of grass according to natural laws which no design has ordered" (Kant 2007 [1790], p. 185). It seems that for beings such as ourselves, used to associate functional complexity with intelligence, it is very difficult to escape the impression of design in nature, and the adaptive complexity of living organisms certainly demands a special explanation. Darwin's theory of evolution by natural selection eventually provided such explanations, fatally undermining Paley's argument, and flagrantly contradicting Kant's pessimism on the matter.

While Darwin was the one who gave the most decisive blow to the design argument by suggesting a natural explanation for adaptive complexity in the living world, many philosophers would agree that David Hume foreshadowed its demise (see Chapter 4), by exposing several problems with the central analogy. In his work Dialogues Concerning Natural Religion (Hume 1998 [1779]), which actually predates Paley's magnum opus by more than 50 years, we find a discussion of the design argument among Philo, the skeptical character that voices Hume's ideas, Demea, the orthodox religious believer, and Cleanthes, the advocate of natural theology.

After Cleanthes has set out the design argument in terms foreshadowing Paley's analogy of the watch, Philo objects that it is dangerous to derive conclusions about the whole of the universe on the basis of a spurious analogy with one of its parts. Given that our experience with design is limited to human artifacts only, we have to proceed with great caution, and it would be presumptuous to take so minute and select a principle as the human mind as the model for the origin of the whole universe.[29] Hume realized that, at least in some cases, appearances of intelligent design can be deceptive. In the words of Philo:

[29] Philo says that "it is a palpable and egregious partiality to confine our view entirely to that principle by which our own minds operate" (Hume 1998 [1779], p. 46).

> If we survey a ship, what an exalted idea must we form of the ingenuity of the carpenter who framed so complicated, useful, and beautiful a machine? And what surprise must we feel, when we find him a stupid mechanic, who imitated others, and copied an art, which, through a long succession of ages, after multiplied trials, mistakes, corrections, deliberations, and controversies, had been gradually improving? (Hume 1998 [1779], p. 36)

In contemplating that "[m]any worlds might have been botched and bungled, throughout an eternity, ere this system was struck out", Hume (p. 36) even comes close to Darwin's crucial insight about the power of natural selection.

Although Hume does not deny that we can discern similarities between nature and human artifacts, he warns us that the analogy is also defective in several respects. And if the effects are not sufficiently similar, conclusions about similar causes are premature. To illustrate this, Philo proposes another possible cosmogony on the basis of the analogy between the world and an animal:

> A continual circulation of matter in [the universe] produces no disorder; a continual waste in every part is incessantly repaired: The closest sympathy is perceived throughout the entire system: And each part or member, in performing its proper offices, operates both to its own preservation and to that of the whole. The world, therefore, I infer, is an animal. (Hume 1998 [1779], p. 39)

Philo further speculates that the world even more resembles a plant, and that it could have come into existence by a process analogous to reproduction or vegetation. While the others protest at his arbitrary speculations, Philo maintains that his analogies, though certainly defective in some respects, are no more so than the machine analogy. "[I]n such questions as the present, a hundred contradictory views may preserve a kind of imperfect analogy, and invention has here the full scope to exert itself" (Hume 1998 [1779], p. 49). Aware of the fallibility and imperfections of human reasoning, Hume remains highly skeptical about the design inference and the machine analogy, even though he was not able to provide a satisfactory explanation for the appearance of design in nature.

In the Origin of Species, Charles Darwin (1859) finally proposed a natural explanation for the phenomenon that inspired Paley but failed to convince Hume. Although the design argument is still of interest to philosophers and historians of science (see Chapter 4), it has been widely discarded in the scientific community. However, the analogy on which Paley based his inference seems to be alive and well, not only in the minds of creationists and IDC proponents, but also in the writings of science popularizers and educators (and even in actual scientific work, as we will see in the next section). Many scientists have actually argued that Paley at least offered an incisive formulation of the problem as there is indeed a hard-to-shake intuition of contrivance and intelligent design in nature. As one of the most ardent defenders and popularizers

of evolutionary theory put it, "Biology is the study of complicated things that give the appearance of having been designed for a purpose" (Dawkins 1991, p. 1). Adaptive complexity, then, is still regarded as something that requires a special explanation.

In textbooks, science educators have presented the comparison of living organisms and man-made machines not just as a superficial analogy, but carrying it out to a considerable level of detail. For example, the cell has been described as a miniature factory, complete with assembly lines, messengers, transport vehicles, etc. Consider the following quote from Bruce Alberts, former president of the National Academy of Sciences:

> The entire cell can be viewed as a factory that contains an elaborate network of interlocking assembly lines, each of which is composed of a set of large protein machines. ... Why do we call the large protein assemblies that underlie cell function protein machines? Precisely because, like machines invented by humans to deal efficiently with the macroscopic world, these protein assemblies contain highly coordinated moving parts. Given the ubiquity of protein machines in biology, we should be seriously attempting a comparative analysis of all of the known machines, with the aim of classifying them into types and deriving some general principles for future analyses. Some of the methodologies that have been derived by the engineers who analyze the machines of our common experience are likely to be relevant. (Alberts 1998, p. 291)

Similarly, in their popular high school textbook Biology, Kenneth Miller and Joe Levine develop an extensive analogy between the living cell and a manufacturing plant Levine and Miller 1994 (Levine and Miller 1994). In Miller's own words:

> The nucleus is the factory's main office, the mitochondria its power plants, the ribosomes its manufacturing equipment, and the Golgi apparatus its shipping and receiving department (Miller 2008, p. 27).

In line with the machine metaphor, scientists have also conceived of the genome as a 'blueprint' for the organism, written in a four-letter alphabet and in a language that scientists have deciphered. In the wake of the Human Genome Project, many scientists have enthusiastically described the human DNA sequence as "the book of life" or "the blueprint for a human being" (for an overview, see Nelkin 2001)). In an interview for Time about the Human Genome Project, biochemist Robert Sinsheimer has described the genome as "the complete set of instructions for making a human being [...] written in the language of deoxyribonucleic acid, the fabled DNA molecule" (Jaroff 1989). According to Miller, machine metaphors are useful because they allow teachers to get across complicated material, since they are "easy to remember and make scientific sense" (Miller 2008, p. 27). However, we will see that analogies between living organisms and machines or programs (what we call "machine-information metaphors") are in fact

highly misleading in several respects. Bearing in mind that metaphors are powerful persuasive tools that can deeply affect the way we look at the world, we think the pervasiveness of extensive machine analogies in science education is in fact unfortunate.

Creationists and their modern heirs of the Intelligent Design movement have been eager to exploit mechanical metaphors for their own purposes (Perakh 2008). For example, Bruce Alberts' description of the living cell as a factory has been approvingly quoted by both Michael Behe and William Dembski, two leading figures in the IDC movement. For IDC proponents, of course, these are not metaphors at all, but literal descriptions of the living world, arching back to Newton's conception of the Universe as a clock-like device made by the Creator. The very fact that scientists rely on mechanical analogies to make sense of living systems, while disclaiming any literal interpretation, strengthens creationists in their misconception that scientists are 'blinded' by a naturalistic prejudice. And of course, the idea of a genomic 'blueprint' is highly congenial to the theistic worldview of IDC proponents.[30] In the creationist textbook Of Pandas and People, which has been proposed by IDC advocates as an alternative to standard biology textbooks in high school, we read that "Intelligent design [...] locates the origin of new organisms in an immaterial cause: in a blueprint, a plan, a pattern, devised by an intelligent agent" (Davis, Kenyon et al. 1993, p. 14).[31]

The analogy between living organisms and man-made machines has proven a persuasive rhetorical tool of the IDC movement (for a thorough examination of IDC, see Pennock 1999; Pigliucci 2002; Shanks 2004). In fact, for all the technical lingo and mathematical 'demonstrations,' in much of their public presentations it is clear that IDC theorists actually expect the analogies to do the argumentative work for them (Young 2004). In Darwin's Black Box, Behe takes Alberts' machine analogy to its extreme, describing the living cell as a complicated factory containing cargo-delivery systems, scanner machines, transportation systems and a library full of blueprints. Here is a typical instance of Behe's reasoning:

> In the main area [cytoplasm] are many machines and machine parts; nuts, bolts, and wires float freely about. In this section reside many copies of what are called master machines [ribosomes], whose job it is to make other machines. They do this by reading the punch holes in a blueprint [DNA], grabbing nuts, bolts, and other parts that are floating by, and mechanically assembling the machine piece by piece. (Behe 2006, pp. 104-105)

---

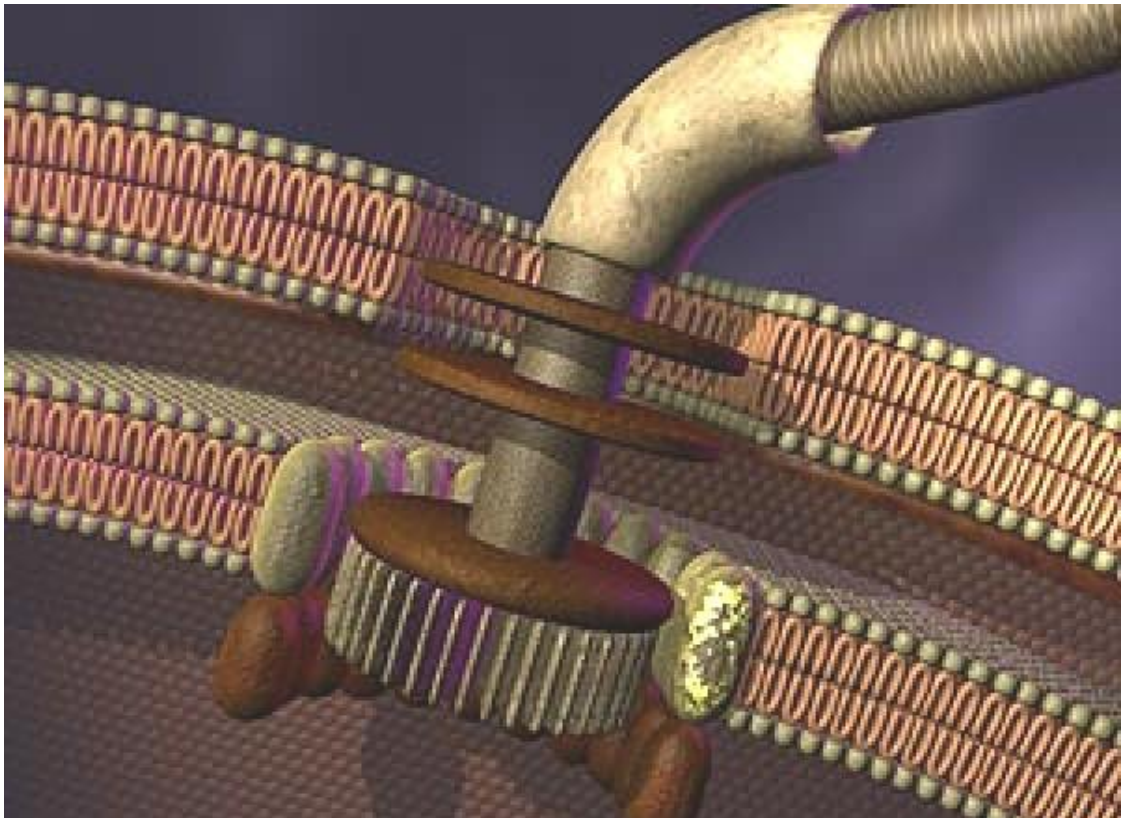
[30] A survey by Condit et al. about the perception of the blueprint and recipe metaphor suggests that deeply religious people prefer the blueprint metaphor precisely because of its theistic connotations (Condit, Bates et al. 2002, p. 312).

[31] Thanks to Stefaan Blancke for this reference.

Behe's favorite model of biochemical systems is a mechanical mousetrap, the familiar variant consisting of a wooden platform, a metal hammer, a spring etc. According to Behe, if any one of these components is missing, the mousetrap is no longer able to catch mice. He has termed this interlocking of parts "irreducible complexity" and thinks it characterizes typical biochemical systems. As Shanks wrote: "[t]he mousetrap is to Behe what the well-designed pocket watch was for Paley" (Shanks 2004, p. 165). But whereas Paley can be excused on the grounds of the state of scientific knowledge in the 18th century, for Behe the situation is a little different. Modern biochemistry, nota bene Behe's own discipline, has revealed that biochemical systems are not like mechanical artifacts at all (Shanks and Joplin 1999).[32] Moreover, even biological systems that are irreducibly complex under Behe's definition pose no problem for evolution by natural selection, see for example Miller (2000) and Chapter 7).

IDC proponents have buttressed their analogies between living systems and mechanical contraptions with a lot of visual rhetoric as well. The flagellum of the bacterium E. coli, the hallmark of the IDC movement, has been represented as a full-fledged outboard rotary motor, with a stator, drive shaft, fuel supply etc. It features on the cover of Dembski's book No Free Lunch and has been used numerous times in presentations and online articles. The idea seems to be that if it looks designed, it has to be designed. But as Mark Perakh has documented, IDC supporters invariably use idealized and heavily stylized representations of the flagellum, in order to make it more resemble a man-made contraption (Perakh 2008).

[32] See our discussion of "brittleness" below.



3.1 A highly stylized and mechanical illustration of a bacterial flagellum, created by Discovery Media Productions, and used as the front cover of William Dembski's *No Free Lunch* (2002)

Another striking example of this visual rhetoric is a new video by Discovery Institute president Stephen C. Meyer[33], which presents a computer-simulated – and again heavily stylized – journey inside the cell, and describes the biochemical processes in terms of "digital characters in a machine code," "information-recognition devices" and "mechanical assembly lines." Meyer commented that evolutionists will have a hard time now dissuading the public from the fact that "the evidence for design literally unfolds before them".

[33] See http://www.journeyinsidethecell.com/

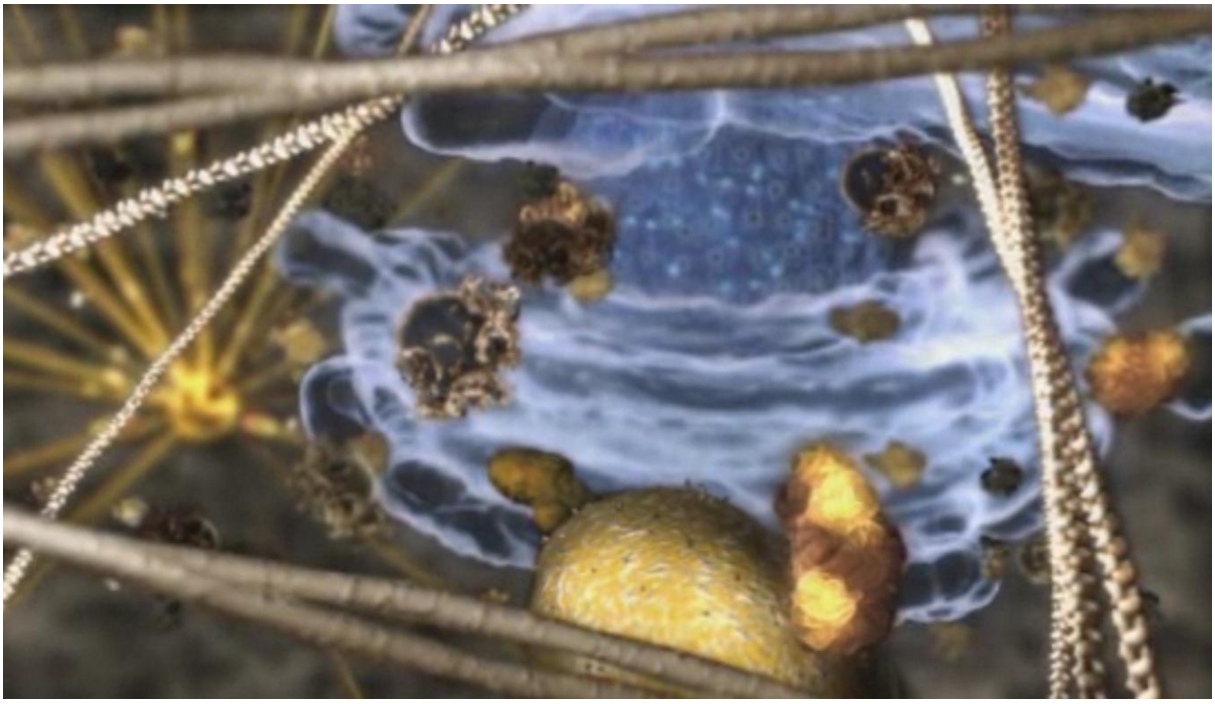


Figure 3 A snapshot from the Discovery Institute's video *Journey inside the cell.*

Of course, the mere observation that creationists have seized on machine metaphors in biology does not suffice to demonstrate that these metaphors do not make scientific sense. However, the fact that they tend to do so systematically, using full-length quotes from respectable scientists, should make us weary of the possible dangers of misleading metaphors. If the rhetoric of the IDC movement is demonstrably based on these mechanical analogies, it can be instructive to reexamine their scientific merits. In the next section, we argue that the machine-information analogy has indeed influenced the way scientists themselves think about biological structure, function, and evolution. By analyzing the consequences of and reactions to this analogy in actual biological research, we show that its scientific merits are very week indeed, and that its place in modern biology has become questionable.

3.2 Another idealized reproduction of the bacterial flagellum, featuring on numerous IDC websites and on the homepage of William Dembski's blog *Uncommon Descent*

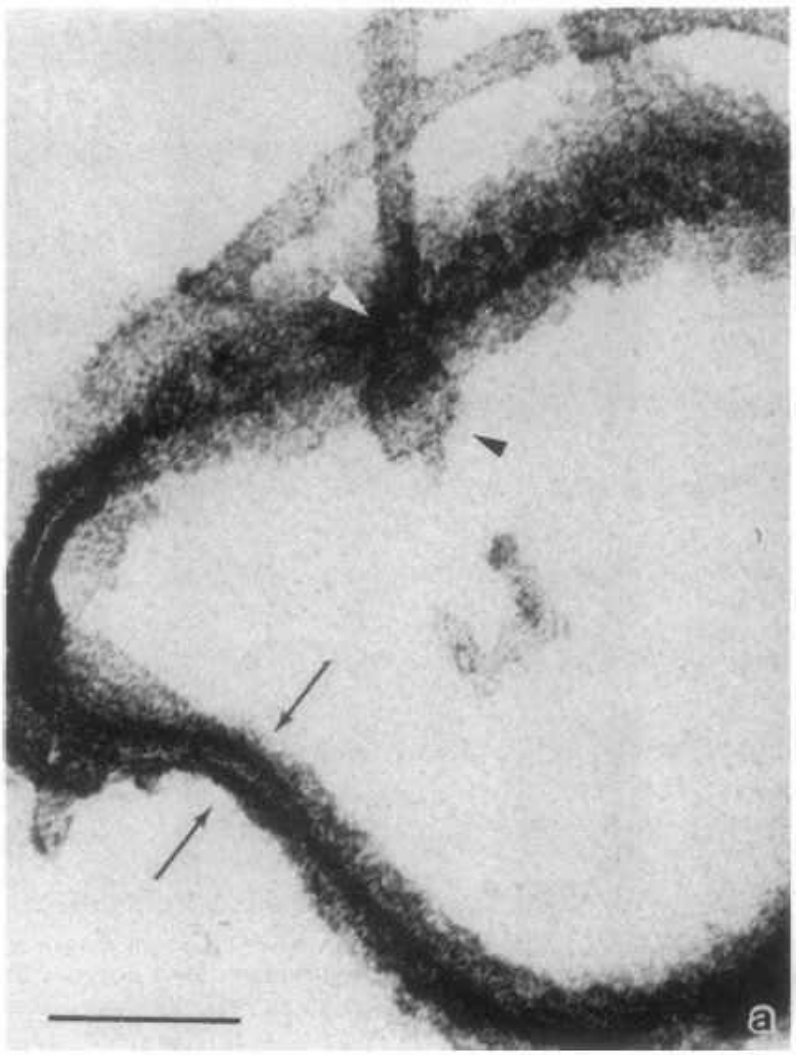


Figure 4 An electronic micrograph of a real bacterial flagellum, which resembles nothing like a man-made contraption.

3.3 Machine metaphors and the practice of biological research

The idea of the organism as a machine, a device that stores and acts on information, has permeated not just education about biological science, but the practice of that science itself. Descartes (1972 [1648]) in L'Homme developed the fundamental idea that living organisms, including human beings (but in the latter only as far as the "vegetative" aspects of the body are concerned) are machines, whose function can be understood in

terms of simple mechanical forces and interactions. This was part of Descartes' overall program of countering Aristotelian science and developing a new physics based on the "mechanical philosophy" that was inspiring Galilei and later found full fruition in Newton, giving birth to what we recognize as modern science. Biology, however, has always also been characterized by the presence of an anti-reductionist, vitalistic streak, which ironically periodically arches back to Aristotle (350BE/1991) and his conception of vegetable and animal "souls" as presented in his De Anima.

This back and forth between mechanistic and vitalistic conceptions of living organisms characterized debates among biologists during the 19th century, when Darwin's ideas can be seen as certainly more mechanistic in nature than, say, those of Lamarck. The vitalistic position gained new prominence again at the beginning of the 20th century, largely through the efforts of Henri-Louis Bergson, but evolutionary biology kept moving steadily in a mechanistic direction throughout the Modern Synthesis of the 1930s-40s (Mayr and Provine 1998). The coup de grace was then given to vitalism by the onset of the molecular revolution, first anticipated in Schrödinger's What is Life? (1992 [1944]) and then definitely playing a determinant role in late 20th century biology after the discovery of the structure of DNA (Watson and Crick 1953).

Ever since the 1950s what we might more properly call the machine-information metaphor has been prevalent in molecular biology and, because of the tremendous success of the molecular revolution, throughout the biological sciences (despite continued pockets of resistance within the more organismally oriented disciplines of biology, chiefly ecology and evolutionary biology). It is sometimes argued by biologists that the use of machine-information metaphors is limited to popular writings of the type discussed in the previous section, and that they do not inform actual research papers. But this is quickly dispelled by an even cursory examination of databases such as Web of Science or PubMed.

Let us take, for instance, the idea that genes are "blueprints," i.e. that they contain the information to build proteins or more broadly any aspect of the phenotype. Hassoun et al. (2009) talk about "differentiation of the principal body axes in the early vertebrate embryo [being] based on a specific blueprint of gene expression" and say that the "mouse and rabbit show distinct structural differences in APD [anterior pregastrulation differentiation] and the[ir] molecular blueprint." Iimura et al. (2009) refer to "the colinear disposition of Hox genes expression domains [which] provides a blueprint for the regionalization of the future vertebral territories of the spine" in vertebrates. Uttamchandani et al. (2009) are confident that the sequencing of the human genomes "provided a wealth of information about the genomic blueprint of a cell" (though they do acknowledge that this does not provide "the entire story" of life and living processes). Rutka et al. (2009) affirm that "the human genome project has been completed providing a blueprint for the human species." Saminathan et al. (2009) maintain that their "neuronal transcripts were further analyzed to provide a genetic

blueprint that can be used by neurobiologists to unravel the complex cellular and molecular mechanisms underlying biological functions." And the list could easily be extended to encyclopedic proportions.

One of the obvious patterns emerging from any such search of the recent literature is that words like "blueprint" are rarely if ever used within the context of organismal (as opposed to molecular) biological research, and indeed there is a sustained effort on the part of (some) ecologists and evolutionary biologists to counter what they see as the hyper-reductionist approach brought about by the molecular revolution. This is particularly evident when we focus on ongoing discussions on the scope and results of "evo-devo" (evolution of development), the relatively new field that is supposed precisely to bridge the gap between organismal and molecular biology, all the while finally bringing developmental biology into the broad fold of the standard theory of evolution known as the Modern Synthesis (e.g. Minelli and Fusco 2005; Hendrikse, Parsons et al. 2007; Muller 2007; Pigliucci 2009). This resistance against the machine-information metaphor, however, cannot and should not be read as a resurgence of vitalism, as no evo-devo author has moved in that direction. Rather, it is the result of a genuine tension between the undeniable successes of the molecular-reductionist approach on the one hand, and the limits that such approach seems to reach when it tackles issues pertinent to the structure and evolution of complex phenotypes.

We argue that part of this tension is in fact the result of, or is at least fostered by, the deployment of the machine-information metaphor as a guiding idea of the molecular biological research program (see the sample of recent references provided above), and that new ways of thinking about development and evolution are building a conceptual vocabulary that increasingly distances itself from the machine-information metaphor. Let us consider two broad categories of examples, what we will be referring to as "the problem of development" and "the problem of environment".

The problem of development has arguably been present in one form or another throughout the history of biology, and in modern shape constitutes the central aim of the research program in evo-devo. If living organisms are sufficiently analogous to programmable machines, then the problem of development largely reduces to identifying which pieces of the "program" (i.e., genes) "control" which parts of the hardware to be assembled (the organism itself). That this is a line of inquiry actually pursued by biologists over the past few decades — and not just a matter of idle talk — is evident from the bewildering literature on "genes for" a particular phenotype, even though the very notion of a gene being "for" a phenotype is in fact justified only in very restricted cases (Kaplan and Pigliucci 2001). While it can certainly be argued that the approach has been successful, in reality that success is largely limited to one of two areas: the identification of few complex phenotypic traits that do show a relatively simple "mapping" between genotype and phenotype, e.g. eye color in vertebrates (Sturm and Larsson 2009), or the cataloguing of large numbers of genes affecting a given

phenotype (often, in the case of humans, a disease-related one), where however each genetic element statistically accounts for a minute fraction of the variation in the phenotype, and often only in a particular subset of populations within a given species (Tian, Gregersen et al. 2008; Wu and Zhao 2009).

Recent advancements in theoretical biology and computational science may help us to articulate the fundamental reasons why the problem of development cannot be solved by more and better Genotype ⇒ Phenotype mapping. Ciliberti et al. (2007) have pointed out that "direct encoding systems", such as human-designed software, suffer from "brittleness", that is, they break down if one or a few components stop working, as a result of the direct mapping of instructions to outcomes. If we think of living organisms as based on genetic encoding systems — like blueprints — we should also expect brittleness at the phenotypic level which, despite the claims of creationists and IDC supporters that we have encountered above, is simply not observed. On the contrary, biological developmental systems tend to be very robust to both internal (i.e., genetic) and external (i.e. environmental) perturbations. In other words, pace Behe and other IDC advocates, removing one component in a complex biochemical pathway typically does not cause the system to "effectively cease functioning" (Behe 2006, p. 39). Indeed, the fact that biological organisms cannot possibly develop through a type of direct encoding of information is demonstrated by calculations showing that the gap between direct genetic information (about 30,000 protein-coding genes in the human genome) and the information required to specify the spatial position and type of each cell in the body is of several orders of magnitude (Stanley 2007). Where does the difference come from?

An answer that is being explored successfully is the idea that the information that makes development possible is localized and sensitive (as well as reactive) to the conditions of the immediate surroundings. In other words, there is no blueprint for the organism, but rather each cell deploys genetic (Johannes, Colot et al. 2008) information and adjusts its status to signals coming from the surrounding cellular environment, as well as from the environment external to the organism itself. The way this works, then, is through two phases: in the signaling phase information is deployed locally, within a given circuit (in computational science models) or cell (in biological systems). The second phase is that of the expression of a particular functional status, which depends on the input received so far by each circuit or cell. In computational science, interestingly enough, this approach is known as "developmental encoding" or "artificial development" and is in fact inspired by a more realistic view of what sort of systems living organisms actually are (Hartmann, Haddow et al. 2007).

One of the most interesting outcomes of shifting our thinking from direct/genetic encoding to indirect/developmental encoding of information is that we then have an immediate link between developmental biology and evolution: not only is research on localized encoding showing it to be a better model for understanding development, but

it turns out that artificial systems based on developmental encoding are much more efficient at searching for more robust solutions to whatever problem is posed to them, i.e. they evolve faster than genetic encoding systems and produce phenotypes that are fault-tolerant — to use software engineering terminology — because they are not brittle (Hartmann, Haddow et al. 2007): giving up talk of blueprints and computer programs immediately purchases an understanding of why living organisms are not, in fact, irreducibly complex.

Developmental encoding has yet another advantage over genetic encoding, which is proving interesting to software engineers while simultaneously shedding light on the way living organisms develop: in the case of direct/genetic encoding, the length of the program grows proportionally to the complexity of the phenotype, which quickly makes the system unwieldy and again slows down its evolution. By comparison, indirect/developmental encoding means that a relatively small number of "instructions" can produce a variety of phenotypes, depending on the interactions among parts of the system and among these and the external environment. Complex phenotypes, then, can be evolved without the necessity to also evolve proportionally large genetic systems (Roggen, Federici et al. 2007). Looking at evolution through these lenses may also provide us with insights into one of the fundamental questions in biology: why did development evolved in the first place? As it turns out, when the phenotypes are simple, genetic and developmental encoding are roughly equally efficient at evolving new solutions, because the genetic system is not too complicated. It is only when more complex phenotypes are favored that the advantage of indirect encoding becomes apparent (Roggen, Federici et al. 2007): perhaps this is why comparatively simpler life forms like bacteria do not need developmental encoding and better approximate the simple Genotype ⇒ Phenotype mapping assumed by the blueprint metaphor. But when evolution began to favor — for whatever reason — more complex, multicellular life forms, a new way of encoding information also evolved.

While much of the preceding discussion was framed in terms of the "problem of development," the second issue, which we referred to above as "the problem of environment", is actually conceptually analogous and can be thought of in a similar fashion. Biologists have known since immediately after the beginning of genetics, in 1900, that the same genotype often develops different phenotypes in different environments. This is difficult to make sense of if one thinks of genomes as a simple blueprint-like reservoir of information. Accordingly, this phenomenon — known as phenotypic plasticity (Pigliucci 2001) — has remained largely in the background of biological research for most of the 20th century.

During the last couple of decades, however, studies of phenotypic plasticity have taken center stage in evolutionary biology, ecology, and even molecular biology, because of the realization of the near-universality of the phenomenon. As usual, the initial approach to the study of the genetic basis of plasticity was guided by the

blueprint metaphor, with researchers attempting to identify and map "genes for" plastic responses of a variety of phenotypes to a variety of environmental conditions. It quickly became clear, however, that plasticity is a complex developmental phenomenon, which requires a more nuanced approach and is no different, in principle, from the study of any other complex outcome of development. Indeed, if standard development can be thought of as the response of indirect/developmental encoding to the internal environment surrounding each cell, then plasticity can be seen as the similar response of indirect/developmental encoding to signals originating in the environment external to the organism (Jablonka 2007).

If the problem of environment is conceptually analogous to the problem of development, and both require a more sophisticated view of how organisms deploy genetic information, then we begin to see the possibility that the notion of genetic "information" itself is not quite so straightforward, and certainly is not one that fits comfortably with ideas like blueprints and machines. While there is no question that the "molecular revolution" has been a central and positive development in biology, and indeed in science in general, throughout the second part of the 20th century, it is also becoming increasingly clear that the overly reductionist approach inspired and fueled by machine-information metaphors is running out of steam and needs to be replaced with more sophisticated and realistic thinking (a kind of reasonable, or non-greedy reductionism, so to speak). Is it then time to retire metaphors like blueprints and machines, and to seek an alternative way to conceptualize biological organisms, or would it perhaps be better to abandon the use of metaphors in this field altogether?

3.4 The search for new metaphors

In their classic work on metaphors, Lakoff & Johnson (1980) argue that the basic function of metaphorical concepts is to structure a new kind of experience in terms of a more familiar and delineated experience. In science as well as in everyday language, metaphors highlight particular aspects of whatever it is we are trying to grasp, but they will inevitably distort others. For example, the image of the 'tree of life,' with new species branching off as budding twigs and extinct species as dead branches, is an instructive approximation of the relations of evolutionary descent. However, it can also foster misconceptions about 'progress' in evolution, or lead to a simplistic conception of speciation events, or to a downplay of horizontal gene transfer and reticulate (i.e., by inter-species hybridization) speciation events. To give one more example, in physical

chemistry the model of the atom as a miniature solar system, with electrons orbiting the nucleus as planets, though still having wide public appeal, is fundamentally inaccurate.

Of course, no metaphor will do its job perfectly, but it is crucial to realize, as Lakoff & Johnson (1980) have shown, that the widespread deployment of a particular metaphor can have a feedback effect on the way we perceive things, not just how we present them to others. In the examples discussed in this paper, the lure of machine-information metaphors in the history of biology has invited scientists to think of genomes as 'blueprints' for organisms, written in the four-letter alphabet of DNA and readable in a manner analogous to a computer code. But as we have argued, the machine-information conception of living systems has led both the public and the scientific community astray.

In response to this problem, some scientists and science educators have proposed several alternative and improved metaphors to characterize the relationship between genotype and phenotype. Biologist Patrick Bateson, for instance, was probably the first to compare the DNA sequence of living organisms with a recipe for a cake (Dawkins and Wong 2005, p. 414). The idea of a genetic recipe has several advantages over the blueprint metaphor, the most important being that it takes into account pleiotropy (one gene affecting more than one trait) and epistasis (gene-gene interactions), and that it is more sensitive to what we termed the problem of environment and the problem of development in the previous section. As a consequence, the simple picture of a one-to-one (or close to) correspondence between particular genes and phenotypic traits is abandoned, which becomes clear when one considers that there is no way to locate particular ingredients in individual crumbs of a cake (Dawkins 1991, pp. 295-296). Accordingly, there is no possibility of reverse-engineering the end product to the set of procedures (the 'recipe') that made the final product possible.

Of course, if carried too far, the recipe metaphor can in turn be quite misleading. To get the desired result, a cook has to lump together different ingredients in the correct proportions, and follow a set of instructions for handling the dough and preparing the oven. But as we saw, developmental encoding is an enormously more complex and very different sort of procedure, which is also highly dependent on epigenetic factors and unpredictable vagaries of the external environment. The expression of specific genes in the course of development resembles nothing like the way a cook handles the ingredients of a recipe. Living organisms are also highly differentiated in a number of functional parts or components (cell types, tissues, etc.), in contrast with the homogenous cake that comes out of the oven. Moreover, the genome is not written in anything like a "language," as in the case of a recipe, and it certainly does not contain a description of the desired end product in any meaningful sense of the word "description".

Condit et al. have discussed the recipe metaphor as an alternative to talk of blueprints, pointing out that it was adopted "with surprising swiftness" by science popularizers and the media in the 1990s (Condit, Bates et al. 2002, p. 303). However, they also remark that, as a new "master metaphor" to capture the relationship between genotype and phenotype, the image of a recipe for a cake has little to recommend either. For example, evoking recipes can invite people to think of the genome as a step-by-step manual that describes "how to make a human," in that sense falling into the same trap as the idea of a blueprint.

That being said, if contrasted with the blueprint metaphor, the recipe metaphor conveys the point about lack of one-to-one correspondence between genes and phenotypes very well, and hence it highlights an important fact about development and the Genotype ⇒ Phenotype map. If the recipe metaphor is used within this restricted context, for example in explicit contrast with the characteristics of a blueprint, it is immediately clear what are the salient points of connection with living systems, and people are less likely to be misled by stretching the metaphor beyond usefulness. If the recipe metaphor is presented as 'the' alternative to the blueprint, however, it is bound to mislead people no less than its rival.

The same point applies to other interesting metaphors that have been proposed in this context, for example Lewis Wolpert's comparison of early embryonic development with the Japanese art of origami (Wolpert and Skinner 1993; Dawkins and Wong 2005). The analogy highlights the circuitous step-by-step development of the early embryo[34], but of course in a piece of origami art the structure is imposed top-down from an intelligent agent, whereas the functional differentiation in the embryo is regulated bottom-up by a complex interaction between genes and environment. Moreover, origami simply is folded to yield the final product, which in a very real sense is already there from the beginning. This is definitely not the way embryos develop, with their ability to respond to local and external environmental fluctuations.

The general problem that we have been discussing seems to us to be not just that one kind of metaphor or another is woefully inadequate to conceptualize biological organisms and their evolution. It is that it simply does not seem to be possible to come up with a metaphor that is cogent and appropriate beyond a very limited conceptual space. Although some of the alternatives are more accurate than the blueprint metaphor (in some respects), we certainly have not found one that we would recommend as a replacement. Should we therefore try to avoid the use of metaphors in biological teaching and research altogether? Or do we simply expect too much from metaphors in science and education?

---

[34] See also the online video demonstration of origami embryology by Kathryn Tosney and Diana Darnell: http://origamiembryo.cba.arizona.edu/

3.5 Conclusion: metaphors as teaching moments in scientific research and education

Analogical and metaphorical thinking is widespread among human beings, although of course different cultures and historical moments inspire people to use different metaphors. After all, a metaphor is an attempt to make sense of novel concepts by pairing them with known ideas to increase our overall understanding. Metaphorical thinking is therefore part of our language, and language is inextricably connected to our thinking, but to put it as Wittgenstein did: "It is, in most cases, impossible to show an exact point where an analogy starts to mislead us" (Wittgenstein 1972, p. 28). Yet a great part of doing philosophy consists precisely in clarifying our language in an attempt to advance our thinking. To quote Wittgenstein (1951 [2009], § 109) again: "Philosophy is a battle against the bewitchment of our intelligence by means of our language." To complicate matters further, there is emerging empirical evidence that the human brain processes metaphors in a specific fashion: research on Alzheimer's patients, for instance (Amanzio, Geminiani et al. 2008), found that impairment of the brain's "executive" function, associated with the prefrontal cortex, leads to poor understanding of novel metaphors (while, interestingly, comprehension of familiar metaphors is unaffected). Metaphorical thinking seems to be a biologically entrenched functional mode of our brains, and may therefore be hard to avoid altogether.

Both science and philosophy have made ample use of metaphorical and analogical thinking, sometimes with spectacularly positive results, at other times more questionably so. Nonetheless, it seems that nowhere is metaphorical thinking so entrenched — and so potentially misleading — as in biology. Given the maturity of biology as a science, and considering that it deals with objects whose nature is not as alien to our daily experience as, say, those of quantum physics, we do not actually see any good reason for clinging on to outdated metaphors in biological education and research for characterizing living organisms, their genomes and their means of development. Taking into account the fact that the machine/information metaphors have been grist to the mill of Intelligent Design Creationism, fostering design intuitions and other misconceptions about living systems, we think it is time to dispense with them altogether. Still, we are also not as naive as to expect that this advise will be followed by scientists and science educators any time soon, precisely because the machine/information metaphor is so entrenched in biology education. What to do then? We propose two approaches, one for science educators, the other for practicing scientists.

In science education, talk of metaphorical thinking can be turned into a teaching moment. Students (and the public at large) would actually greatly benefit from

explanations that contrast different metaphors with the express goal of highlighting the limitations intrinsic in metaphors and analogies. So, for instance, science educators and writers could talk about the human genome by introducing the blueprint metaphor, only to immediately point out why it does not capture much of what genomes and organisms are about; they could then proceed to familiarize their students and readers with alternative metaphors, say the recipe one, focusing on its differences with the original metaphor while of course not neglecting to point out the (different) deficiencies of the new approach as well. The goal of this process would be to foster a cautious attitude about metaphorical thinking, as well as to develop a broader understanding of how unlike commonsense modern science really is. On the latter point, it is interesting to note, for instance, that a popular refrain among evolution or global warming deniers is that "simple common sense" shows that the scientists are wrong, a position that ignores the proper weight of technical expertise in favor of a folk understanding of nature. It is therefore crucial that the public appreciates the limitations of common sense thinking about science.

There is an analogous teaching moment that can be brought to bear when research scientists engage in unbridled metaphorical thinking: we could refer to this as a philosophy-appreciation moment. Scientists are notoriously insensitive to, or even downright dismissive of, considerations arising from the history and philosophy of their discipline, and often for good practical reasons: modern science is a highly specialized activity, where there is barely enough time to keep up with the overwhelming literature in one's own narrow field of research, and certainly not enough incentive to indulge in historical readings or philosophical speculation. Nonetheless, historians and philosophers of science can easily show the pitfalls of metaphorical thinking (by using well-documented historical examples) and even get across to their colleagues some basic notions of philosophy (by analyzing the effects of particular metaphors on the development of specific lines of scientific inquiry). None of this will quickly amount to overcoming C.P. Snow's (1993 [1959]) infamous divide between "the two cultures," but it may bring about better understanding and appreciation of philosophy by scientists, and perhaps even help science see new horizons that have hitherto been obscured by a superficially illuminating metaphor.

Chapter 4.
Where the Design Argument Goes Wrong: Auxiliary Assumptions and Unification

Ever since the creation of the world his invisible nature, namely, his eternal power and deity, has been clearly perceived in the things that have been made. – Romans 1:20

If only God would give me some clear sign! Like making a large deposit in my name in a Swiss bank. – Woody Allen

Abstract.[35] In *Evidence and Evolution* (2008), Elliott Sober reconstructs the biological design argument in the framework of likelihoodism, and proceeds to demonstrate that it is defective for intrinsic reasons. We argue that Sober's thesis on the adoption of auxiliary hypotheses is too restrictive, as it commits him to rejecting types of everyday reasoning that are valid. We propose a more lenient account of the choice of auxiliaries, based on the explanatory virtue of unification and the avoidance of gerrymandering. If only the design argument satisfied certain theoretical requirements, it could be rendered compelling in ways that violate Sober's restriction concerning the choice of auxiliaries. The present account shows that the design argument indeed fails, but not for the *intrinsic* reasons adduced by Sober. By the same token, Sober's critique of the

[35] This chapter has been provisionally accepted in *Philosophy of Science*.

argument from imperfections and the *argument from evil* against design is off the mark, and fails to appreciate the weight of empirical evidence against the design hypothesis.

4.1 Introduction

Who gave the decisive deathblow to the argument from design on the basis of biological complexity? Both philosophers and biologists are divided on this point (Oppy 1996; Dawkins 1986; Sober 2008; Dennett 2007). Some have claimed that the biological design argument did not falter until Darwin provided a proper naturalistic explanation for adaptive complexity; others maintain that David Hume had already shattered the argument to pieces by sheer logical force several decades earlier, in his *Dialogues Concerning Natural Religion* (Hume 2007 [1779]). Elliott Sober has been among the philosophers who maintain that, as Hume was not in a position to offer a serious alternative explanation of adaptive complexity, it is hardly surprising that "intelligent people strongly favored the design hypothesis" (Sober 2000, p. 36). In his most recent book, however, Sober (2008) carefully develops what he thinks is the most charitable reconstruction of the design argument, and proceeds to show why it is defective for intrinsic reasons (see also Sober 2002). Accordingly, Sober argues that the design argument can be rejected even without the need to consider alternative explanations for adaptive complexity

> To see why the design argument is defective, there is no need to have a view as to whether Darwin's theory of evolution is true.

We argue that Sober's reconstruction suffers from several problems. His requirements regarding the choice of auxiliary hypotheses and his proposed independence relations are overly restrictive, as they commit him to rejecting types of reasoning that are obviously valid. We develop an alternative and more lenient account of auxiliary assumptions, based on the explanatory virtue of unification and the avoidance of gerrymandering. In our view, if only the design argument satisfied certain theoretical requirements, it could be rendered compelling in ways that violate Sober's restriction concerning the choice of auxiliaries. We conclude that the design argument does not

[36] As we will see in 4.2.1, this statement is suprising given Sober's proclaimed contrastive approach to hypothesis testing.

suffer from any intrinsic flaws, but it has simply collapsed under the weight of evidence and has been outcompeted by evolutionary theory, which is all the more damaging to the epistemic status of the design hypothesis. Theoretical immunizations by design theorists and historical examples from natural theology are discussed to support this thesis. An important corollary of our view is that Sober's objections against the argument from evil and the argument from imperfections, which have been leveled *against* the design hypothesis ever since Darwin's seminal work, are equally misguided.[37]

4.2 Likelihoodism and the design argument

4.2.1 Reconstructing the design argument

In his reconstruction of the design argument, Sober wants to arrive at "the strongest, most defensible, version of the argument" and then to show why he thinks the argument is "defective" (2008, p. 113). Sober's reconstruction has three features we should keep in mind. First, it is probabilistic, not deductive. Second, it is contrastive: he does not want to evaluate the design hypothesis in isolation, but only against competing hypotheses (but see 4.5.1 for Sober's ambiguity). Third, he favors a 'likelihood approach' over a Bayesian approach, because he refuses to assign prior probabilities to Darwin's theory of evolution, or to the existence of an intelligent designer, since these merely reflect "a subjective degree of certainty" (Sober 2008, p. 121). Sober applies the law of likelihood to William Paley's *Natural Theology* (1802), in which Paley pursued the analogy between the human eye and a pocket watch to drive home the design argument. Sober (2008, p. 122) arrives at the following reconstruction, where 'ID' is the hypothesis of intelligent design and 'Chance' is the old Epicurean hypothesis of pure chance:

> Observation O favors ID over Chance if and only if Pr(O | ID) > Pr (O| Chance)

This likelihood reconstruction encounters one immediate objection. The value of Pr (O | ID) can be artificially raised to unity by tuning the hypothesis to the observations. For example, if "ID++ = there exists an omnipotent supernatural Creator for whom the creation of the bacterial flagellum is number one priority" and "O = there exists a

[37] For a thorough critique of Sober's likelihood reconstruction of the cosmological design argument, see Weisberg (2005).

bacterial flagellum", then the likelihood P (O | ID++) equals one. But why not build the observational outcome in the competing hypothesis instead? For example, if "Chance ++ = a chance-process produced the bacterial flagellum", then Pr (O | Chance++) is likewise 1. As is clear from these examples, the mere fact that the likelihood of some contrived hypothesis equals one, does not make it any more plausible.

In Sober's words: "[w]ithin a likelihood framework, there is no beating a hypothesis that *entails* the observations" (Sober 2008, p. 131). If we allow the introduction of favorable (or unfavorable) assumptions to the central hypothesis, *in casu* assumptions about the intentions and attributes of the designer, we are left with no way in which an observation O can discriminate between the competing hypotheses. The evidential significance of the observation O "will be thoroughly obscured *if we build the observational outcome into the theories we wish to test*" (Sober 2008, p. 132, emphasis in original).

4.2.2 Restrictions on auxiliary hypotheses

If we want to avoid this problem, we somehow have to introduce restrictions on the choice of auxiliary assumptions for our central hypothesis. Sober's proposed solution is to demand "an *independent reason* for believing assumptions about goals and abilities" (Sober 2008, p. 144, 2002). More specifically, the introduction of an auxiliary hypothesis "must be justified without assuming H_1 or assuming H_2 or assuming O", or more specifically, without assuming that H_1, H_2 or O are true (Sober 2008, p. 145).

> [S]uppose you are on a jury. Jones is being tried for murder, but you are considering the possibility that Smith may have done the deed instead. Evidence is brought to bear: A size 12 shoe print was found in the mud outside the house where the murder was committed, as was cigar ash, and shells from a Colt .45 revolver. Do these pieces of evidence favor the hypothesis that Smith is the murderer or the hypothesis that Jones is? It is a big mistake to answer these questions by inventing assumptions. If you assume that Smith wears a size 12 shoe, smokes cigars, and owns a Colt .45 and that Jones wears a size 10 shoe, does not smoke, and does not own a gun, you can conclude that the evidence favors Smith over Jones. If you make the opposite assumptions, you can draw the opposite conclusion. [...] What is needed is independently attested information about Smith's and Jones's shoe sizes, smoking habits, and gun ownership. (Sober 2008, p. 145)

In relation to the design argument, this means that we cannot simply attribute intentions and motives to the designer if we don't have any independent justification for doing so. For example, from the fact that humans have eyes, we cannot conclude

that the intelligent designer, if such a being exists, must have had the intention for equipping humans with eyes:

> What is needed is evidence about what God would have wanted the human eye to be like, where the evidence does not require a prior commitment to the assumption that there is a God and also does not depend on looking at the eye to determine its features. (Sober 2008, p. 146)

Without independently justified auxiliary hypotheses, we are left with a designer without attributes, and the likelihood that such a designer wanted to (and was able to) create the world we observe cannot be calculated. Thus, the design argument fails.[38]

4.3 Criticism

4.3.1 Background knowledge

Sober's solution effectively prevents the practice of building observations into one's hypothesis, but we argue that it does much more than that and hence is too restrictive. Consider the murder scenario described by Sober, but in a somewhat different light. Does the available evidence provide support for the hypothesis that someone committed a murder in the first place? Suppose the landlord is nowhere to be found, we find blood stains and broken glass in his bedroom, and we possess all the other evidence Sober alludes to. If a detective wants to assess the murder hypothesis, we submit that she is justified in making the additional assumption that the hypothesized murderer, *whoever* it was, wears a size 12 shoe, smokes cigars and used a Colt .45.

> O = a size 12 shoe print, cigar ash, and shells from a Colt .45 revolver were found in the bedroom.
> H = the victim was murdered.
> A_1 = the murderer wears a size 12 shoe, smokes cigars and used a Colt. 45.

[38] In his *Philosophy of Biology* Sober made a more provisional assessment: "Thus far, no argument has been stated that allows one to know which auxiliary assumptions should be adopted. Perhaps this will change, but until that happens, creationism cannot be tested" (Sober 2000, p. 54).

What justifies our adopting auxiliary hypothesis A_1? In the first place, we are informed by background knowledge (K) on human beings wearing shoes, occasionally smoking cigars and even less occasionally murdering people, and on the Colt. 45 producing specific shells. But note that K, by itself, does not warrant our adopting A_1. Only the conjunction of K with O does. Does the choice of A_1 'depend' on looking at O in a way that is not allowed by Sober? It seems so. At this point Sober's requirement commits him to rejecting types of everyday reasoning that are obviously valid, but there is a charitable way to reconstruct his argument, by fine-tuning the dependence relation as follows:

> If you want to construct an auxiliary hypothesis A_n for testing H with respect to O, then, though your adoption of A_n may be informed by O, it must be so in conjunction with *at least* one other, independent reason. By contraposition, if your choice of A_n is solely informed by O, then you are not conducting a proper test of H.

Sober's intrinsic objection against the design argument can then be rephrased: as there is no independent background knowledge available about the designer over and above the empirical data we possess of, the biological argument does not get off the ground. We argue that this weaker version of Sober's argument is still indefensible, and that if only certain evidential standards were fulfilled, the design argument would not be in need of independent background knowledge.

4.3.2 Ruling out uninteresting assumptions

If Sober is right, design theorists are never justified in adopting such or such auxiliary, because they are completely in the dark as to the identity of the designer. But this seems to be too restrictive. In a preparatory stage of investigation, in which the theorist tries to bring her candidate hypothesis "into contact with the observation" (Sober 2008, p. 145), she concentrates on eligible auxiliary hypotheses, and she pays no attention to those that are extremely unlikely to yield the data we want to explain. If the detective wants to consider the murder hypothesis, and he finds Colt. 45 shells around the blood stains, he makes the additional assumption that someone murdered the victim with a Colt. 45, even if, at that point, the victim has not been found and no further evidence supports his tentative hypothesis.

Sober's criterion seems to frustrate even this kind of tentative fleshing out of competing hypotheses. For example, suppose that William Paley, reflecting on the origin of the human eye, constructed the following design hypothesis, conjoined with two additional assumptions:

H = The human camera eye was created by an intelligent designer.
A_1 = The designer is interested in creating camera eyes.
A_2 = The designer is capable of designing something as complex as the camera eye.

The adoption of both A_1 & A_2 seems reasonable enough, since their negation is completely uninteresting, in the sense of being very unlikely to yield the data in question:

$\sim A_1$ = The designer has no interest at all in creating camera eyes.
$\sim A_2$ = The designer is a bungler completely incapable of producing anything as complex as the camera eye.

Evidently, the likelihood of both H & $\sim A_1$ and H & $\sim A_2$, viz. Pr (O | H & $\sim A_1$) and Pr (O | H & $\sim A_2$) is extremely low. If we follow Sober's approach, however, this gives us no reason for adopting A_1 & A_2, because, in the absence of background knowledge about the designer, the independence rule is violated. But is this really where Paley's design argument goes off the rails?

If we pause to think about it, there do not seem to be many ways of justifying the introduction of an auxiliary *except* by taking the observations into account which we set out to explain. Merely having independent reasons for accepting an auxiliary is not sufficient. Take for example:

A_1* = Naive set theory suffers from Russell's paradox.

Arguably, Pr (O | ID & A_1*) = Pr (O | ID) and Pr (O | Chance & A_1*) = Pr (O | Chance). Even if we have (very good) independent reasons for accepting A1*, there is no use incorporating it as an auxiliary, because it has no bearing on our observations in any way.[39] In short, unless we take into account the observations with which we want to bring the hypothesis into contact, we have no idea how to exclude such uninteresting auxiliary as A_1*.

[39] The fact that A_1* is not an empirical statement is irrelevant here, since the same holds for A_2* = Rome is the capital of Italy.

4.4 Introducing auxiliaries

4.4.1 Gerrymandering and spurious unification

The problem central to Sober's concern is the practice of gerrymandering hypotheses by inventing *ad hoc* auxiliaries to fit the data. This type of reasoning is pervasive in much creationist writings, and its problems were already spelled out by Darwin, in his discussion of the theory of special creation:

> On the ordinary view of the independent creation of each being, we can only say that so it is;—that it has pleased the Creator to construct all the animals and plants in each great class on a uniform plan; but this is not a scientific explanation. (Darwin 2006, p. 677)

As Darwin noted elsewhere in the Origin of Species, the theory of special creation amounts to "restating the fact in dignified language" (Darwin 2006, p. 336). It is designed to yield the known observations *and nothing more*. By contrast, the main explanatory merit of evolutionary theory lies in its power to yield a "consilience of inductions" (Whewell 1840), by bringing together a wide array of facts from different domains and explaining them as following from the same basic principles: blind variation, heritability and selective retention (Kitcher 1985).

The design theorist might object that his hypothesis also accomplishes this kind of unification, as every observation in the natural world is subsumed under the explanation of "God's will".

> H = God wants it to be the case that O_1...O_n.

As Kitcher (1981, p. 528) has pointed out, however, in such an explanatory pattern the "nonlogical vocabulary which remains is idling" (Kitcher 1981, p. 528) The pattern does not impose constraints on the sentences that can be derived by using it, and thus it is able to accommodate any observation whatsoever. In Sober's vocabulary, the only reason for adopting the assumption that God really wants a specific fact O_1 to be the case depends on looking at O_1 and nothing else. Thus, in reality the design theorist simply posits a new divine disposition for each and every observation, and not a single unifying explanation. As Sober (2008, p. 181) writes, "the fact that the model postulates a single *designer* is besides the point".

Sober's principle that "auxiliary assumptions must be justified without assuming that O is true" effectively undermines this type of "spurious unification" (Kitcher 1981, p. 528), together with other ways of gerrymandering, but does it leave any room for reasonable consideration of auxiliaries? The following *reductio* he offers in support of his

thesis is ineffective (Sober 2008, p. 145): If we assume that O is true, then so is the disjunction "either H_1 is false or O is true". If we take this disjunction as auxiliary hypothesis A_1, then H_1 & A_1 entails O, even if H_1 has nothing to do with O. Thus, according to Sober, we cannot allow our auxiliary hypothesis to depend on O.

However, from the fact that one intuitively illegitimate move happens to violate Sober's rule, it does not follow that *any* violation runs into similar problems. To think otherwise is to commit the fallacy of affirming the consequent. What Sober needs is an argument to the effect that, *whenever* his rule is violated, we are indeed dealing with a move that is epistemically suspect. In the next sections, we demonstrate how the design hypothesis could be made compelling in ways that violate Sober's restriction concerning the choice of auxiliaries.

4.4.2 Unification

A hypothesis can derive empirical support either by accommodating known observations in particular ways, or by successfully predicting new observations. Predictivists attach special epistemic status to successful predictions, but some philosophers have questioned this different assessment (Harker 2008). We will first focus on the case of accommodation, as we think that the 'mere' accommodation of known data *in an appropriate way* would already make the design argument convincing.

The bugbear of accommodation is the temptation of the theorist to "overfit" the data, which consists of sacrificing the simplicity of one's hypothesis in order to attain a maximal fit with the available data (Hitchcock and Sober 2004). However, even philosophers who attribute special epistemic value to prediction acknowledge that accommodation need not be problematic, only that prediction guards the theorist against the temptation of overfitting.

> It is possible to accommodate data without overfitting them, but when one is accommodating data, the temptation to overfit is always present. By contrast, when one accurately predicts new data that were not used in formulating one's theory, there is no opportunity to overfit those data. (Hitchcock and Sober 2004, p. 20)

An appropriate measure against overfitting consists of balancing simplicity against fit with data, so that any loss of simplicity must be offset by a *sufficient* gain in fit with data (not just any gain of fit, see also Forster and Sober 1994; Leplin 1975). The ideal hypothesis, if any such is allowed by the available observations, is one that is both sufficiently simple and achieves a maximum data fit. For example, the murder hypothesis H is to be preferred if and to the extent that the detective, on adopting some suitable and simple set of auxiliaries $A_1...A_n$, succeeds in *unifying* the available

circumstantial evidence $O_1...O_n$ in a way that cannot be accomplished as successfully *without* assuming H. Recall the observations: a size 12 shoe print, cigar ash, and shells from a Colt .45 revolver are found in the bedroom, we can see blood stains and broken glass on the floor, and the landlord is nowhere to be found.

> H = the landlord has been murdered.
> A = the murderer wears a size 12 shoe, smokes cigars and used a Colt. 45.

It is not difficult to invent other hypotheses with suitable auxiliaries that also entail the observations. For example, "H* = the landlord left for an unexpected walk", conjoined with the following auxiliaries: somebody just threw a stone through the window, the shells from the Colt .45 dropped out of a visitor's pocket, the landlord just slaughtered a pig in his house before his unexpected walk, etc. Or, alternatively, someone with inscrutable intentions has planted all the evidence. It is clear, however, that the murder hypothesis is superior, because it succeeds in unifying all the available data under a simple assumption.

This is not to say that H outcompetes every possible hypothesis. For example, "H**= the landlord went underground" can account for the data if conjoined with the auxiliary "A**= in order to fake his own death, the landlord has left the shells, the blood stains, ...". Arguably, H** & A** is not far more complex than H & A, and it is equally unifying. Therefore, H** is an admissible competitor for H.

How does this translate to the biological design hypothesis? Consider William Paley's *Natural Theology* (1802), which made a deep impression on the young Darwin. The main argument in *Natural Theology* states that adaptive complexity in the living world bears the mark of a designing intelligence:

> Arrangement, disposition of parts, subserviency of means to an end, relation of instruments to a use, imply the presence of intelligence and mind. (Paley 1802, p. 12)

Perceptive of the explanatory virtue of unification, Paley enumerates a wide variety of examples of contrivance and usefulness in nature, and he points out the coherence of animal body plans:

> [I]n comparing the eyes of different kinds of animals, we see in their resemblances and distinctions one general plan laid down, and that plan varied with the varying exigencies to which it is to be applied. (Sober 2008, p. 33)

Apart from noting such similarities, however, Paley seems unable to discern any overall intentional plan in the creator's work, making only vague gestures in that direction. For example, a consideration of the bountiful diversity of nature "might induce us to believe that *variety* itself [...] was a motive in the mind of the Creator, or with the agents of his will" (Paley 1802, p. 372).

Accordingly, Paley is unable to infer much about the designer's attributes and specific intentions, except for the point that he must have been at least as powerful and wise to be able to create all things we currently observe:

> The attributes of such a Being, suppose his reality to be proved, must be adequate to the magnitude, extent, and multiplicity of his operations. (Paley 1802, p. 474).

In the penultimate chapter of *Natural Theology*, Paley attempts to demonstrate at least the goodness of the creator. Tellingly, however, he makes recourse to convoluted rationalizations to explain away the preponderance of evil in the world, notably to the argument that God's ways are inscrutable to humans (see 4.4.4). In the end, Paley's does not flesh out his design hypothesis any further, and he places his money on the explanatory *necessity* of a designer, for even a single instance of purposeful contrivance:

> [w]ere there no example in the world of contrivance except that of the eye, it would be alone sufficient to support the conclusion which we draw from it, as to the necessity of an intelligent Creator. (Paley 1802, p. 81)

Modern IDC advocates have made little progress since Paley. To the extent that they have made attempts at all towards unification, they have mainly accomplished one of the "spurious" sort, attributing every particular observation to God's will and explaining that He moves in mysterious ways (see section 4.4.4). What is interesting for our critique of Sober's likelihood reconstruction, is the fact that the design argument *might* have achieved genuine unification, if only its advocates had succeeded in subsuming a wide array of natural phenomena under the assumption of a simple and distinct creative intention on the designer's part (or a simple set of intentions). If only a few 'parameters' in the design hypothesis were to provide an elegant explanation for phenomena that resist any conceivable naturalistic explanation, it seems that our worries about overfitting would be assuaged. The fact that the choice of auxiliaries about the designer's intentions and attributes ($A_1...A_n$) would depend on the observations we set out to explain ($O_1...O_n$), without the support of independent background knowledge, would then be of no concern.

In what way could the design argument achieve such genuine unification? Reflecting on the vast number and variety of beetle species on earth, the biologist J. B. S. Haldane once quipped that the Creator, if He exists, has an "inordinate fondness for beetles". Assuming, for the sake of the argument, that Haldane was making a serious theological point, as it stands his design argument is not very persuasive. Suppose, however, that Haldane happened to discover that beetles have minuscule Hebrew letters written on their shields, forming edifying Biblical messages. Let's say subsequent research demonstrates that beetles all over the world display these microscopic patterns, that they are encoded in beetle DNA, and that the fossil record suggests that beetles displayed these remarkable features even before humans arose on the scene.

The scenario is rather outlandish, but it will suit our purposes. There is no way in which Hebrew letters, as opposed to meaningless scribble, could confer any selective advantage on beetles, either through natural selection or sexual selection. It is even less plausible that the phenomenon would be the result of genetic drift or would be the by-product of other evolutionary adaptations. In general, it is very hard to see how the explanatory repertoire of the naturalistic scientist, consisting of blind and unguided processes, could succeed in explaining anything like the existence of Hebrew beetle decorations (Clarke's third law notwithstanding, see section 1.5).

In the described case, the design hypothesis, conjoined with an auxiliary hypothesis about the designer's abilities and intentions, would allow us to explain otherwise puzzling phenomena.

> H = Beetles are created by an intelligent designer
> A_1 = The intelligent designer has the ability to create beetles, is inordinately fond of them, and he has used their bodies to inscribe his Word

One could object that, even in such an unlikely event, all the available evidence for naturalistic evolution still stands, and one anomaly does not suffice to undermine a well-substantiated scientific theory (Oppy 1996, p. 534). The point is well taken but ineffective, as we could easily fancy a world in which *all the phenomena* of biology would converge on the intelligent work of a creator who, judging from his works, bears a suspicious likeness to the Judeo-Christian God. For example, suppose that *all* living organisms in this world bore an autograph in Hebrew, unique for the species to which they belong, and that all the characters together formed the words of the Old Testament. Suppose, moreover that we would not witness any of the examples of imperfections, rudimentary organs and botched designs that are currently viewed as betraying an evolutionary heritage (see section 4.4.4). Or if this is not sufficient, think away the fossil record, the biogeographical and anatomical evidence for evolution, the evidence from genetics and embryology, etc. Surely there must be *some* point at which the evidence would tilt in favor of intelligent design at the expense of evolutionary theory. And so it should be. Which theories we can reliably accept about the world, depends for a large part on contingent matters of fact, on how the world looks like. An adherent of Sober's approach, however, would be unmoved even by such a fanciful scenario, because the adoption of auxiliary A_1 (the properties of the Judeo-Christian God) still depends upon looking at O_1...O_n (without independent background knowledge).

This example illustrates that the problem with the biological design argument as it stands is not so much that it relies on observations of living organisms to provide the theorist with some clues as to the character and intentions of the alleged designer, but that it yields *nothing beyond* those observations. Thus, although we agree with Sober that we need some "independent" reasons, broadly construed, for adopting auxiliaries A_1...A_n, over and above the mere observations we set out to explain, we think Sober has

construed these reasons too narrowly, neglecting the role played by explanatory unification. Sober mistakenly thinks that violating his independence condition always amounts to gerrymandering, apparently because he has extrapolated from a special problem with the construction of auxiliaries to a general assessment of design reasoning.[40]

In fact, our approach is more faithful to Sober's commitment to contrastive hypothesis testing than Sober's own treatment of the design argument: the design argument is currently outcompeted by evolutionary theory (for a recent overview, see Dawkins 2009; Coyne 2009b), but if only design theorists would come up with evidence that defies all explanatory efforts in a naturalistic framework, and that is elegantly explained on some suitable design hypothesis (in the sense discussed above), they would certainly deserve our attention. It will be clear that this assessment is *all the more* damaging to the intelligent hypothesis (see 4.5.2).

Interestingly, in his discussion of the model selection approach, Sober himself hints at the theoretical possibility of a design model of the living world that accomplishes such a unification. This is somewhat at odds with his assessment of the design argument as suffering from the "devastating objection" (Sober 2008, p. 126) about its auxiliary assumptions. From the perspective of the model selection framework, in order to evaluate the design hypothesis we have to "collect different observations together and view them as consequences of a single plan that the designer(s) has in mind" (Sober 2008, p. 182). We think this is a much more promising tack, but Sober quickly brushes the idea aside: "How should this be achieved? I don't know: this is a task for intelligent-design theorists to address." True enough, but it is also the task of the philosopher assessing the design argument to find out whether such unification would be possible *in principle*. If this is indeed viable, and if we are right that the attributes of the designer

---

[40] An additional problem of Sober's approach concerns the different non-trivial ways in which we can separate the central hypothesis from auxiliary hypotheses. For example, returning to Haldane's beetles, we could reconstruct different design arguments:

H = God created the world and all living beings separately.
O = There are a lot of beetles.
A_1 = God has an inordinate fondness for beetles.

An alternative reconstruction would be to break H further up in a core hypothesis and a number of auxiliary hypotheses, for example:

H* = an intelligent being X created the world.
A_2 = X created all living beings in the world separately.
A_3 = X is omnipotent, benevolent and omniscient (and God is the only person with these attributes).

Where does the design argument go wrong, according to Sober? Depending on how we slice up the cake, different propositions will count as auxiliary hypotheses. If we take up the first reconstruction, Sober will find only the additional assumption A_1 about God's fondness for beetles problematic (because it depends on O), whereas in the second reconstruction, the very attribution of omnipotence and benevolence to X, and the proposition about X's *modus operandi*, will be disallowed by Sober (because we don't have independent reasons for accepting A_2 and A_3).

would have to be inferred from the observations we set out to explain, then Sober's intrinsic argument against the design argument is off the mark.

4.4.3 Prediction

In the murder investigation we discussed, the detective *need* not predict the exact location of the murder weapon in order to convince the jury of his case, although of course such a feat would surely be impressive. Likewise, although the design hypothesis *need* not predict novel phenomena, this would of course be a way of boosting the plausibility of the design hypothesis dramatically. As we noted in the previous section, predictivists attach special epistemic status to successful predictions. Predictivism comes in many flavors, and some of these flavors have faced some important criticisms, most notably coming from Sober himself (Hitchcock and Sober 2004). Hitchcock & Sober distinguish between global and local predictivism. The first "maintains that a theory which successfully predicts some observation will always be superior to one that accommodates the same observation" (Hitchcock and Sober 2004, p. 3). For the latter, prediction is only sometimes superior to accommodation. Hitchcock & Sober's sympathies lie with the local variant, as they show that there are cases in which accommodation is better than prediction. For the sake of the argument, however, suppose we are confronted with a strong predictivist who would be unimpressed by the ability of the design hypothesis to unify and accommodate known data. Is there any 'possible world' in which the design hypothesis can also achieve predictive success in addition to explanatory unification?

Suppose that many different organisms bore an autograph in Hebrew, unique for the species to which they belong. Suppose also that these form a very large part of the Old Testament, except for some missing quotes. Then the hypothesis H and the auxiliary A_1 of the previous section may be used to predict that there exist organisms we have as yet not discovered or not studied carefully enough, which bear the requisite inscriptions (maybe some verses from the book of Jonah on the fin of a new whale species). If we are able to predict what are the missing inscriptions, this furnishes us with an extra reason to accept H & A_1, in addition to their presumed unificatory power.

It is not entirely clear, however, how Sober's restrictions apply to predictions of new data, as opposed to accommodation of known observations. If the design argument would allow us to make successful predictions of phenomena that have a very low probability on any non-contrived naturalistic hypothesis we can think of, would Sober still refuse to accept it? In the prediction case, the observation O that we use to test our competing hypotheses cannot enter into our considerations for choosing auxiliaries $A_1...A_n$, because, by definition, O has not been observed yet. In what sense is the "independence" of $A_1...A_n$ to be understood? Is it acceptable if our justification of $A_1...A_n$

depends on other observations that are already known? If so, why does Sober not leave room for such cases of predictive success in setting up his intrinsic argument against design? In any case, our argument does not hinge on Sober's approach disallowing some forms of successful prediction, as we have already demonstrated that it precludes valid forms of explanatory unification.

4.4.4 Imperfections and evils

Many philosophers and scientists, starting with Darwin himself (for recent examples, see Avise 2010b, 2010a; Coyne 2009b, pp. 86-91), have argued that the clumsy and botched works of nature provide evidence *against* the design hypothesis. According to Sober, both the *argument from imperfections* and *argument from evil* (Sober 2004)[41] fall victim to the same objection which he leveled against the design argument itself, namely that they make unwarranted assumptions regarding the character and intentions of the alleged designer. For example, discussing Stephen Jay Gould's famous argument about the clumsy design of the panda's pseudo-thumb (Gould 1980), Sober charges Gould with simply "inventing assumptions" about the designer to reach a pre-established conclusion (Sober 2008, p. 128).

Sober's criticism is off the mark for both a specific and a more general reason. First, Sober fails to see that these arguments are put forward *in the particular context* of the widespread belief in a benevolent and omnipotent Creator with a purposeful creation plan. As soon as one accepts these traditional assumptions about the designer, as most theists do, including intelligent design theorists (Forrest, Barbara C. and Gross, Paul R. 2007), the pervasiveness of botched design and especially the existence of needless suffering *is* most damning (Mackie 1955; Hume 2007 [1779]). It goes without saying that, if one relinquishes some of the traditional attributes of God, the argument from evil no longer has any force. But the same does not apply to many instances of the argument from imperfections, which brings us to the second and more general problem with Sober's argument.

Even if a design theorist is not committed to any particular religious doctrine about the designer's attributes, the existence of puzzling imperfections and rudimentary organs, together with the countless instances of ineffective and wasteful processes, should worry her nonetheless. These senseless and botched structures present a challenge not just to the traditional theological account (as in the case of the argument

[41] Sober (2008, pp. 164-167) rehearses the same line of reasoning, but he writes that he is not so sure anymore whether this puts the biological design argument on a par with the argument from evil.

from evil), but to *any* attempt at subsuming the phenomena of the living world under a coherent design plan, and thus to any attempt at genuine unification. This is not the proper place to enumerate examples or to give a full account of the argument from bad design, but let us briefly point out its logical impact. The thrust of the argument from imperfections is to *disintegrate* the design hypothesis and its auxiliaries. More specifically, the pervasiveness of imperfections of the particular kind that only make sense on an evolutionary understanding, forces the design theorist either to invent a particular intention on the part of the designer for each new observation, or to state that the designer must have wanted the world to look as though it evolved. In other words, the argument from imperfections challenges the sensible auxiliaries of the design hypothesis, and leaves over only the contrived ones (Kitcher 1993, pp. 18-25).

The surest indication that the unification of biological phenomena under the design hypothesis, conjoined with some suitable auxiliaries, is an all but impossible challenge, is the fact that those who are eager to make a scientific case for design have never took up the challenge to do so (Dennett 2007). Indeed, William Paley himself made only vague gestures in that direction. Instead of fleshing out their design hypothesis, IDC theorists have insisted that the designer is inscrutable and his intentions unfathomable (see for example Johnson 1991, p. 67; Behe 2006, p. 223).[42] They have even accused their critics of making unwarranted assumptions about the intelligent designer (Nelson 1996). For example, Michael Behe wrote:

> Another problem with the argument from imperfection is that it critically depends on a psychoanalysis of the unidentified designer. Yet the reasons that a designer would or would not do anything are virtually impossible to know unless the designer tells you specifically what those reasons are. (Behe 2006, p. 223)

Surprisingly, but in conformity with his view on auxiliary assumptions, Sober grants that this is a "good reply by creationists" (Sober 2007, p. 4). However, we submit that Behe's response is an all too convenient way of insulating the design argument against empirical objections without adding any empirical substance to the theory (Pennock 1999, p. 249). Behe's insistence on the inscrutability of the designer is not a sign of sensitivity to a pressing epistemological problem which his critics have overlooked, but it is an epistemological retreat that is symptomatic of a degenerated research program (Boudry and Braeckman 2010). As Philip Kitcher noted:

---

[42] From a political perspective, the IDC movement wants to brush aside theological quarrels and fight for a common cause. The 'minimal' design hypothesis is interesting because IDC advocates think it allows them to circumvent the Establishment clause against the teaching of religion (Forrest, Barbara C. and Gross, Paul R. 2007).

> As the evidence accumulates, creationists increasingly must take refuge in responses Darwin saw as unsatisfactory evasions, appealing to the thought that these properties of life are unfathomable mysteries. (Kitcher 2007, p. 58)

The evasive arguments of modern creationists indicate that there is no non-contrived way to flesh out the design hypothesis that will stand up to the facts, no matter what auxiliary hypotheses one adopts. Taking into account that the living world, and especially the peculiar examples of 'bad design', looks very much like the kind of world we would expect if there was no design at all but only mindless natural processes at work, the biological design hypothesis is effectively dead.

4.5 Discussion

4.5.1 Conclusion

Sober has correctly identified the main problem with the likelihood reconstruction of design reasoning, but his solution does not hold water. To demand that auxiliary hypotheses be justified "independently" of the available data one sets out to explain is overly restrictive, and commits one to rejecting forms of obviously valid reasoning. Even on a charitable reconstruction of his independence relation, Sober's intrinsic objection against the design argument fails, as he has mistakenly identified the creationist practice of gerrymandering as an inevitable trap into which all observation-based introduction of auxiliaries must fall.

In our view, advocates of Intelligent Design are perfectly free to construct auxiliary hypotheses about the intentions and attributes of their designer, provided that these assumptions are elegant and unifying, and are not just tailored to individual observations (which they often are). In fact, *pace* Sober's likelihood approach, this is what a reasonable critic would *demand* from them (Dawes 2007, pp. 78-79; Pennock 1999, pp. 199-201; Dennett 2007). As long as design theorists fail to flesh out their hypothesis, we are left with an unnamed and unknown designer, and we can do nothing beyond restating the facts in dignified language. Not surprisingly, therefore, design theorists have insisted that the whole affair is unfathomable, thus dodging the issue altogether.

Sober's treatment of the design hypothesis sits uncomfortably with his intended "contrastive" approach to hypothesis testing: "*to test a hypothesis requires testing it against alternatives*" [emphasis in original] (2008, p. 52). This departure from the ethos of contrastive evaluation is all the more remarkable since, as we have seen, Sober (2000, p.

36) himself has earlier noted that, before a serious alternative explanation for adaptive complexity became available, it was not surprising that "intelligent people strongly favored the design hypothesis" and that "Darwin entirely altered the dialectical landscape of this problem." However, in *Evidence and Evolution,* Sober (2008, p. 125) rejects that very claim in almost exactly the same wordings.[43]

4.5.2 Taking the design argument seriously (and then rejecting it)

Ever since the design argument was formulated, there have been philosophical attempts to demonstrate that it is guilty of some fundamental flaw of reasoning, and that we do not need an alternative explanation to see why this is so. Spinoza and Hume's Philo were among the first to argue against the program of natural theology (for more recent examples of this approach, see Oppy 1996; Pigliucci 2002; Scott 2004). We think that the design argument was difficult to resist before the advent of Darwin's theory, even though Hume had already pointed out some of the damaging problems it faces; note, however, that this historical assessment does not necessarily follow from the argument developed in this paper. Even if one accepts the claim that the design inference is not intrinsically defective, one is still free to maintain that the empirical evidence as it was available to natural theologians before Darwin never favored it (for a critical discussion of the evidential warrant of the design argument before Darwin, see Oppy 1996; Gliboff 2000).

Given the philosophical consensus view that the biological design argument is a failure, is it really important to quarrel over where exactly it goes wrong? We think it is. The fate that befell the design argument illustrates a number of important philosophical issues regarding the choice of auxiliary hypotheses, the problem of gerrymandering, and the explanatory virtue of unification. Moreover, different diagnoses of the design argument are wedded to different assessments of its epistemic status. One unexpected consequence of typical *a priori* or *fundamental* objections to the design argument is that, ironically, they are less damaging to the design hypothesis than *a posteriori* objections (Boudry, Blancke et al. 2010a). If we accept Sober's critique of auxiliary assumptions, not only is the design argument stillborn even before any empirical evidence can be brought to bear on it, but the empirical arguments *against* design will not get off the

[43] In *Evidence and Evolution*, Sober writes that, according to a common opinion among biologists, "Paley reasoned correctly [...] but that the dialectical landscape shifted profoundly when a third hypothesis [Darwin's theory of evolution] was formulated." On the next page, he rejects this position and claims that Paley's design argument has always been flawed. In fact, we think Sober's later position is more consistent with his critique of auxiliary assumptions, which he already developed in (2000). After all, if correct, his argument applies equally to Paley.

ground either (Sober 2008, pp. 126-128). If advocates of design are not allowed to make unjustified assumptions about the designer's attributes, then neither are their critics. Hence, Sober's symmetric critique unwittingly suggests that the critics are equally unjustified in rejecting intelligent design as the advocates are in defending it. We think this is mistaken, and it is a conclusion which Sober would want to avoid.

Part 2 – Method in Madness

Chapter 5.
Immunizing Strategies & Epistemic Defense Mechanisms

A successful pseudoscience is a great intellectual achievement. Its study is as instructive and worth undertaking as that of a genuine one. - Frank Cioffi (1998, p. 115)

Abstract.[44] An immunizing strategy is an *argument* brought forward in support of a belief system, though independent *from* that belief system, which makes it more or less invulnerable to rational argumentation and/or empirical evidence. By contrast, an epistemic defense mechanism is defined as a *structural feature* of a belief system which has the same effect of deflecting arguments and evidence. We discuss the remarkable *recurrence* of certain patterns of immunizing strategies and defense mechanisms in pseudoscience and other belief systems. Five different types will be distinguished and analyzed, with examples drawn from widely different domains. The difference between immunizing strategies and defense mechanisms is analyzed, and their epistemological status is discussed. Our classification sheds new light on the various ways in which belief systems may achieve *invulnerability* against empirical evidence and rational criticism, and we propose our analysis as part of an explanation of these belief systems' enduring appeal and tenacity.

[44] This chapter is based on Boudry and Braeckman (2010). Excerpts of this paper were presented at the Fourth Conference of the Dutch-Flemish Association for Analytic Philosophy at the Catholic University of Leuven (2010).

5.1 Introduction

Skeptics of pseudoscience and the paranormal have been puzzled and sometimes exasperated about the enduring popularity of beliefs that are either very implausible or impossible from a scientific and rational perspective (Benassi, Singer et al. 1980; Shermer 2002; Hines 2003). Although many of these belief systems have been thoroughly debunked, the critical efforts of skeptics are mostly unavailing. In this paper, we discuss the remarkable recurrence of *immunizing strategies* and *defense mechanisms*, which play an important role in the tenacity of these belief systems. We define an **Immunizing Strategy** as an argument brought forward in support of a belief system, though independent *from* that belief system, which makes it more or less invulnerable to rational argumentation and/or empirical evidence.[45] By contrast, an **Epistemic Defense Mechanism** is defined as an internal *structural feature* of a belief system, which has the same effect of deflecting rational arguments and adverse evidence.

5.1.1 The demarcation problem

The idea of immunizing strategies sometimes surfaces in the philosophical debate about the demarcation problem. Karl Popper famously argued that the most distinctive feature of the scientific attitude is the willingness to take bold empirical risks, and that a theory can only be regarded as scientific to the extent that it is open to empirical refutation. Of course, resorting to immunizing tactics to protect one's theory from falsification is doing exactly the opposite of taking empirical risks, and hence, according to Popper's view of science, it is the hallmark of pseudoscientific thinking. However, naive falsificationism has been widely abandoned in philosophy of science (Laudan 1983), and the enthusiasm for the demarcation project has waned significantly. A more sophisticated philosophy of science accepts that every scientific research programme builds up a "protective belt" of auxiliary hypotheses around its "hard core" claims (Lakatos and Musgrave 1970; Lakatos 1968). Thus, to a certain extent 'immunizing strategies' can be found in bona fide science as well, and scientists are certainly not

[45] In our immunization metaphor, theory-external arguments are introduced from outside as a means of protection. In that sense, the metaphor is more in line with immunization as a health policy (vaccination) than with the body's internal immune system.

immune to the use of dubious arguments to deflect valid criticism (Hines 2003; Park 2002).

Nevertheless, Popper's basic insight about the value of boldness in conjecture making is still valuable, and survived in a more sophisticated form in Lakatos' philosophy of science. In Chapter 6, we will see that, if the sole effect of introducing an auxiliary hypotheses is to protect the core theory from refutation, without yielding some form of theoretical progress, we are dealing with a degenerating research programme, and hence with bad science. In line with this view, the systematic reliance on defensive maneuvers, evasive arguments and *ad hoc* excuses is still widely regarded as a telltale feature of pseudoscientific discourse (e.g. Derksen 1993; Hines 2003).

5.1.2 Overview

In this chapter we are not so much involved with immunizing strategies as a proposed solution to the demarcation problem. Instead, we present a descriptive classification of different types of immunizing strategies and defense mechanisms, and we discuss their epistemological status. To be sure, many of the belief systems we discuss are traditionally regarded as pseudoscience, but as we are not particularly interested in traditional demarcationism, we will also offer examples from different domains, for instance cult belief systems, pseudo-philosophy, magic and religion. Indeed, one of the main purposes of this paper is precisely to draw attention to the pervasiveness of immunizing strategies and epistemic defense mechanisms across widely different domains.

We provide an overview of the different ways in which a belief system can be rendered immune from criticism and adverse evidence, but our list is not intended to be exhaustive. In discussing these examples, it will transpire that it is often difficult to separate the theory-as-such from the 'immunizing strategies' used by its defenders. Thus, the strict distinction between immunizing strategies and defense mechanisms will be called into question.

5.2 Theory change and degenerating research programmes

How exactly do immunizing strategies and defense mechanisms relate to the problem of theory change in science, in particular to progressive/degenerating research

programmes and pseudoscience? The history of science has witnessed numerous examples of theories that are widely abandoned today but that were initially promising and respectable in the scientific community. What often happens is not that a single crucial experiment discredits a theory in one fell swoop, but that a broader research programme runs into ever more anomalies and conceptual problems, increasingly requiring the sort of changes that Lakatos denoted as "degenerative". Max Planck wrote that science advances "one funeral at a time", but while the sudden demise of a well-established theory has been observed (as when physicists abandoned Newton's theory of physics almost overnight following Eddington's famous eclipse observations), most theories "die the death of a thousand excuses". Examples include phrenology, the theory of 'cold fusion' in physics, Lamarckism in biology, Marxist theories about the law of profit and the demise of capitalism, or more recently, the Duesberg hypothesis on the non-infectious nature of AIDS. Although it is often difficult to identify a moment when a theory or a research programme collapses under the weight of anomalies, it is widely acknowledged that there are certain indications of a degeneration into bad science or pseudoscience.

On the one hand, advocates of a theory may resort to certain generic strategies for protecting a cherished theory from mounting adverse evidence: cherry-picking the data, shooting the messenger, distorting findings, special pleading, discrediting the methods employed in research with unwelcome results, accusing the new 'orthodoxy' of a hidden agenda etc. These generic methods can be broadly construed as 'immunizing strategies', but they are not particularly interesting from a philosophical perspective, and we will not be much concerned with them in this paper.

On the other hand, a theory-in-crisis is often belatedly *modified* by its advocates so as to be less vulnerable to refutation, by introducing *ad hoc* elaborations (see Chapter 6) and special clauses that explain away apparent failures and reduce the empirical content of the theory. As every scientific theory makes use of a protective belt of auxiliary hypotheses, these amendments can seem scientifically respectable at an early stage, and there is often no clear point at which they collapse into pseudoscientific immunizing strategies.

Insofar as it is possible to separate the original theory from later (pseudoscientific) modifications, we prefer to use the term 'immunizing strategies', which are *brought forward* at some point to rescue the original theory from refutation. However, in more complicated cases, these protective strategies progressively become integral part of the theory proper (see examples of parapsychology and psychoanalysis in 5.3.5.1). They are not any longer 'strategies' to which its advocates resort when the theory runs into trouble, but they have become integrated in the explanatory resources and conceptual structure of the theory. We designate these internal, structural features as epistemic 'defense mechanisms'.

In some interesting cases, as we will see, the defense mechanisms of a belief system even follow naturally from its conceptual nucleus (see the example of conspiracy theories in 5.3.3). As a result of their inbuilt defense mechanisms, these belief systems exhibit a self-perpetuating rationale, and are particularly interesting from an epistemological perspective. This topic will be further developed in Chapter 11. We will also briefly return to the conceptual distinction between immunizing strategies and defense mechanisms in 5.4.1.

5.3 Immunizing strategies & epistemic defense mechanisms

5.3.1 Conceptual equivocations & moving targets

Pseudoscientists often make use of conceptual equivocations to transform their theory into a moving target. This may be achieved in two different ways: either one makes a series of ambiguous and open-ended claims that are construed in such a way that one can conveniently switch back and forth between specific and broad interpretations. Alternatively, one defends a theory that appears specific and exciting on a first inspection, but when it runs into trouble, one belatedly deflates it to make it trivial or uninteresting. The two immunizing strategies are sometimes difficult to distinguish, and some successful pseudoscientists will use them in tandem.

5.3.1.1 Multiple endpoints

Skeptics have often remarked that astrologers and soothsayers shy away from making bold and specific statements. For example, the predictions in a horoscope typically have multiple endpoints (Gilovich 1991, pp. 58-59), so that they can be matched retrospectively to a wide range of events. In fact, a 'good' horoscope contains predictions that are amenable to both specific interpretations and a range of broad and metaphorical ones. Inevitably, people will perceive some matches with real events, and they will immediately see this as the *intended* or *real* interpretation of the prediction, ignoring the other ways in which it might have been 'borne out'. As a result, they will be

unduly impressed by its accuracy.[46] Even if not all predictions yield an uncanny match with a real event, the astrologer – or the naive interpreter – can always resort to one of the broad interpretations, thus avoiding the impression that a prediction has failed. Naturally, people will tend to remember the predictions belonging to the first category. Thus, the technique of multiple endpoints creates an in-built asymmetry between what will count as hits and misses for astrological predictions, the effect of which is to immunize astrology against refutations.

As an example of this asymmetry, consider the case of Nostradamus' prophecies, which, as is well-known, allow for almost unlimited allegorical and metaphorical interpretation (Marks 2000, pp. 262-266; Hines 2003). As soon as interpreters have found a 'fit' with actual historical events, the congruency seems so compelling that they are unable to read the prophecies in any other light. The abundance of quatrains and the problem of multiple endpoints guarantee that people will find a lot of tenuous matches, some of which even look striking. As for the other predictions, one can readily persist that these have yet to be borne out, or that their 'true' meaning has not yet been discovered.

As there is nothing in astrological theory that dictates the use of equivocations and multiple endpoints, we may regard this technique as an *immunizing strategy* used by some astrologers to forestall predictive failure, as opposed to a *defense mechanism*. On the other hand, the practice is so common that it has become inseparable from the field of astrology, and several authors have devised convenient rationalizations for it. For example, Nostradamus explained that he deliberately obscured his predictions so as to avoid persecution by the Inquisition.

5.3.1.2 Deflationary revisions

As an example of the second type, consider the case of the Jehovah's witnesses who, after the prediction of the Second Coming of Christ in 1873-74 failed to come true, argued that Christ *had* returned as predicted, albeit as an invisible spirit being (Zygmunt 1970, p. 931). Zygmunt has demonstrated that, over the course of history, Jehovah's witnesses have consistently "redefined [failed prophecies] in retrospect in a manner which provided nonempirical confirmation" (Zygmunt 1970, p. 934). Often enough, these took the form of deflationary revisions of the original prediction. In Evelyn Waughs novel *Brideshead Revisited*, quoted by philosopher Frank Cioffi (1998, p. 220), the

[46] It is a well-known psychological finding that people have difficulties assessing the specificity of ambiguous statements once they have found a fitting interpretation. For example, people will rate the results of a bogus personality test as accurate descriptions of themselves, even if these results contain only vague and ambiguous claims that are applicable to virtually anyone, a phenomenon that is known as the Barnum effect or Forer effect.

character Rex Mottram is questioned by his Jesuit instructor in order to ensure his good Catholic Faith.

> 'Supposing the Pope looked up and saw a cloud and said "It's going to rain," would that be bound to happen?' 'Oh yes, Father.' 'But supposing it didn't?' He thought a moment and said, 'I suppose it would be sort of raining spiritually, only we were too sinful to see it.'

Cioffi has documented an interesting case of belated deflationary revisions in Freud's theory of the libido. As is well-known, Freudian psychoanalysis makes the sweeping claim that the root of all neuroses is to be found in repressed 'libido'. Freud's intended interpretation was clearly sexual. For example, one can only understand why fathers threaten their sons with *penile amputation* if one accepts that the desires of the sons were very carnal indeed. Freud elevated this sexual etiology of all neuroses to a central dogma of psychoanalysis, and he derided others when they were compromising on this point. When presented with empirical difficulties, however, Freud resorted to just such a fuzzier interpretation, widening the scope of the libido concept so as to make it encompass "what Plato meant by 'Eros' and St. Paul by 'love'" (quoted in Cioffi 1998, p. 16). For example, in Freud's explanation of the 'war neuroses' following the First World War, only a deflationary interpretation of 'libido' as a general kind of self-love allowed him to maintain his sweeping universality thesis. In the case of Freudian psychoanalysis, the equivocations surrounding the concept of libido (and other pseudoscientific concepts, see section 5.3.2 and also Chapter 10) were arguably always an integral part of the theory, so that one may properly speak of a *defense mechanism*.

The strategy of belated deflationary revisions is also rampant in a great deal of postmodernist and social constructivist literature, where it is used in tandem with a maneuver in the opposite direction. André Kukla has coined these strategies "switcheroos" and "reverse-switcheroos":

> One commits a switcheroo by starting with a hypothesis that's amenable to a range of interpretations, giving arguments that support a weak version, and thenceforth pretending that one of the stronger versions has been established. (2000, p. x)

What Kukla terms "reverse switcheroos" corresponds to what we term a deflationary reinterpretation:

> you put forth a strong version of the hypothesis, and when it gets into trouble, you retreat to a weaker version, pretending that it was the weaker thesis that you had in mind all along. Switcheroos and reverse switcheroos can be performed in tandem, and the cycle can be repeated ad infinitum. A judicious application of this strategy enables one to maintain an indefensible position forever. (2000, p. x)

A skilled pseudoscientist switches back and forth between different versions of his theory, and may even exploit his own equivocations to accuse his critics of misrepresenting his position. Philosopher Nicholas Shackel has termed this strategy the "Motte and Bailey Doctrine" (Shackel 2005; see also Fusfield 1993), after the medieval defense system in which a stone tower (the Motte) is surrounded by an area of open land (the Bailey):

> For my purposes the desirable but only lightly defensible territory of [...] the Bailey, represents a philosophical doctrine or position with similar properties: desirable to its proponent but only lightly defensible. The Motte is the defensible but undesired position to which one retreats when hard pressed. (Shackel 2005, p. 298)

Analogous to Kukla's analysis of switcheroos, Shackel argues that a successful application of this strategy requires a "systematic vacillation between the desired territory and retreating to the Motte when pressed" (Shackel 2005, p. 298). Again, this retreat to the Motte corresponds to what we call deflationary revisions.

Some recent examples of these Motte and Bailey strategies can be found in the literature of Intelligent Design Creationists. In Chapter 7, we will argue in detail that the central concept of "irreducible complexity" introduced by Intelligent Design advocate Michael Behe vacillates between an empirically adequate but somewhat trivial observation, and an exciting but completely unfounded claim (see also Chapter 8 for similar equivocations regarding the concept of teleology). Behe writes that a system is irreducibly complex when the removal of any one of its components lead to a breakdown in functioning. He has argued that "any precursor to an irreducibly complex system that is missing a part is by definition non-functional" (Behe 2006, p. 39), and that therefore evolution by natural selection is ruled out. However, evolution by natural selection often works by indirect routes and by co-opting existing systems to perform other functions. When pressed on this point, Behe retreats to a deflationary interpretation of irreducible complexity, which simply amounts to the claim that some biological systems cease functioning when one or more components are removed. But after he has given arguments for this defensible but uninteresting position, Behe again proceeds to use his concept as though it posed a major problem for evolutionary theory. This equivocation, which allows Behe to keep on moving the goalposts, is inherent in the very definition of irreducible complexity, so that one may regard it as a defense mechanism of Behe's Intelligent Design Creationism (Boudry, Blancke et al. 2010b).

The work of William Dembski, another leading theorist of the IDC movement, is similarly based on bait-and-switch strategies. In the use of his notion of complex specified information (CSI), Dembski systematically switches back and forth between Shannon's mathematical definition of information, which is simply a measure of

randomness, and the common notion of information as "meaningful message" (Perakh, Mark 2004, pp. 64-75).

5.3.2 Postdiction and feedback loops

Unobservable entities are routinely invoked in scientific explanations, and there is nothing wrong per se with theories that make use of them. However, if particular defense mechanisms are present, belief systems about unobservable entities and their causal workings are completely impervious to falsification. The recipe for such a belief system is as follows: postulate the existence of certain invisible or imponderable causes to account for a range of phenomena, and maintain that the working of these causes can only be inferred *ex post facto* from their effects. Provided that the effects themselves are hard to assess, and that the causal relations in the belief system are not sufficiently specified, believers can get entangled in subtle feedback loops between theory and observations, which keep the belief system forever outside the reach of empirical refutation.

As an example, consider the belief in the efficacy of rituals and magical interventions. Anthropologists have noted that the question whether a performed ritual is 'genuine' is often underspecified by its constitutive components, and can only be determined *ex post facto* dependent on the expected outcome. If the result is successful, one can infer that the intervention was the right one and was properly performed. If it was not, obviously 'something must have gone wrong' during the intervention, or the intervention was not of the appropriate type. Indeed, the very idea of a failed ritual loses any meaning, because any apparent failure 'shows' that it was just not performed properly, or not with the right material, or that some other and equally invisible force interfered with the ritual (additional immunizing strategies are possible). According to anthropologist Evans-Pritchard, belief in ritual efficacy is protected by a whole repertoire of "secondary elaborations" for explaining away particular failures in the expected effects (Evans-Pritchard 1965). In this way, the general causal principles *themselves* remain immune from disconfirmation.

As a result, the taxonomic identification of objects as having certain magical properties, or the identification of a person as a 'real' shaman, tends to feed back into causal assumptions, engendering a form of vicious circularity. As Pascal Boyer noted: "taxonomic assumptions are the basis of causal expectations, and conversely, causal expectations lead to innovations or corrections in the taxonomic identification" (Boyer 1994, p. 144). For example, the efficacy of a magic spell to chase away evil spirits is assessed on the basis of how the patient's condition develops, but the question whether the patient is now really liberated from these evil spirits is itself determined by the

'genuineness' of the magic spell (which may depend on the conditions of the exorcism or the reputation of the healer).

As another example, consider the belief in the therapeutic power of healing crystals, chakra stimulation or even homeopathy. On the one hand, the causal relations in these belief systems are always underspecified: what kind of crystal is appropriate for which patients, how long it takes for chakras to open, what kind of homeopathic medicine is suitable for which patient, etc. Different interventions are 'allowed' by the belief system, and the one that coincides with the moment of recovery can be used to construe the apparent cause. On the other hand, the therapeutic effects themselves are often difficult to ascertain straightforwardly. For example, what exactly are the visible results of having one's 'energy levels restored' again, or having one's 'chakras released', according to alternative therapists? As a result of these defense mechanisms, causal inferences feed back into each other in a way that always protects the belief system from refutation.

The technique of postdiction also underlies an elegant rationalization for data mining, confirmation bias and explaining away null results in parapsychology (Gilovich 1991, p. 21). Parapsychologists have tried out a whole range of different experimental set-ups and procedures to summon psi phenomena. When confronted with a pattern of alternate successes and failures, many parapsychologists have explained that psi is an elusive and unpredictable force. As a result, they find it easy to interpret the patterns of hits and misses *ex post facto* as the result of the intermittent workings of psychic powers, and to explain away failures as due to settings that were simply not psi-conducive. James Randi recounts the remarkable case of a water diviner who, after being queried why he did not count his failures in a series of experiments, replied that "obviously, when I fail, the powers aren't working at that time, and, after all, I'm counting percentages on the cases where I'm divining, not when I'm just guessing!" (Randi 1981, p. 13), Randi notes that even many highly regarded psi experiments include a series of preliminary "warm-up" sessions. In this context, the technique of postdicting psi-activity is very tempting: "Ok, that was just warming-up" – "It seems I'm getting a little tired" – "There are obviously bad vibes around that are distracting me". On a more academic level, postdiction is often used for rationalizing the practice of data-mining. As Richard Wiseman noted, if a pool of parapsychological experiments is sufficiently extensive and heterogeneous, it is not difficult to "'[explain] away' overall null effects by retrospectively identifying a subset of studies that used a certain procedure and yielded a significant cumulative result" (Wiseman 2010, p. 38).

Depending on one's understanding of parapsychological theory, the practice of postdiction can be regarded as an immunizing strategy or a defense mechanism. If one holds that the characterization of psi as elusive and unpredictable is just a pseudoscientific excuse that has nothing to do with parapsychology proper, one may regard it as an immunizing strategy. By contrast, if one maintains that the elusive and

unpredictable nature of psi is a central thesis of parapsychology (see 5.3.5.1), one may properly call it a defense mechanism of the belief system.

A beautiful example of postdiction can also be found in, once again, Freudian psychoanalysis. In the etiology of psychological illness, Freud hypothesized that there is an unobservable 'quantitative factor' in the patient's libidinal economy that had to be taken into account. After all, according to psychoanalysis, 'normal' persons harbor the same repressed wishes and complexes that are found in neurotic patients. The difference between the two groups is only to be found in the quantitative factor in their mental economy, which ultimately determines if and when an unconscious complex will develop into neurosis. Tellingly, Freud admitted that this factor could only be inferred *ex post facto* to account for the unexpected presence or absence of any given symptom.

> We cannot measure the amount of libido essential to produce pathological effects. We can only postulate it after the effects of the illness have manifested themselves. (Freud 1924, p. 119)

5.3.3 Conspiracy thinking

5.3.3.1 Turning the evidence on its head

Conspiracy theories are very interesting from an epistemological perspective, and certainly deserve a more extensive discussion than the one we can offer within the confines of this chapter (Clarke 2002; Keeley 1999). For our present purposes, we want to highlight the fact that all conspiracy theories share a fundamental template of epistemic defense mechanisms. Conspiracy theories purport to provide an explanation of a historical event that differs markedly from the received view or official account. According to conspiracy theorists, the event in question was brought about by a group of actors who have been secretly pulling the strings behind the scenes, and who have tried to cover up their actions by spreading a false story. This false account is the received view, which they are trying to fool us into believing. However, the conspirators have not been completely successful, and they have left traces that allow the conspiracy theorist to reveal their evil plot.

Conspiracy theorists point to incongruities and anomalies in the official account of events, and try to account for these by constructing a unifying alternative explanation. Brian Keeley (1999, p. 118) has termed these the "errant data" with which conspiracy theories are constructed, and he distinguishes two classes: data that are unaccounted for on the official account, and data that actually contradict it.

If the conspiracy hypothesis should fail to be confirmed by further investigations, however, or if new evidence should turn up that flatly contradicts it, conspiracy

theorists typically turn the evidence on its head, arguing that an apparent accordance with the official story is of course predicted by their theory. After all, successful conspirators may be expected to deliberately lay out forged evidence to lead us astray, to cover up the traces of the secret plot, to bribe those who witnessed the cover-up, etc. As Clarke (2002, p. 135) notes: "the apparent plausibility of the nonconspirational received view is a consequence of the success of the cover story or cover-up, according to conspiracy theorists". This pattern of epistemic defense mechanisms, in which any apparent contradiction can be turned into a confirmation, is a common feature of all global conspiracy theories.[47]

Thus, confronting ardent conspiracy theorists with adverse evidence and eyewitness accounts is generally to no avail. In the believers' eyes, this apparent evidence merely constitutes further proof of the cunning and power of the conspirators. The epistemological situation of the conspiracy thinker reminds one of the hollow face illusion. Either way we look at the mask of a hollow face, from the front or from behind, we always 'see' a normal convex face staring at us (Gregory 1997). In a similar way, no matter how the evidence stares at the conspiracy theorist, he or she will always 'see' the action of conspirators.

[47] For an analogy between conspiracy thinking and Freudian psychoanalysis from an epistemological perspective, see section 10.3.3.

Figure 5 The hollow face illusion. To the left is a mask of a normal convex face, while the right picture is the same mask viewed from behind. Although the left face is concave, we have a strong visual bias to interpret both images as convex faces.

5.3.3.2 Explaining the motives for disbelief

A special defense mechanism implicit in conspiracy theories allows the believer to explain the existence of disbelievers within the framework of the belief system itself. For example, Sigmund Freud thought he was able to account for the 'resistance' of his opponents in psychoanalytic terms:

> Psycho-analysis is seeking to bring to conscious recognition the things in mental life which are repressed; and everyone who forms a judgment on it is himself a human being, who possesses similar repressions and may perhaps be maintaining them with difficulty. They are therefore bound to call up the same resistance in him as in our patients; and that resistance finds it easy to disguise itself as an intellectual rejection and to bring up arguments like those which we ward off in our patients by means of the fundamental rule of psycho-analysis. (Freud 1957, p. 39)

Thus, people attack psychoanalysis because they themselves harbor the repressed wishes and complexes revealed by the theory. Being under the spell of unconscious forces, the critics are not even aware of their unconscious motivations, because these

are 'disguised' for them as rational arguments.[48] As a consequence, any objection, however seemingly reasonable, can be dismissed by the psychoanalyst as unconscious resistance in disguise (Gellner 1985). Hence, it is the perfect joker card of the pseudoscientist. Defenders of Marxism sometimes use a similar immunizing argument, labeling criticism from outsiders as manifestations of 'bourgeois class consciousness', thus demonstrating the very theory the critics were objecting to.[49]

The argument from resistance is not just a form of rhetoric that some psychoanalysts happen to resort to in the face of valid criticism – rather, it is "an imperative emanating from the heart of the psychoanalytic vision" (Crews 1986, p. 14). Indeed, if Freud's model of the human mind were accurate, we would *expect* the kind of disguised resistance he was alluding to. Hence, in our terminology, the argument plainly is an epistemic *defense mechanism* of the psychoanalytic belief system.

Essentially, the argument from resistance is structurally identical to any form of conspirational suspicion that takes the attacks of critics as lending further support to the belief system. This style of reasoning is remarkably widespread, even outside classical conspiracy theories. For example, many creationists believe that evolution is an invention of the devil to deceive faithful Christians and lure them into disbelief. For example, Henry Morris, co-author of the seminal work *The Genesis Flood* that sparked the Young Earth Creationism movement in the 1960s, actually believes that the theory of evolution was given by Satan himself to Nimrod, at the Tower of Babel. Morris wrote that "[b]ehind both groups of evolutionists one can discern the malignant influence of 'that old serpent, called the Devil, and Satan, which deceiveth the whole world'" (Morris 1963, p. 93).

In his conspiracy book on UFOs and alien abductions, history professor David Jacobs explains that the evidence for his views is so weak and sketchy because the aliens have carefully installed a "wall of secrecy" (Jacobs 1998, p. 117): they "cloud" the experience of their abductees, implant false memories, and they "perceptually alter potential witnesses" (Jacobs 1998, p. 112). In this way, skepticism and disbelief is easily explained: "The aliens have fooled us. They lulled us into an attitude of disbelief, and hence complacency, at the very beginning of our awareness of their presence" (Jacobs 1998, p. 258). An even more extreme example of this defense mechanism is found in the way Scientology members handle criticism from outsiders. Notoriously, Scientologists systematically try to silence their critics by spreading false allegations and smearing

[48] It is not even clear that 'we' concocted those arguments rather than a mental entity that is independent from 'us', which is precisely what caused Wittgenstein to remark that Freud had made an "abominable mess" of the reasons and causes of our behavior.

[49] According to Popper (2002), in contrast with Freudian psychoanalysis, Marx's initial theory was predictive and not without scientific merits, and it degenerated into pseudoscience only when some of his defenders resorted to ad hoc revisions and immunizing tactics.

their reputation.[50] In an internal policy letter, founder Ron L. Hubbard makes clear that critics can only have one incentive for attacking Scientology:

> There has never yet been an attacker who was not reeking with crime. All we had to do was look for it and murder would come out. [...] They fear our Meter.[51] They fear freedom. They fear the way we are growing. Why? Because they have too much to hide. (Foster 1971, p. 134)

Interestingly, this alleged motive for attacking the Church is explained by the theory of Dianetics in terms that are almost identical to Freudian psychoanalysis. According to Scientologists, people naturally have a 'reactive mind' full of unconscious impressions and traumas that are called 'engrams'. Members of the Church are called Clears, because they are liberated from the influences of this reactive mind. The non-members of the Church, however, which are called 'pre-clears', are still struggling with their engrams, and they will try anything to hide them from view. Hence their attacking Scientology.

5.3.4 Changing the rules of play

By undermining the standards of reasoning employed in a rational debate, one can safeguard one's position from valid criticism. In many instances of this immunizing strategy, the very attempt at criticism is condemned as fundamentally misguided. Sometimes, reasons for this short-circuiting of criticism are dictated by the belief system itself, in which case we have to do with an epistemic defense mechanism (see our account of Lacanian psychoanalysis in Chapter 9).

For example, according to postmodernist philosophers and radical social constructivists, there are no objective canons of rationality, only different social constructions of rationality that are all equally valid. Therefore, occupying a position at all amounts to pretending that some positions are more defensible than others, which is already misguided. Therefore, the postmodernist tries to occupy what philosopher Nicholas Shackel has termed the "No-Position Position" (Shackel 2005, pp. 311-319). This conveniently allows him to "use normative notions of rationality while evading accountability to rational standard" (Shackel 2005, p. 312). The self-excepting nature of the "No-Position Position" reminds one of what the philosopher David C. Stove termed the Ishmael Effect, after Ishmael's epilogue to Melville's Moby Dick: "and I only am

[50] Ron L. Hubbard wrote: "Don't ever defend. Always attack. Find or manufacture enough threat against them to cause them to sue for peace. Originate a black PR campaign to destroy the person's repute and to discredit them so thoroughly they will be ostracized..." (Foster 1971).
[51] The E-meter is an instrument used by Scientologists to measure stress and detect engrams.

escaped alone to tell thee". It refers to the claimed ability of some philosophical theory to escape the fate to which it condemns all other discourse.[52] Because the postmodernist pretends not to be accountable to any normative notion of rationality, the very act of criticizing his 'position' misses the point, and thus postmodernism is rendered completely immune to criticism. Shackel has meticulously demonstrated that, while being obscured by the insinuation of the "No-Position Position", self-refutation is inevitable in postmodernist discourse.

The postmodernist rejection of reason is an extreme example of stonewalling, but one can find other exponents of this argument, which are not so sweeping as to entail self-refutation. For example, in discussions about alternative medicine one often hears the claim that each person or patient is "radically unique", thus frustrating any form of systematic knowledge about diseases and treatments. Of course, advocates of unproven medical treatments use this argument as a way to deflect the demand for randomized and double-blind trials to substantiate their therapeutic claims (Williams 1980; Gordon 1996). If each patient is radically unique, there is no point in lumping patients together in one treatment group and statistically comparing them with a control group. Homeopathy, for example, "considers the single patient as indivisible and unique [...] as not accessible to the method of measuring" (Guttentag 1940, p. 1177). Indeed, the whole idea of a classification systems of diseases is perceived by many advocates of alternative medicine as a form of greedy reductionism that eradicates the human subject. The argument is so convenient that it has been borrowed as an immunizing strategy by countless alternative therapists, including, inevitably, psychoanalysts (Boudry 2009a).

5.3.5 Invisible escape clauses

A last popular immunizing strategy of pseudoscientists – and in some cases taking the form of a full-blooded defense mechanism – consists in the systematic *ad hoc* invocation of invisible or imponderable causes that conveniently account for such a pattern of observations as would have been expected if the theory were false. As in the case of conceptual equivocations above, the availability of these escape clauses is initially obscured, and they come out of the closet only when the theory runs into trouble, giving the pseudoscientist's initial claims a spurious sense of empirical boldness. Again, we can distinguish two subtypes:

[52] The problem is also similar to the Mannheim paradox: if all discourse is ideological, how is it possible to have non-ideological discourse about ideology?

5.3.5.1 Tailoring around the phenomena

In the first subtype, the pseudoscientist invokes an invisible cause that is neatly tailored to an observational pattern of apparent failure, thus protecting the theory from refutation. An extreme example of this strategy is the so-called Omphalos hypothesis by Philip Gosse (1857), a variant of creationism according to which God forged all the geological evidence for an ancient universe to test our faith. In fact, this is a limiting case of a conspiracy theory, in which there is only 'inverted' evidence for the theory: all the observations point in the direction of an old universe, which is exactly what one would expect from a deceitful divine being intent on testing our faith.

One among many interesting examples of this immunizing strategy in parapsychology is the idea of negative psi emitted by skeptical minds and experimenters in general (Wiseman 2010), which is a popular excuse when psi experiments fail (see for example Sheldrake 1995). Some authors have given it impressive labels like "catapsi", which is defined as "the generation of 'static' that cancels out regular psi powers within its range" (Bonewitz 1989, p. 55). The idea that the presence of inquisitive minds disturbs paranormal phenomena already occurred to Franz Anton Mesmer and his fellow magnetizers, who believed that the skeptical presence weakened the force of the magnetic fluid. The instructions for magnetizers of Joseph P. F. Deleuze were clear enough: "Never magnetise before inquisitive persons!" (quoted in Mackay 1974 [1841], p. 290).

Parapsychologists have also invented the "error phenomenon" (Rao 1968), which refers to the finding that, when there is an error in the methodology or procedure of an experiment, this leads to better results, because these errors tend to activate psi (Humphrey 1996, p. 152). The famous psychical researcher John Beloff argued that psi phenomena are "actively evasive" (Beloff 1994, p. 7) and he has coined the "decline effect" (1994, p. 11) to describe the puzzling tendency of psychics to lose their powers as they are tested more extensively. Some parapsychologists have hypothesized that the primary function of psi is to "induce a sense of mystery and wonder", which allegedly explains its elusive character (Kennedy 2003, p. 67). Several other elaborate immunizing strategies have been devised for explaining why psi seems to actively avoid corroboration, some of which border on paranormal conspiracy theories (for an overview and discussion, see Kennedy 2001). Again, insofar as one takes these concepts to be an integral part of parapsychology, they are no longer immunizing strategies, but they have to be characterized as full-fledged defense mechanisms.

However one decides the question, all of these fanciful concepts and explanations have one thing in common: they seem designed so as to mimic the observations one would expect if the alleged psi phenomena were due to deception, trickery and methodological defects. They function as simple escape clauses for experimental failure, rendering psi theory immune from falsification.

5.3.5.2 Imponderabilia

The second subtype is related to the first, but with a different emphasis. Sometimes a pseudoscientist belatedly adds an extra factor to his theory that confounds the initial expectations it generated. A good example is the astrologer's belated invocation of the formation of stars at the moment of conception – which is of course very hard to determine – when his prediction on the basis of the birth date has failed. Another example is the amusing suggestion by believers in Bigfoot that the creature is possibly "extradimensional", so that any failed attempt to catch it can be explained by arguing that Bigfoot has escaped "into another dimension" (Zuefle 1999, p. 27; for the same trick with aliens, see for example Mack 1995).

As we will see in Chapter 10, Freudian psychoanalysis contains a host of escape clauses and methodological joker cards that make the theory eminently resilient to potential disconfirmations (Cioffi 1998; Esterson 1993). To give just one example, consider the way in which the already reviewed 'quantitative factor' in the patient's libidinal economy confounds empirical expectations initially engendered by the theory. Cioffi quotes several passages in which Freud leaves the reader with the impression that he has offered assessable hypotheses about the traumatic sexual events that predispose one to neurotic illness. On later occasions, however, Freud admits that some people who fall ill have experienced none of these events, after which he resorts to the imponderable quantitative factor that can only be inferred ex post facto. As Cioffi (1998, p. 119) writes, "our hopes that Freud might be placing a limit on the kinds of events or states which are conducive to the onset of neurosis and might then go on to tell us what these are, are dashed" when we read this kind of pseudoscientific elaborations of the theory.

5.4 Discussion

5.4.1 Remarks

In philosophy of science, some authors have emphasized that it is imperative not to confuse the theory-as-such with the immunizing tactics of its defenders (Grünbaum 1979, 2008). In regard to Freudian psychoanalysis, Adolf Grünbaum has insisted that the falsifiability of the theory-as-such be distinguished from the tenacious unwillingness of some psychoanalysts to face adverse evidence. Although we agree that, inasmuch as

possible, it is important to make the distinction Grünbaum insists on, in general it is not clear who is authoritative to decide where the theory-as-such ends and where the pseudoscientific immunizing strategies of its defenders begin (Cioffi 1998, p. 300). Consequently, there is often no objective way to distinguish immunizing strategies from internal defense mechanisms. For example, the immunizing strategies used by parapsychologists to account for apparent experimental failure in fact follow naturally from the intrinsically elusive nature of the alleged psi force. Who has the authority to decide whether this characterization of psi has nothing to do with proper parapsychology, or whether it is an integral part of parapsychological theory? When is the belated deflation of a theoretical claim really a revision of the original, and when is it just an elucidation based on equivocations that were always there? And what if immunizing gambits and conceptual joker cards emanate directly from the core conceptual structure of the theory, as in Freudian psychoanalysis (see section 10.3.2)? We will return to this problem in relation to the problem of *ad hoc* reasoning in Chapter 6, where we try to distinguish between core theory and auxiliaries.

Throughout this paper, we have used intentional language such as 'strategies', 'evasions' and 'maneuvers' for describing the ways in which pseudosciences and other belief systems are immunized against disconfirming evidence and rational criticism. In Chapter 11 (Boudry and Braeckman forthcoming), however, we argue that there need not be much conscious deliberation involved in such moves. The overall impression of strategic convenience we are left with when confronted with immunizing strategies and defense mechanisms may well derive from the latter's internal epistemic rationale, rather than from conscious deliberation and strategic planning on the part of believers.

5.4.2 Conclusions

In this paper, we reviewed several ways in which a belief system can achieve epistemic invulnerability against falsification and rational criticism: (1) the use of conceptual equivocations & moving targets, either through the technique of multiple endpoints or that of deflationary revisions; (2) the postdiction of invisible causes and unassessable effects; (3) the double evidential standard of conspiracy thinking, including the practice of explaining disbelief; (4) the practice of changing the rules of play in a rational debate and thus short-circuiting any form of criticism; (5) the invocation of invisible escape clauses, either by tailoring the theory around the phenomena or by invoking imponderable causal factors that confound expectations. As we noticed throughout our discussion, these techniques can be found across widely different domains: parapsychology, pseudo-philosophy, belief in magic, conspiracy theories, alternative medicine, religious cults, etc.

At the outset of this paper, we have distinguished immunizing strategies, which are brought forward by proponents from without a belief system, and epistemic defense mechanisms, which are structural parts of the belief system itself. However, in running through our classification, we have found that this distinction is sometimes difficult to maintain. First, an ad hoc elaboration that was introduced at some point to rescue a belief system from apparent falsification may gradually develop into an integral part of that belief system. In this way, the distinction between immunizing strategies and epistemic defense mechanisms is blurred. Second, although in some cases an evasive maneuver can be easily detached from the theory-in-itself, in other cases escape maneuvers were already implicit in the conceptual structure of the theory. Contra Grünbaum, there is not always a clear point at which the theory-as-such ends and the immunizing tactics of its defenders begin.

Our classification shows that proponents of belief systems that are either highly implausible or impossible from a scientific perspective, are in fact well-prepared to withstand the impact of empirical refutation and the force of critical argument. In Chapter 11, we further develop this account from a cognitive and psychological perspective, explaining why immunizing strategies and epistemic defense mechanisms are so pervasive, and why rational arguments are generally unavailing in debating believers (Boudry and Braeckman forthcoming).

Chapter 6.
The Hypothesis that Saves the Day. Ad hoc Reasoning in Pseudoscience

It is error only, and not truth, that shrinks from inquiry.
– Thomas Paine

Abstract.[53] What is objectionable about *ad hoc* hypotheses? Ever since Popper's falsificationist account of adhocness, there has been a lively philosophical discussion about what constitutes adhocness, and what, if anything, distinguishes legitimate auxiliary hypotheses from illicit *ad hoc* ones. This paper draws upon distinct examples from pseudoscience to provide us with a clearer view as to what is troubling about *ad hoc* hypotheses. Our approach retains the colloquial, pejorative connotation of adhocness, and calls attention to the way in which the context of a theoretical move bears on the charge of adhocness. We also discuss the role of motivations and aims implicit in the concept of adhocness, and the way *ad hoc* moves draw on theory-internal rationalizations.

[53] This paper is due to appear in *Logique et Analyse* (Boudry in press)

6.1 Introduction

In both academic and popular discussions on the scientific status of controversial theories, a hypothesis or explanation is often rejected as being *ad hoc*. In philosophical discussions about the demarcation project, the practice of resorting to *ad hoc* moves in the face of anomalous data is often regarded as a distinguishing feature of bad science or pseudoscience. However, traditional analyses of adhocness along Popperian lines have been confronted with several shortcomings, and more recent developments in philosophy of science have complicated the picture outlined by Popper and his followers. As with many concepts that are intuitively clear at a first glance, it has proven surprisingly difficult to unpack the notion of adhocness.

In this paper, we draw on distinct examples of bad reasoning from disciplines that are widely regarded as 'pseudoscience', in order to clarify what is objectionable about *ad hoc* moves. Rather than rehearsing the standard examples from the history of science (e.g. the postulation of an extra-Uranian planet, Pauli's neutrino hypothesis, the Lorentz-Fitzgerald contraction hypothesis), on which philosophical opinion is divided, our strategy is to focus on specimens of reasoning that are *blatantly* fishy in a sense that we intuitively recognize as *ad hoc*. Taking these clear-cut examples as a starting point, we may be better able to explicate what underlies our intuitions of 'adhocness', and we should be better capable of evaluating more complicated examples. We discuss the motivational and psychological component of *ad hoc* reasoning and rely on the concepts of immunizing strategies and epistemic defense mechanisms, as explored in Chapter 5.

6.2 Falsificationism and *ad hoc* reasoning

Karl Popper famously argued that the distinguishing feature of the scientific attitude is the willingness to make bold empirical conjectures and subject them to successive attempts at refutation. According to Popper, a theory can only be regarded as scientific if it forbids certain states of affairs, and the paragon example of a scientific theory is one that takes the boldest empirical risks. A hypothesis can be corroborated if it survives attempts at falsification, but when it runs against empirical observations, it has to be abandoned. However, sometimes an auxiliary assumption is incorporated in the theory in order to *rescue* it from falsification. According to the traditional Popperian view, this resort to *ad hoc* reasoning is illegitimate and even the hallmark of pseudoscience:

> Such a procedure [...] rescues the theory from refutation only at the price of destroying, or at least lowering, its scientific status. (Popper 2002, p. 48)

In accordance with the Quine-Duhem problem of underdetermination, however, philosophers of science after Popper have acknowledged that, in order to bring a hypothesis into contact with reality, one always needs a number of auxiliary hypotheses. In other words, hypotheses are always tested in conjunction and never in isolation. If a "bundle" of hypotheses is tested and the observations do not accord with what was predicted, from a logical point of view any one of the auxiliary hypotheses (or the core hypothesis) could be blamed. Indeed, when scientists devise a test for such a conjunction of hypotheses, what counts as the central hypothesis under test and what counts as background knowledge is a matter of methodological decision. Lakatos has attempted to correct Popperian falsificationism taking this problem into account:

> No theory forbids some state of affairs specifiable in advance; it is not that we propose a theory and Nature may shout NO. Rather, we propose a maze of theories, and Nature may shout INCONSISTENT. (Lakatos 1968, p. 162)

Falsificationists after Popper – and Popper himself in his more cautious moments – have allowed for modification of auxiliary hypotheses in the face of refutation, provided that the latter are independently testable and do not *reduce* the empirical content of the theory. If these conditions are not met, according to the falsificationist, the auxiliary hypothesis has to be discarded as *ad hoc*. Thus, a more sophisticated falsificationist philosophy of science accepts that every scientific research program builds up a "protective belt" of auxiliary hypotheses around its "hard core" claims (Lakatos and Musgrave 1970; Lakatos 1968). As such, adjustments and revisions in the face of empirical anomalies are not necessarily problematic. Scientists routinely resort to auxiliary hypotheses to rescue a theory from apparent refutation, and significant progress has been made by doing so. The example of Leverrier's and Adams's successful postulation of an extra-Uranian planet (Neptunus) to account for the perturbations in the orbit of Uranus provides a case in point.

However, the explication of adhocness in terms of reduced empirical content is still wanting. Although an auxiliary may not have *known* testable consequences at the time of its introduction, further developments and new experimental procedures may render it testable after all (Grünbaum 1976). Bamford (1993) argues that Popper equivocates between the colloquial sense of adhocness on the one hand and a technical definition on the other hand, calling attention to instances of genuine scientific progress in which the appeal to an '*ad hoc*' auxiliary did *not* increase the empirical content of the original theory. According to Bamford, the very idea of a hypothesis which has no testable consequences other than the observation it was introduced to account for, is difficult to make sense of in any case. Finally, even if it were possible for an auxiliary not to be

independently testable from the main hypothesis at all, this is not to say that such a modified theory would be necessarily false. At most, one may argue that adhocness is generally not *conducive* to scientific progress.

In light of these and other problems, some authors have abandoned the project of explicating adhocness in a pejorative sense, and have portrayed the concept as a neutral methodological or epistemic attribute. Grünbaum (1976) offered a purely descriptive hierarchy of three senses of adhocness, and Laudan (1977) has even suggested that the extent to which a theory allows for *ad hoc* moves is a *redeeming* feature of that theory. However, others have unpacked the notion of adhocness in a way that retains the colloquial, pejorative meaning. For example, Leplin (1975) has reserved the term for auxiliary hypotheses the introduction of which results in a loss of theoretical simplicity that is *not* off-set by a proportionate gain of fit with the data (see also Kitcher 1982).

6.3 Adhocness in pseudoscience

Although Popper's demarcation criterion has been decidedly out of philosophical fashion for several decades, the charge of adhocness as a strategy for evading falsification, and thus as the very antithesis of empirical boldness, is still often leveled against controversial theories. As we saw in Chapter 5, an increasing dependence on *ad hoc* moves is widely regarded as a telltale sign of pseudoscientific discourse (e.g. Derksen 1993; Hines 2003; Carroll 2003; Pigliucci 2010; Kitcher 1982). Unless these critics of bad science are entirely misguided, there must be *some* virtue to empirical boldness and non-adhocness that requires further explication, even if Popper's demarcationism has been overly simple.

The problem is that many major episodes in the history of science in which the charge of adhocness played a role prove to be ambiguous. In real-life scientific disputes, it is not always clear when exactly a theory starts to require too much gerrymandering to accommodate anomalies. Examples include phlogiston theory, phrenology, Lamarckist evolution, or the steady state model of the universe. In many of these cases, there is no clear point at which the theory collapses under the weight of empirical anomalies, and when it ceases to be rational to defend it. Often enough, crucial experiments and fatal anomalies are only to be identified in retrospect, when the theory has been entirely abandoned. This shows that the concept of adhocness is somewhat complicated and may be a matter of gradients, but not that it is fundamentally incoherent.

Indeed, if we accept that non-adhocness is an epistemic liability, and given that we are dealing with serious scientific disputes, we can reasonably *expect* that charges of adhocness are somewhat ambiguous and open to discussion. If advocates of a theory X are confronted with empirical anomalies, they will reasonably try to save X in a way that is least open to the charge of adhocness, at least as long as this is possible. When the evidence against a theory is so overwhelming that it would require blatant *ad hoc* moves to save it, that theory is unlikely to be the subject of serious scientific debate for a long time, and will either disappear or persist merely on the fringes of science. Notorious pseudosciences are only rarely the subject of philosophical discussions about adhocness, but this is unfortunate. By having a look at blatant examples of *ad hoc* reasoning from the hinterland of pseudoscience, we may be better able to make sense of charges of adhocness in more complicated cases.

6.3.1 Escape Clauses

In section 5.3.5, we discussed a number of immunizing strategies and epistemic defense mechanisms for protecting parapsychology against adverse evidence. These theoretical moves can be reconstructed as *ad hoc* auxiliaries to the central hypothesis of psi forces. For example, some parapsychologists have proposed a negative form of psi (catapsi) emitted by inquisitive observers that cancels out regular psi activity (Wiseman 2010; Humphrey 1996). In the same vein, parapsychologists have explained away evidence by relying on the "error phenomenon" (Rao 1968) and "decline effect" (1994, p. 11), or by claiming that psi is "actively evasive" (Beloff 1994, p. 7) (for further details, see section 5.3.5).

The most conspicuous feature shared by these theoretical moves is that they are conveniently tailored around some empirical anomalies, while at the same time being too vague and non-specific to allow for novel predictions, and failing to provide any form of explanatory unification (see section 4.4). With this explanatory resource at their disposal, parapsychologists are in a position to explain away not just one particular anomaly, but potentially *any number of negative experimental results*. Hence, the error phenomenon and catapsi effect account for precisely what would have been expected if the alleged psi phenomena were entirely due to deception, sloppy experimental design and methodological defects.

What do we mean when we reject such moves as *ad hoc*? What is problematic is not just the logical relation between main hypothesis and auxiliary, but rather the circumstances in which the latter is introduced, and its (potential) range of application. To further illustrate this point, consider again the case of an astrologer who belatedly invokes the formation of stars at the moment of *conception* when his prediction on the basis of the birth date has failed. Or similarly, consider the defender of biorhythm

theory who resorts to the hypothesis that some people are "arrhythmic" some of the time when his predictions do not fit the observed patterns (Carroll 2003, p. 7). Why are we entitled to reject these moves of *ad hoc*? Because we realize that, by the same token, *any* type of observation can be handled, and hence that there are *no constraints* on the use of the (type of) auxiliary hypothesis in question. In Freudian psychoanalysis, the so-called "hereditary factor" was often invoked when traces of allegedly repressed infantile desires were unforthcoming. This meant that, if the patient had not personally experienced the sexual traumas required by the theory, Freud argued that traces of "phylogenetic memory" – recollections of ancient traumas inherited over the generations – could fulfill the same role (Cioffi 1998, p. 108).

The other side of the coin, as far as the 'no constraints' objection is concerned, is that the concept on which the *ad hoc* move relies is indeterminate enough to be conveniently *ignored* as long as the data align with the theory. Consider the same astrologer who, as long as his birth date-based predictions do not run against the facts, does not bring up the moment of conception as an allegedly crucial factor, and boasts that his predictions have been successful. Or consider Freud's neglect of phylogenetic inheritance when he wanted to explain a patient's symptoms in terms of the *absence* of a powerful father threatening with castration. Whereas such concepts are invoked on particular occasions to explain away predictive failure, they are inconsequentially ignored on other occasions. Thus, there is an *asymmetry* in the use of the auxiliary hypothesis in question.

As is clear from these examples, the broader context of a theoretical move is fundamental with regard to its being *ad hoc*. Although the presumption of *ad hoc* reasoning is strongest when we have actually witnessed several instances of opportunistic and inconsequential use of an auxiliary hypothesis, suspicion can also be warranted on the basis of a single case, provided that the lack of constraints is immediately obvious, and that it is clear that the hypothesis can yield nothing beyond explaining away particular failures.

6.3.2 Patterns of systematic *ad hoc* reasoning

When we dismiss a theoretical move as being *ad hoc,* we seem to be saying that its introduction is only motivated by the particular anomaly encountered, and that there are no further theoretical reasons at play. In an important sense, however, this assumption has to be corrected. In some cases, *ad hoc* reasoning is not just based on spurious extensions or modifications of a theory, but a systematic pattern of *ad hoc* reasoning emerges from the very epistemic structure of a theory. In Chapter 5, we

distinguished immunizing strategies and epistemic defense mechanisms as two ways in which a theory can achieve invulnerability from criticism and empirical refutation. [54] In the case of epistemic defense mechanisms, the theory itself proffers ample opportunity for rationalizations and evasive responses in the face of counterevidence. One of the ways in which this may happen is if structural features of a theory invite a systematic pattern of *ad hoc* reasoning.

Consider the phenomenon of postdiction and feedback loops we discussed in section 5.3.2: an unobservable force or cause is posited to account for a range of observations, while believers have no precise understanding of the causal mechanisms involved, and at the same time go about inferring the activity of these invisible forces *ex post facto* from their effects. In shamanistic healing practices, the success of magic rituals depends on a number of elaborate procedures, taboos and prescripts, the violation of which may interfere with the workings of the ritual. As the proper performance of the ritual is partly inferred from the therapeutic outcome itself, believers are invited to interpret unexpected outcomes not as a failure of the ritual per se, but as an apparent sign that 'something must have gone wrong': the intervention was not of the appropriate type, a taboo has been breached, the ritual was not performed properly, evil magic has interfered with the ritual. Again, similar moves may be made by a scientist making an unexpected observation under a microscope, but in such a case there are indirect and independent checks available.

In his seminal study on witchcraft and magic among the Azande, Evans-Pritchard has termed these *ad hoc* moves "secondary elaborations" (Evans-Pritchard 1965), which have the effect of making the belief systems internally coherent and impervious to refutation (a second class of such secondary elaborations is available when the alleged effect is itself extra-empirical or otherwise difficult to observe, see section 5.3.2).

In the field of parapsychology, a poor causal understanding of alleged paranormal forces likewise invites a systematic *pattern* of *ad hoc* reasoning in the face of alternate successes and failures. Feedback loops between the interpretation of observations and the theoretical characterization of *psi* will almost inevitably arise. Psi-activity may be postdicted whenever the results are successful, and because of the underspecified causal mechanisms of psi, virtually any contextual factor may be invoked as an *ad hoc* explanation for failure (Wiseman 2010). Note that, according to parapsychologists, *psi* powers may disappear over time or work only intermittently, and that this can only be

[54] The use of intentional language such as 'strategies', 'evasions' and 'maneuvers' for describing the ways in which pseudoscience and other belief systems are immunized against disconfirming evidence and rational criticism is not always to be taken literally. In Chapter 11, we will argue that the overall impression of *strategic convenience* we are left with when confronted with immunizing strategies and defense mechanisms may well derive from the latter's internal *epistemic rationale*, rather than from conscious deliberation and strategic planning on the part of believers.

inferred from experience. What is objectionable about parapsychology and other magical belief systems is not so much the introduction of a single auxiliary hypothesis, but rather a systematic *practice* of *ad hoc* reasoning, which is part and parcel of the pseudoscientific theory itself. In other words, we find that there are simply *no constraints* on the use of such evasive auxiliaries.

6.4 Discussion

6.4.1 Unobservability

As many of the examples of *ad hoc* reasoning we have reviewed rely on unobservable entities and forces, one may wonder whether unobservability has any bearing on the accusation of adhocness. Leplin (1975) has argued, against the received view of *ad hoc* reasoning, that unobservability in principle mitigates the condition of being de facto unobservable. Indeed, the failure to detect X is less damaging if we possess of good, independent reasons to presume that X must be intrinsically unobservable. By contrast, if there is no good theoretical reason for X's unobservability, then we have no excuse for *failing* to detect X. For example, astronomers have postulated the existence of an invisible planet, Vulcan, or an invisible asteroid belt, to account for anomalies in the orbit of Mercury. The main reason why these auxiliaries to the Newtonian model of the solar system were quickly discarded, however, was that no satisfactory explanation was given for why these objects should be invisible to observers on earth. In other cases, such as the postulation of quarks and other subatomic particles, auxiliaries can derive support from other considerations: accordance with well-established theories, indirect experimental support, unification of a range of data etc. In the pseudoscientific examples we reviewed, however, the activity of X in question is merely postdicted on the very phenomena which X was supposed to account for, without any offset in the form of other explanatory virtues. These considerations justify the labeling of the invocation of X as *ad hoc*, not so much the alleged unobservability of X.

6.4.2 Adhocness and psychological motivation

An important problem confronting the reconstruction of adhocness in a pejorative sense is the distinction between the methodological and psychological components of

the concept, both of which are often conflated in traditional discussions of *ad hoc* reasoning (Bamford 1993; Bamford 1999). As Laudan (1977, p. 117) writes, underlying many discussions of adhocness is the conviction that "there is something suspicious about any change in a theory which is motivated by the desire to remove some anomaly." Is our discussion of *ad hoc* reasoning in pseudoscience guilty of this confusion as well?

In the received view, an auxiliary is typically dismissed as *ad hoc* because its introduction is said to be *motivated to* explain away some bothersome anomaly. This is misleading, because a theoretical solution is not invalid simply because it happens to fulfill the desire to rescue a theory. Even deliberately searching for a suitable auxiliary assumption to that end is not a bad thing *per se*. For example, even if a scientist introducing auxiliary X *intended* to explain only a particular anomaly, this does not preclude that X has other (perhaps unintended) testable consequences (Bamford 1999, pp. 379-380). A scientist may just be lucky that his personal desire (to prove a theory right) is borne out by nature. That being said, few would deny that a strong desire to protect a theory may be a good *proxy* for explanatory problems with the auxiliary hypothesis one comes up with in the face of anomalies. If A is desperate to save one particular theory at all costs, and B has no stakes in the debate, it is more likely, other things being equal, that the resolution A comes up with is more biased, compared to B's take on the issue. Thus, it is not unreasonable to suspect A's solution as having been motivated *solely* by the desire to persevere in his cherished ideas.

The same problem pertains to any account of motivations with regard to theory choice. For instance, the very fact that A is motivated by his religious views to favor a hypothesis does not immediately invalidate A's choice, since strong religious motivations are compatible with sound reasoning. However, if A favors a theory that fits well with A's religious views, at least the suspicion may be raised that A's religious zeal has trumped epistemic considerations. More importantly, in absence of good epistemic reasons, A's religious views constitute the best candidate for explaining why A has endorsed such or such view. Thus, if understood properly, there is nothing wrong with dismissing a theory as "ideologically motivated", or with objecting to an *ad hoc* move because it is "simply motivated by the desire to save the theory". The latter should be viewed as an elliptic expression, which means nothing more than that, in the *absence of proper epistemic reasons*, the only plausible motive is the desire to rescue the theory at all costs.

6.4.3 Core theory and *ad hoc* auxiliary hypotheses

In most philosophical accounts, an auxiliary is labeled as *ad hoc* in relation to some original theory or core hypothesis. In philosophy of science, however, some authors

have emphasized that it is imperative not to conflate *ad hoc* maneuvers for rescuing a theory with the theory-as-such. Notably, in regard to Popper's verdict on Freudian psychoanalysis as an unfalsifiable pseudoscience, Adolf Grünbaum commented:

> The (revocable) falsifiability of the theory-as-such in the context of its semantic anchorage is a logical property of the theory itself, whereas the tenacious unwillingness of the majority of its defenders to accept adverse evidence as refuting is an all too human property of those advocates. (Grünbaum 1979, pp. 137-138)

According to Grünbaum, the latter point merely comes down to the "*sociological* objection that Freudians are evasively unresponsive to criticism of their hypotheses" (Grünbaum 2008, emphasis in original). Although we agree that, inasmuch as possible, it is important to make the distinction Grünbaum insists on, often it proves difficult to separate the theory-as-such from the *ad hoc* reasoning and immunizations used by its defenders (Boudry and Braeckman 2010; Cioffi 1985, 1998). For example, the pattern of *ad hoc* reasoning found in parapsychology has been given a theory-internal explanation, viz. the elusive nature of psi forces, its wonder-inducing function, its intermittent efficacy, etc. For many parapsychologists, this feature of psi is one of the central findings of their research program (Kennedy 2003, 2001). Whom do we have to consult to find out whether the resort to the shyness of psi is an *ad hoc* maneuver which has nothing to do with proper parapsychology, or whether it is an intrinsic part of parapsychological theory?

Grünbaum's own example of Freudian psychoanalysis is particularly instructive in this regard. In Chapter 10, we will see that the practice of *ad hoc* reasoning on the part of Freudian psychoanalysts emerges naturally from (i) the flexible and indeterminate conceptual structure of the theory (ii) Freud's dynamic system of repressions, inversions and projections between antagonistic mental subsystems, in which anything can stand for either itself or its counterpart; (iii) the countless theoretical joker cards and immunizing gambits in psychoanalytic theory, which are almost impossible to resist in interpretive practice (Macmillan 1997; Cioffi 1998; Boudry and Braeckman 2010). If Freudian theory *itself* is the source of systematic *ad hoc* reasoning, we submit that there is something more going on than the mere "tenacious unwillingness of the majority of [psychoanalysts]" to accept disconfirming evidence. Although reasons of space prevent us from providing a detailed discussion of this problem, we hope that these examples have at least challenged the idea that the purported distinction between the theory-as-such and its advocates' lapses into *ad hoc* reasoning is not a straightforward matter.

6.5 Conclusion

What, if anything, distinguishes *ad hoc* reasoning from the legitimate introduction of auxiliary hypotheses in science? In order to get a clearer view on this matter, we have tried to find species of reasoning that are blatantly '*ad hoc*' in the colloquial and pejorative sense, borrowing examples from disciplines that are widely regarded as pseudoscience.

First, we have reviewed the opportunistic and inconsequential resort to invisible escape clauses that are conveniently tailored around observational anomalies. Second, we have shown that the structure of some pseudosciences is such as to invite a systematic pattern of *ad hoc* reasoning, by making use of secondary elaborations regarding ambiguous causal mechanisms and unassessable effects.

As these examples from pseudoscience make clear, an explanation or auxiliary hypothesis deserves to be labeled *ad hoc* if it *merely* explains away particular anomalies without yielding any form of theoretical progress or empirical unification. The concept of adhocness, in the pejorative sense we have construed here, is sensitive to the particular context in which a theoretical move is made. An accusation of adhocness is more damning if it can be shown that there are no demonstrable constraints on application (i.e. the move can be used to explain away any bothersome anomaly) and if it is being used inconsequentially (i.e. it is conveniently ignored on other occasions). Precisely because the introduction of an *ad hoc* auxiliary serves no proper epistemic or explanatory goals, the move is often condemned as *psychologically motivated* to save the theory at all costs. This way of framing the complaint of adhocness is strictly speaking inaccurate, but it is understandable as a shorthand. If a theorist is strongly motivated to rescue a theory at all costs, this is a good reason to be suspicious, and if no proper epistemic reasons for his move are apparent, we may safely conclude that we are dealing with a case of wishful thinking.

If we put this motivational component of adhocness into proper perspective, and if we move beyond a naïve falsificationist view about unobservability and testability, the concept of non-adhocness may be retained as capturing an important liability for theorizing and explanation.

Chapter 7.
Irreducible Incoherence and Intelligent Design: A Look into the Conceptual Toolbox of a Pseudoscience

Si la nature nous offre un nœud difficile à délier laissons le pour ce qu'il est et n'employons pas à le couper la main d'un être qui devient ensuite pour nous un nouveau nœud plus indissoluble que le premier. - Denis Diderot

Abstract.[55] The concept of Irreducible Complexity (IC) has played a pivotal role in the resurgence of the creationist movement over the past two decades. Evolutionary biologists and philosophers have unambiguously rejected the purported demonstration of "intelligent design" in nature, but there have been several, apparently contradictory, lines of criticism. We argue that this is in fact due to Michael Behe's own incoherent definition and use of IC. This paper offers an analysis of several equivocations inherent in the concept of Irreducible Complexity and discusses the way in which advocates of the Intelligent Design Creationism (IDC) have conveniently turned IC into a moving target. An analysis of these rhetorical strategies helps us to understand why IC has gained such prominence in the IDC movement, and why, despite its complete lack of scientific merits, it has even convinced some knowledgeable persons of the impending demise of evolutionary theory.

[55] This chapter has previously been published in *Quarterly Review of Biology* (Boudry, Blancke et al. 2010b).

7.1 Introduction

Until its dramatic legal defeat in the Kitzmiller v. Dover case, Intelligent Design Creationism (IDC) had been one of the most successful pseudosciences of the past two decades, at least when measured in terms of cultural influence. It is interesting to explore the way this species of creationism had achieved this success, notwithstanding its periodic strategic setbacks as well as its complete lack of scientific merits. Of course, a full explanation would include religious, socio-political, and historical reasons; instead, in this paper, we will take a closer look at the conceptual toolbox and rhetorical strategies of the ID creationist. As a case study, we will concentrate on the central concept of "irreducible complexity" (IC), but other examples can be found that support our assertions (see Chapter 8). Our analysis shows that the conceptual equivocations inherent in the concept of IC — in particular its potential to be employed as a moving target in discussions (Boudry and Braeckman 2010) — may help further our understanding of the superficial appeal of the design argument based on IC, which is only the most recent in a long series of creationist challenges mounted against evolutionary theory.

7.2 Irreducible complexity

In 1996, biochemist and IDC proponent Michael Behe introduced the infamous concept of irreducible complexity (IC) in his book *Darwin's Black Box.* Although Behe's critics univocally agreed that he failed to demonstrate evidence of "intelligent design" in nature, there have been several, seemingly inconsistent lines of criticism. Some evolutionary scientists claim that biological systems do sometimes exhibit IC as Behe defines it, but they deny that this poses a problem for evolutionary theory (e.g. Orr 1997; Shanks and Joplin 1999; Miller 2000), while others maintain that Behe has never demonstrated the existence of bona fide instances of IC in nature (e.g. Pigliucci 2002; Forrest, Barbara C. and Gross, Paul R. 2007). Pennock (1999, pp. 264-272)concurs with this criticism but grants the possible existence of biological IC systems, arguing that, in any case, these would not threaten evolutionary theory.

We propose that this seemingly contradictory criticism is in fact due to Behe's own disjointed definition and misleading use of IC. First, we introduce Behe's concept and briefly recount the most important empirical objections against it. Then, we analyze the

conceptual equivocations inherent in Behe's approach on several levels (see also Dunkelberg 2003). Finally, we argue that these kinds of equivocations allow Behe and his IDC fellows to make a moving target out of their theory, hence serving to insulate it against criticism.

7.2.1 The evolution of irreducible complexity

The concept of IC has an interesting "evolutionary" pedigree (Forrest, B. C. and Gross, P. R. 2007, p. 302). In the 1970s and 1980s, young-Earth creationists used similar terms to describe biological systems that were alleged obstacles to evolutionary theory. In 1974, Henry Morris, founder of the Institute for Creation Research and father of the Creation-Science movement, argued in his influential book Scientific Creationism that "The problem is simply whether a complex system, in which many components function unitedly together, and in which each component is uniquely necessary to the efficient functioning of the whole, could ever arise by random processes" (Morris 1974, p. 59). In 1980, young-earth creationist Ariel Roth argued that "Creation and various other views can be supported by the scientific data that reveal that the spontaneous origin of the complex integrated biochemical systems of even the simplest organisms is, at best, a most improbable event" (Roth 1980, p. 83). Behe has simply adapted these creationist notions to his own ends. Consider his definition of IC in Darwin's Black Box:

> By *irreducibly complex* I mean a single system composed of several well-matched, interacting parts that contribute to the basic function, wherein the removal of any one of the parts causes the system to effectively cease functioning. An irreducibly complex system cannot be produced directly (that is, by continuously improving the initial function, which continues to work by the same mechanism) by slight, successive modifications of a precursor system, because any precursor to an irreducibly complex system that is missing a part is by definition non-functional. An irreducibly complex biological system, if there is such a thing, would be a powerful challenge to Darwinian evolution. (Behe 2006, p. 39)

7.2.2 Redundant and irreducible complexity

Behe then proceeds to argue that many biological systems exhibit IC, especially at the subcellular level (e.g., the bacterial flagellum). However, many critics have noted that the components of a typical biological system manifest considerable functional overlaps and redundancy. Contrary to Behe's assertions, living systems are often quite robust to perturbations, despite, or even because of, their complexity (Ciliberti, Martin et al.

2007). Overall these system exhibit what has been termed "redundant complexity" (Shanks and Joplin 1999). For example, if we eliminate one or even several elements from the blood clotting cascade, which Behe cites as an instance of an IC system, the system still manages to perform its function, albeit not as swiftly or efficiently as before. From the perspective of evolutionary theory, this is hardly surprising; natural selection is a clumsy and opportunistic process that tinkers with the available material. Thus, the widespread phenomenon of redundant complexity makes it perfectly clear that evolution by natural selection can gradually produce increasingly complex systems without the guidance of an intelligent designer (see below).

7.3 Equivocations

7.3.1 Conceptual double life

To be sure, it is not difficult to find examples of biochemical systems in which the removal of just one part damages the whole system. But consider Behe's phrases "effectively ceases functioning" and "by definition non-functional." There are two possible reconstructions of his definition: 1) the term "functioning" refers exclusively to the basic function currently performed by the whole system (e.g., the rotary motion of the bacterial flagellum) and does not pertain to other possible functions, in other contexts, when one or more components are removed; and 2) the phrases "effectively ceases functioning" and "non-functional" include any function that the impaired system or one of its components may perform in other contexts. In principle, it is not very hard to discover whether a system exhibits IC in the first, weak sense. Leaving aside the ambiguity regarding the natural "parts" into which the system must be decomposed (Dunkelberg 2003; Sober 2008, pp. 135-160), it suffices to knock out these parts one after the other to see if the system can still perform its basic function. Again, evolution by natural selection is perfectly capable of producing complex functional systems exhibiting IC in this weak sense. For example, Lenski et al. (2003) used a population of "digital organisms" (i.e., computer programs) to simulate the evolution of a complex functional system. By performing a series of knockout experiments on one of the complex functions that emerged from their simulation, Lenski et al. were able to determine how many genomic "instructions" were involved in its functioning. The researchers found that the function "depends on many interacting components" (2003:141), the removal of any of which causes the system to break down.

In fact, only an IC system in the second, strong sense would be an obstacle to evolutionary theory, because it would rule out evolutionary precursor systems and function shifts of the system's components. However, it is hard to see how Behe could even begin to demonstrate the existence of such a system without defaulting to the classical "argument from ignorance" (Pigliucci 2002, p. 67). Interestingly, Behe has disingenuously taken advantage of this very ambiguity in answering his critics.

In his initial definition, Behe seems to intend the weak interpretation, but he then proceeds to use the concept in a line of reasoning that only makes sense under the strong interpretation. Precisely because the bacterial flagellum is IC, Behe tells us, it could not have evolved by means of random mutation and natural selection. However, when critics object that the system's components may well be able to perform other functions in other contexts, thus pointing to the possibility of indirect evolutionary pathways, Behe switches back to the weak definition and claims that his critics have misrepresented his argument.

7.3.2 A conceptual mousetrap

Robert Pennock (1999, p. 267) objected to Behe's design argument that "even if a system is irreducibly complex with respect to one defined basic function, this in no way implies that nearby variations might not serve other nearby functions". Reasonably, Pennock construes Behe's argument in a sense that is intended to preclude any functional intermediate on a direct or indirect evolutionary path to the current system:

> Behe claims that there could never be any functional intermediates that natural selection could have selected for on the way to *any* irreducibly complex system, but he can't get the empirical conclusion from his "by definition" conceptual argument (Pennock 1999, pp. 267-268)

Pennock's reasoning is correct, of course, but in the afterword to the tenth anniversary edition of Darwin's Black Box, Behe (2006, p. 258) retorts that "Pennock [simply] substituted his own concept of irreducible complexity for mine," whereupon he shifts back to the weak version of the concept, which merely rules out direct improvements on the system: "On the contrary, on page 40, I point out that, although irreducible complexity does rule out direct routes, it does not automatically rule out indirect ones" (see also Ratzsch 2005). Thus, Behe protests that Pennock has "overlooked important qualifications" (Behe 2001, p. 707) and has simply "constructed his own rigid straw man definition for IC." But Behe himself has boldly stated that any IC system is a "powerful challenge to Darwinian evolution" (2006, p. 39), and that "[w]e know of no other mechanism, including Darwin's, which produces such complexity" (1996, p. 25). Thus, the fact that Behe's own qualifications are inconsistent with his boastful presentation of

IC as a major stumbling block for evolution is hardly Pennock's problem. Behe did acknowledge that Pennock exposed another weakness in the definition of IC, owing to its focus on already functioning systems rather than on the evolutionary development of such systems. Although he promised to "repair this defect in future work" (Behe 2001, p. 695), so far Behe has not lived up to that promise, instead seeming to ignore the problem altogether.

The neglect of evolutionary development in Behe's definition is hardly a trivial matter, however, and his concession concerning indirect routes is quite an important one, which seems to be completely absent from his original definition (see also Sober 2008, pp. 161-162). As early as the beginning of the 20th century, geneticist Herman Muller explained how biological systems that depend on the complex "interlocking" actions of many different components could come about by evolutionary processes: "Many of the characters and factors which, when new, were originally merely an asset finally became necessary because other necessary characters and factors had subsequently become changed so as to be dependent on the former" (Muller 1918, pp. 463-464). Thus, redundant complexity can eventually generate IC (under the weak interpretation). More recently, biochemist and molecular biologist A. G. Cairns-Smith proposed the analogy of "scaffolding" in the construction of an arch to explain the evolution of systems that are IC according to Behe (Cairns-Smith 1986; see also Orr 1997; Pennock 2000). A classical stone arch is IC in the weak sense, because the structure will collapse as soon as one removes either the keystone or one of the other stones. The support of scaffolding is necessary in building a stone arch, but once the arch is completed, the scaffolding can be safely removed. In a similar vein, a biochemical structure may have functioned as a scaffold in the evolution of an IC system before becoming dispensable and disappearing. That is, "Before the multitudinous components of present biochemistry could come to lean together they had to lean on something else" (Cairns-Smith 1986, p. 61).

Behe has performed a similar conceptual sleight of hand in dealing with the objections of molecular biologist Kenneth Miller (2000). Miller accepts that some biological systems are IC as Behe defines it (weak version), but he objects to the anti-evolutionist conclusions that Behe derives from IC. As a counterexample of Behe's claim, Miller offers a plausible reconstruction of the evolutionary history of the five-part auditory apparatus in mammals, which he argues fulfils the definition of IC. Miller demonstrates that the individual parts of the auditory apparatus—mallens, incus, and stapes—evolved from the rear portion of the reptilian jaw. It is important to note that before they migrated to the middle ear and were adapted for their new purposes, these structures were indeed perfectly functional. Therefore, Miller concludes that Behe's statement (2006, p. 39) that "any precursor to an irreducibly complex system that is missing a part is by definition non-functional" is plainly wrong. Miller challenges strong IC and demonstrates the crucial point, which is that the "interlocking necessity [of the

parts of the final working system] does not mean that the system could not have evolved from a simpler version" (2000, p. 139).

Behe, however, has responded by asserting that Miller "concocted his own, private definition of irreducible complexity, and then argued against that" (2006, p. 259). It is quite possible, he goes on to explain, that individual components of an IC system can perform functions in different contexts. Thus, according to Behe (2006, p. 260), Miller has "redefined irreducible complexity to mean that none of the component parts of an IC system could have its own function separate from the system".

Yet again, the equivocation is in Behe's definition, not in Miller's criticism. Bearing in mind that Behe treats IC as if it were an insurmountable obstacle for evolution, which is already clear from the very wording of the term "irreducible," the critic naturally confronts Behe's claim of "non-functionality" by pointing to the different functions performed by evolutionary precursors of IC systems, which may or may not have contained parts of the current system.

After all, if we bear in mind that biological systems can be adapted over the course of evolution for another function than that for which they were originally selected—for instance, by being integrated as part of a new system performing a different function—then Behe's non-functionality claim becomes either trivial (weak version) or plainly wrong (strong version).

7.3.3 Dembski's conceptual remedy

In *No Free Lunch* (2002), Behe's creationist ally William Dembski proposed to remedy the conceptual problems of IC. Dembski believes that the concept of IC is "salvageable" (2002, p. 280), and after a series of modifications, he arrives at the following new definition:

> Definition IC_{final} – A system performing a given basic function is *irreducibly complex* if it includes a set of well-matched, mutually interacting parts such that each part in the set is indispensable to maintaining the system's *basic, and therefore original, function*. The set of these indispensable parts is known as the *irreducible core* of the system. (Dembski 2002, p. 285, emphasis in original)

Accordingly, Dembski argues, the IC of a system is a straightforward empirical question:

> Individually knocking out each protein constituting a biochemical system will determine whether function is lost. If it is, we are dealing with an irreducibly complex system (Dembski 1999, p. 148).

Clearly, Dembski has "fine-tuned" the concept of IC in the direction of the weak interpretation, restricting the definition to the basic, original function of the system.

His updated version has the merit of conceptual clarity (but see Perakh 2002), but, in remedying Behe's conceptual ambiguity, Dembski actually takes the sting out of the whole argument. IC thus conceived is perfectly consistent with indirect and circuitous routes, scaffolding, and exaptations. So what is all the fuss about? The collapse of IC in Dembski's hands illustrates that the conceptual ambiguity he was trying to salvage was actually very *convenient* for Behe.

7.4 Moving the goalposts

7.4.1 Never enough

Despite Dembski's remedy, other equivocations in the concept of IC have yet to be resolved. Having failed to provide an objective criterion that makes evolutionary accounts impossible, the IDC proponent retreats to a weaker probabilistic claim; as the number of individual components in an IC system increases, the plausibility of a gradual succession of slight modifications becomes vanishingly small. "The strength of the inference depends on the number of parts, and the more intricate and sophisticated the function, the stronger is our conclusion of design" (Behe 2006, p. 265). Leaving aside the problems regarding this alleged correlation between the numbers of parts and the strength of the design inference, which are amply documented by Pennock (1999, p. 270), we still *seem* to be left with a testable statement. If we can find a well-functioning precursor for one of the systems discussed by Behe (or for one of its components), or if we can construct a plausible evolutionary pathway for one of Behe's examples, the "probability" argument collapses.

Behe's claim has indeed been tested against the facts and has been found wanting (Miller 2000; Lenski, Ofria et al. 2003; Young and Edis 2006; Forrest, B. C. and Gross, P. R. 2007). In response to these demonstrations, however, IDC proponents belatedly "reinterpret" their initial claims in order to lift them out of the critic's reach. A first strategy to this end consists of shifting the burden of proof from plausible evolutionary pathways to the actual evolutionary story, and thus to protest that the broad outlines of a plausible evolutionary account amount to nothing more than Darwinian wishful thinking and speculation. The same bait-and-switch technique can be discerned here: IC is constantly boasted as a point of principle for ruling out the possibility of evolutionary explanations, but as soon as it is challenged on that ground through a discussion of

plausible evolutionary scenarios, ID creationists contend that they were talking about actual evolutionary pathways all along.

When they are confronted with tangible evidence of actual evolutionary history, IDC theorists resort to a second strategy, shifting their design claims to the remaining parts of the evolutionary puzzle, as if the "real" problem was always there. For example, Kenneth Miller (2004) beautifully demonstrated the structural similarities between one component of the flagellum and the so-called type III-secretory system. He convincingly argued that the former is a very plausible evolutionary precursor of the latter, which has been co-opted by evolution to perform a new function (see also Pallen and Matzke 2006). In response to this embarrassing demonstration, Behe (2001, pp. 689-690) simply shifted his attention to the complexity of the newly discovered system by itself, while at the same time stubbornly insisting that the assemblage of these precursors into the flagellum system is still impossible without the helping hand of a Designer (Behe 2004, p. 359).

In light of these evasions, one may wonder whether there is any amount of comparative genetic evidence, or any level of evolutionary reconstruction, that would make Behe and his allies abandon their design claims. Because of the sloppiness of the probabilistic IC claim, which is not based on any serious quantification of probabilities, IDC theorists can continue to raise the evidential bar up to a point where the concept of IC is lifted outside of the empirical domain altogether. Indeed, when pressed on the available scientific knowledge of a particular complex system that he cites, Behe has made it clear that only a complete, quantitative, and fully-detailed description of what *actually* happened over the course of the ages would convince him of its evolutionary origin (Behe 2007). In his testimony at the Dover trial, Behe conceded:

> Not only would I need a step-by-step, mutation by mutation analysis, I would also want to see relevant information such as what is the population size of the organism in which these mutations are occurring, what is the selective value for the mutation, are there any detrimental effects of the mutation, and many other such questions. (2005, p. 19)

But this is an absurd demand, which is never met in any other scientific domain, and is certainly not met by ID creationists themselves when they propose "design" as an alternative explanation. Indeed, despite his demand for such a high level of evidence for the evolution of what he claims are IC systems, Behe himself has been completely unwilling to flesh out his design hypothesis to any degree at all, insisting that the motives and character of the designer are in fact inscrutable, and he provides us with no clue as to his *modus operandi*. As for Behe's request for fully detailed knowledge about evolutionary history, Pigliucci (2002, p. 240) has warned biologists not to be overconfident in taking up creationist challenges, and not to mistake partial reconstructions and plausible scenarios for a complete understanding of evolutionary

development. Indeed, evolutionary theorists are better advised to explain why the burden of proof insisted on by creationists is absurd, and to point out that scientific knowledge will never be complete in this respect.

In any case, what is disingenuous in Behe's presentation is that this preposterous challenge to offer a complete and step-by-step evolutionary account of IC systems is not spelled out from the beginning, but is a belated revision of his original claim, based on ambiguities in his definition of IC. In *Darwin's Black Box*, Behe leaves us with the impression that the unevolvability claim of IC is in principle easy to challenge, but when his critics take up the gauntlet, as we saw in the discussion with Pennock and Miller, Behe simply dodges and weaves like a hunted rabbit. Thus, what remains of Behe's argument boils down to the same old "argument from personal incredulity" (Dawkins 1991, p. 38), which is a far cry from the "objective criterion" for design that IDC theorists had promised.

It is interesting to note that the same pattern of reasoning has always been rampant in traditional creationist arguments regarding the so-called "gaps" in the fossil record. Creationists claim that they would readily accept evolution if only the "missing links" between the taxonomic groups turned up in the fossil record, but, whenever such a fossil is found, they complain that the intermediate is not really the ancestor of the present organism—an impossible demand for the fossil record—or even that Darwinists now face an ever bigger hurdle, because they are left with two gaps to explain. The latter principle has been coined "Gish's law" by geologist Robert S. Dietz (1983), after young-earth creationist Duane Gish.

7.4.2 Falsification and failure of instantiation

The design argument based on IC always allows for a final retreat. Suppose we can provide IDC proponents with a fully-detailed description of the evolution of the bacterial flagellum. Even if their stubborn insistence on the flagellum's exhibiting IC would at that point become absurd even in their own eyes (although one can never be too sure about that), they would surely regard this not as a refutation of IDC as such, but merely as a specific case in which IC turns out not to be instantiated. The expectation that this particular biological system would exhibit IC and hence be one of those unmistakable traces of design would simply be disappointed, and the search for new obstacles to evolution could begin.

In fact, this is what the history of the creationist movement is all about: if the case for evolution by natural selection becomes too overwhelming, creationists typically drop their favorite examples of complexity and come up with fresh ones, whose evolutionary origins are still relatively obscure (Pennock 1999, pp. 171-172). For example, the traditional objection against evolution used to be the vertebrate eye. Nowadays, the

evolutionary development of the vertebrate eye is well-understood and it has become an outdated argument against evolutionary theory. It is not even a particularly difficult example for evolutionary theorists, as it involves relatively straightforward selection pressures.

As the evolutionary history of the bacterial flagellum and the blood clotting system are being unraveled, the next generation of creationists can always disclaim the examples of their IDC forebears, and a new round of pointless arguments can begin—although they would at least have to admit that their former "design criterion" was defective because it generated false positive[s]. However, the retreat into unknown territory cannot go on indefinitely. In fact, as Robert Pennock (1999, p. 171) remarked, the current preoccupation of IDC theorists with invisible biochemical niceties such as the propeller system of *E. coli* bacteria indicates "just how far creationists have had to retreat to find significant explanatory gaps in evolutionary theory".

7.4.3 Bait-and-switch strategies

The most conspicuous feature of the concept of IC is not so much its ambiguity, but the discrepancy between what it seems to promise and what it eventually delivers, as far as testable empirical claims are concerned. On first reading Behe's argument, the unsuspecting reader may be left with the impression that Behe really sticks his neck out and presents evolutionists with a clear empirical challenge. However, this apparent rigor of the IC concept as an objective criterion for design, which arguably makes it appealing to anti-evolutionists, evaporates upon closer inspection. Under the weak interpretation, the concept describes a well-known phenomenon in the living world that is unproblematic for evolutionary theory. Under the strong interpretation, IC systems would indeed confront evolutionary theory with serious problems, but Behe has not given us an inkling of how we could ever demonstrate whether a system qualifies as IC in this sense. Indeed, it would require ruling out any conceivable evolutionary history, and would thus amount to showing that no part or precursor of the system in question is able to perform any other function, in any other situation and at any time.

This allows for an interesting bait-and-switch strategy, which one could summarize as follows: "First, present evidence for weak IC in the living world, then pretend that strong IC has been demonstrated and continue to equate IC with 'unevolvability.' If challenged on empirical grounds, jump back to the weak version and claim that your critics are misrepresenting your argument. Switch the IC claim to subsystems and assembly of components, keep raising the standards of evidence, and reassert that all this directly follows from the simple objective criterion of IC. Finally, when really

pressed against the wall, give up this particular system and quickly find a new one. Repeat the circle ad libitum".

7.4.4 Further equivocations

Behe's concept of IC is not the only instance of conceptual equivocation in the IDC literature. Two examples may be the subject of further research. First, when writing about "information," William Dembski surreptitiously switches between its standard interpretation in information theory, in which it is a measure of the randomness in a system, and its colloquial use in the sense of "meaningful message" (Perakh, Mark 2004, pp. 64-75). This ambiguity allows him to fool the reader into believing that the "information" encoded in DNA, for example, points in the direction of an intelligent designer. For a similar discussion regarding the term "teleological", see Blancke, Boudry & Braeckman (2010).

A second example is the IDC response to the series of mousetraps that John McDonald devised to refute the claim that gradualist evolution of IC systems is impossible (with the mechanical mousetrap as a paradigm example). Instead of admitting to their lack of imagination, IDC theorists have responded by complaining about the intelligent guidance used in constructing this evolutionary progression of mousetraps (Behe 2004, pp. 364-366). Amazingly, they argue that McDonald's mousetraps unwittingly demonstrate that an IC system always requires an Intelligent Designer. But this reply illegitimately shifts the discussion—which is actually about a human artifact and thus is, in any case, irrelevant—from the IC of a system to the blind and unguided character of evolution.

7.5 Conclusion

Although the IDC movement has been damaged, in terms of its credibility, by the Kitzmiller v. Dover case, it does not show clear signs of disappearing. As Forrest and Gross note in an afterword to their meticulous study of IDC's politics and religious ideology, the movement has simply changed its strategy once again. After their recent legal setbacks, they have been forced to drop overt talk of "intelligent design" and to adopt code words like "academic freedom" and teaching "the strengths and weaknesses of evolution" instead (Forrest, Barbara C. and Gross, Paul R. 2007, p. 337). "Creationists

never give up. They merely change their strategy with each new defeat" (Forrest, B. C. and Gross, P. R. 2007, p. 309).

As was apparent from its conception, the rapid success of the IDC movement was never driven by its arguments but by its religious ideology, which was epitomized in the so-called Wedge document of IDC's home base, the Discovery Institute (Forrest, Barbara C. and Gross, Paul R. 2007). Beyond religious motivation, one can point to sociological, cultural, and political factors to account for the remarkable success of IDC (outside the scientific community, to be sure), and it is plausible that the persistence of anti-evolutionary sentiments and the continuing appeal of the design argument is also a function of deeply rooted cognitive dispositions and hard-to-shake teleological intuitions about the world (Kelemen and Rosset 2009; Kelemen 2004).

Anti-evolutionism can take many different forms, however, and not all of them can achieve equal cultural success (see Chapter 11). In this paper, we have analyzed some of the rhetorical strategies that Behe and other IDC theorists have used for presenting their challenge to evolution and for deflecting valid criticism. On the one hand, we claim that Behe's presentation of IC has the appearance of an objective design criterion, which makes it superficially more respectable than the age-old "argument from personal incredulity." On the other hand, the equivocations that are built into the definition of IC allow it to be used as a moving target (Boudry and Braeckman 2010), and as a kind of conceptual chimera that is hard to pin down by critics. These considerations partly explain why the concept of IC was hailed by the movement as the ultimate challenge to evolutionary theory, and why, despite its complete lack of scientific merits, it has convinced even some knowledgeable persons of the impending demise of evolutionary theory. As Robert Pennock wrote:

> We think of creationism as a cluster of ideas that reproduces itself by spreading from mind to mind and struggling with competing ideas for a home among a person's beliefs. Sometimes it loses out to more powerful rival ideas, but sometimes it finds receptive mental soil, takes root and waits to be passed on again. (1999, p. 1)

Indeed, in the past two decades the concept of IC seems to have found receptive mental soil among anti-evolutionists. An analysis of the convenient conceptual equivocations inherent in IC, as well as of the rhetorical strategies with which IC has been presented, helps us to understand this remarkable fertility.

Chapter 8.
Even Better than the Real Thing: Simulation of Biological Evolution under Attack

Abstract.[56] Intelligent Design theorist William Dembski (2002) has argued that the first No Free Lunch theorem, formulated by Wolpert and Macready (1997), renders Darwinian evolution impossible. Several of Dembski's critics have pointed out that the theorem is in fact irrelevant to biological evolution, but now mathematician Ronald Meester (2009) maintains that the theorem still applies to *simulations* of evolutionary processes. According to Meester, the theorem shows that simulations of Darwinian evolution, being typically set in advance by an intelligent programmer, are laden with teleology and therefore fail to capture properly Darwinian evolution. As a result, they are uninformative as to how complex biological adaptations arise in nature. We argue that Meester's use of the term "teleological" is particularly sloppy, and that he equivocates on the notion of "tailoring around" a fitness function, which is central to the NFL theorem. Meester is confused both about the NFL theorem and about the distribution of fitness functions in real life, and in the end his position is indistinguishable from that of Intelligent Design Creationism (IDC) advocate William Dembski.

[56] This chapter is loosely adapted from Blancke, Boudry & Braeckman (2010), which offers a specific rebuttal of Meester's claims in the journal *Biology & Philosophy*.

8.1 Introduction

In *No Free Lunch. Why specified complexity cannot be purchased without intelligence* (2002), the leading IDC theorist William Dembski argues that the first NFL theorem prohibits the evolution of complex adaptations by Darwinian evolution. This theorem, first published by Wolpert and Macready in 1997, establishes that no optimization algorithm can outperform a random search when averaged over all fitness functions. In this way, the theorem rules out the possibility of a universal, free-for-all algorithm outperforming a random search on any fitness function. For an algorithm to perform better than mere chance over a particular fitness function, the algorithm has to be *tailored around* that fitness function.

On the basis of this mathematical result, Dembski has argued that, for the algorithm of natural selection[57] (NS) to outperform blind chance, extra information about the particular fitness function is required. This search for the required information input, if we are to be believe Dembski, is even harder to accomplish than the original search for fitness optima performed by NS. Dembski has termed this the *displacement problem*. In order to avoid infinite regress, extra 'information' must be supplied by an intelligent designer. More specifically, the parameters of the environment have to be fine-tuned by the designer to allow NS to work successfully.

Dembski's book has met with devastating critiques. Some of Dembski's critics (e.g. Wolpert 2002; Shallit 2002) complained that his writings are so vague that it is almost impossible to pinpoint his actual position. As concerns Dembski's use of the NFL theorem, most critics concur that it is entirely irrelevant to biological evolution (Häggström 2007; Perakh 2002; Perakh, M. 2004; Rosenhouse 2002; Sarkar 2007). Darwinian evolution is the result of NS acting over specific fitness functions, i.e. the ones that we find in realistic biological environments; biological evolution is simply not concerned with averaging over *all possible* fitness functions. This means that, within this particular setting, nothing prevents NS from outperforming random search. As the NFL theorem does not apply for specific subsets of all possible fitness functions, in practice it is irrelevant to biological evolution.

[57] See Dennett (1996) on NS as an evolutionary algorithm, and Wolpert and Macready's (1997) account of NS as an optimization algorithm

8.2 No Free Lunch and simulations

In a recent paper in *Biology and Philosophy,* Ronald Meester (2009), a Dutch mathematician and IDC sympathizer[58], has subscribed to this critique: "it is simply not the case that a biological fitness function can be viewed as an average over all possible fitness functions. [...] Therefore the NFL theorem simply does not apply" (2009, p. 464). Unlike other critics, however, Meester believes that the "algorithmic 'NFL way' of thinking about evolution is very meaningful when it concerns computer simulations of certain evolutionary processes" (2009, p. 468).

To support his argument, Meester discusses "two examples of the NFL theorem in action" (2009, p. 464). Both involve the use of a search algorithm to find a particular target in the form of a letter string. As in Richard Dawkins' (1986) illustration with a line from Shakespeare's *Hamlet*, an algorithm combining random variation and selective retention is shown to outperform mere chance in finding the target string (in Dawkins' case, ME*THINKS*IT*IS*LIKE*A*WEASEL; in Meester's version, the word YES). Meester draws on the NFL theorem to make the following argument:

> [the researcher's] algorithm is too efficient to be the result of averaging over all fitness functions; it is not likely that he chooses his fitness function uniformly at random over all possibilities at the start of each new search. No, it is reasonable to conclude that he uses the fitness function corresponding to the word YES, *and* that he uses the search algorithm associated with that word. Again, note that the conclusion is twofold: we know that he uses special fitness functions and we know that his search algorithm is tailored around his choice in order to get an efficient algorithm. (Meester 2009, p. 466)

According to Meester, the programmer has already selected particular fitness functions and a suitable algorithm for reaching the target *in advance*, making sure that the algorithm is "very carefully tailored" (p. 471) around the fitness functions. This makes the whole affair "intrinsically" (p. 468) or "necessarily teleological" (p. 471). No simulation of the evolution of complex biological adaptations, no matter how sophisticated (e.g. Lenski, Ofria et al. 2003), escapes this conclusion. As Darwinian evolution is supposed to be non-teleological and undirected, computer simulations cannot deliver any understanding of real-world evolution. There are, however, several difficulties with Meester's position that radically undermine his conclusion.

[58] For further documentation of Meester's endorsement of IDC, see Blancke et al. (2010).

8.3 Difficulties

8.3.1 Setting a target

Meester acknowledges that NS can be understood as an algorithmic search procedure (Meester 2009, p. 464). By implementing this algorithm in a computer simulation, however, Meester wants us to believe that, all of a sudden, it ceases to be same Darwinian algorithm. But why would this be so? The mere fact that we are dealing with a computer simulation instead of a real-life situation is irrelevant, for algorithms are substrate neutral. It does not matter whether an algorithm is implemented in a biological environment or in a silicon-based digital one. As long as the conditions of variation, differential survival and heredity apply, evolution by NS is bound to take place, irrespective of the medium (Dennett 1996). Running the algorithm of random variation and selection on a computer simulation does not alter its non-teleological character.

On closer inspection, however, Meester's argument equivocates between two senses of the term "teleological", which he applies indiscriminately to programs, simulations and algorithms. In one sense, "teleological" stands for "being aimed at a target" or constructed "with insight into the future" (p. 470). If the programmer is modeling an algorithm with a preset target in mind, that makes his simulation "intrinsically teleological".

> [...] the programmer chose his program with insight in the future goal he wanted to reach and hence the simulation was intrinsically teleological. (p. 468).

By contrast, modeling bacteria resistance to antibiotics is deemed unproblematic by Meester, as their goal is "not to reach a special target, but instead to compare the 'typical' behavior of related systems" (p. 470). In that sense of the word, however, the question of the "teleological" nature of simulations has no bearing on their validity. Of course the programmers have a "goal" in mind, but as long as they make sure that the *algorithmic process* itself, in particular the source of variation, is undirected, this does not affect the validity of the simulation. The algorithmic process is not 'teleological' in any interesting sense at all.

If Meester's argument were to hold water, then almost any simulation or model of natural processes and phenomena, no matter how careful and sophisticated, would become 'teleological' and hence uninformative about the actual natural events. Weather forecasts, for example, are set up by intelligent humans with specific goals in mind, which would make them intrinsically teleological and hence unsuitable to talk about real weather phenomena, because the latter are thought of as undirected, natural

processes. Now, if any simulation is "vulnerable" to Meester's critique, this ultimately raises the question as to why Meester singles out simulations of biological evolution as his main target.

8.3.2 Careful tailoring?

Arguing on the basis of the NFL theorem, Meester claims that the search algorithm in a simulation is always "very carefully tailored" (p. 471) around particular fitness functions to get at a specific target. He points out that the programmers have not chosen the fitness functions in the simulation "at random over all possibilities" (p. 466). This is no wonder, however, because *neither are they in the biological world*. Fitness functions in real life exhibit a significant amount of what Häggström (2007) terms "clustering properties", which means that the fitness values of two highly similar DNA sequences are not statistically independent. In particular, "similar DNA sequences will tend to produce similar fitness values" (2007, p. 228), allowing a search algorithm like NS to perform much better than blind chance. In other words, there are "slopes" around a local fitness optimum which NS can work with. As Perakh puts it:

> The evolutionary algorithms, both designed by intelligence and occurring spontaneously, deal with given, specific fitness functions and have no need to search the information-resource space [i.e. all possible fitness functions]. (Perakh, M. 2004, p. 170)

Once we realize that the NFL theorem applies only when we average the performance of our algorithm over *all* possible fitness functions, then we see that the theorem is in fact a rather trivial mathematical result. With some justification, Häggstrom has written that the NFL theorem is a "fancy (and more general) way of phrasing the following fact":

> If we spread a well-shuffled deck of cards face-down on a table and wish to find the ace of spades by turning over as few cards as possible, then no sequential procedure for doing so is better than any other. (Häggström 2007, p. 226)

The same point applies to the search algorithm itself, which is "tailored" only in the sense that it is specifically programmed to mirror the *actual* biological search algorithm, i.e. random variation and selective retention. In fact, what Meester objects to in these simulations is precisely what makes them successful simulations in the first place, i.e. that they mimic real life condition. Meester thinks his position is different from Dembski's because he (Meester) tries to restrict the "NFL-way of thinking" to *simulations* of evolution. However, because his argument is based on a misunderstanding of the

distribution of fitness functions in *real life*, his position is actually *indistinguishable* from Dembski's.

8.4 Conclusion

Meester uses the term "teleological" in a sloppy and misleading way. After pointing out that simulations are designed by intelligent agents with particular goals in mind, Meester pretends that he has shown the algorithmic process in a simulation to be "teleological" in nature. Meester thinks the NFL theorems are relevant to evolution by NS because, following in Dembski's footsteps, he misunderstand the nature of actual fitness functions in real life. Furthermore, Meester equivocates on the technical term "tailoring around" to suggest that simulations are nothing but cheap tricks in which the outcome is already built in the program in advance. Pace Meester, a Darwinian algorithm does not cease to be Darwinian if simulated in a computer program. Meester's argument on the basis of NFL fares no better than Dembski's, and in fact their positions are all but indistinguishable.

Chapter 9.
Obscured by Lacanian Clouds? Epistemic Defense Mechanisms and the Declarative Fallacy

> *Having it both ways is essential to the appeal of postmodernism, for it is precisely by apparently speaking simultaneously of two different concepts with the same word that the appearance of giving a profound but subtle analysis of a taken-for-granted concept is created. – Nicholas Shackel (2005, p. 304)*

Abstract.[59] Many philosophers, cultural theorists and psychoanalysts begin their project of elucidating Jacques Lacan's writings with explanations of and justifications for their master's obscure voice. Interestingly, these arguments often draw on theory-internal resources, and seem to take the reader hostage: the style and presentation of Lacan's writings are assumed to reflect his profound theoretical insights about unconscious meaning and language. We review several *epistemic defense strategies* for coping with the conceptual obfuscations of Lacanian theory and the obscurity of Lacan's discourse. Finally, we analyze the hermeneutic situation created by Lacanian obscurantism in terms of the *Declarative Fallacy*.

[59] This chapter is a revised version of a paper that has been written together with Filip Buekens. A number of examples from the Lacanian literature are drawn from Buekens (2006b).

9.1 Introduction

In a now famous summary discussion in the even more infamous *Livre Noir de le Psychanalyse* (The Black Book of Psychoanalysis), published in 2005, the French philosopher Mikkel Borch-Jacobsen (2005, p. 180) argued that psychoanalysis, in its many (often incompatible) versions, is an *empty theory* ("une théorie vide"): the key concepts, he argued, have become 'empty signifiers', that could be interpreted at will (similar conclusions were anticipated by Cioffi 1998). Particularly in Lacanian psychoanalysis the vicissitudes of such central concepts like *jouissance*, the Other, the *objet petit a* or enigmatic claims as that "the unconscious is structured like a language" or that "The woman does not exist" are such that no one really understands what they mean or what they have meant in the hands of their originator. In the first part of this paper we explore a number of (post-)Lacanian arguments to the effect that the incoherence of key concepts in Lacanism reflects the very nature of the unconscious, and that the obscurity of Lacan's writings reflects his own profound insights about the unconscious and language. Such arguments can best be seen as examples of *epistemic defense mechanisms* explored in Chapter 5 (Boudry and Braeckman 2010), which offer a theory-internal rationale for fending off criticism. Appeal to this epistemic defense mechanism is provoked by committing what we call the *Declarative Fallacy*.

9.2 Let there be more light

In Chapter 5, we argued that epistemic defense mechanisms draw on internal conceptual resources to fend off criticism and adverse evidence. In practice, the distinction between immunizing strategies and epistemic defense mechanisms may prove difficult to maintain. An argument that was initially developed as an immunizing strategy may over time be integrated in the structure of the theory, thus blurring the distinction between the theory proper and the immunizing strategies that protect the theory that engenders it. Alternatively, defenders may find a way to short-circuit criticism that loosely draws on theory-internal resources, yet cannot be reasonably construed as an *intrinsic* part of the theory.

From the perspective of the critical outsider, to rely on arguments that already presuppose the truth of the theory one tries to defend is obviously an instance of circular reasoning. But this observation does not reflect the persuasive force of such

arguments in the eyes of those who use the argument and accept the propositions or theories they are supposed to defend. The conspiracy theorist, for whom the very existence of adverse evidence attests to the extreme cunning of the conspirators, and in the eyes of whom disbelievers are themselves involved in the evil plot, offers a prime example. To argue that conspiracy theorists already *presuppose* the truth of the conspiracy, and are thus engaged in a form a circular reasoning, will evidently be to no avail in debating them. Indeed, the fact that a belief system can account for seemingly adverse evidence, that it makes the existence of disbelievers intelligible (and predicts their existence) is a very compelling fact for he who is already committed to that belief system. In this chapter, we will apply this perspective to the writings of Jacques Lacan and to different rationalizations of his obscurity that have been put forward.

9.3 Obscured by clouds

Interpreters of Lacan have proposed several theory-internal explanations for the notorious obscurity of his pronouncements (Buekens 2005) and other problematic features of his discourse. An intriguing version of the argument directly appeals to Freud's conception of dreams as decipherable rebuses in the *Traumdeutung* (Freud 1953a, 1953b), of which Lacan said: "Cet ouvrage ouvre avec l'oeuvre sa route royale à l'inconscient" (Lacan 1966, p. 509). We are told that Lacan's work is a 'rebus', just as dreams are: underneath a manifest dream content is hidden a latent dream content the correct reconstruction of which will reveal a web of repressed desires that cast a shadow over the dreamer's life (Lacan 1966, p. 470). Freud's model of the dream and its alleged meaning can easily be applied to Lacan's writings:

> It does not seem unfair to characterize Lacan's writings in this way [as a rebus] ...[f]or their substance deals with the nature of the unconscious as Freud understood it, hence with that dimension of human experience that lies beyond the kern of conscious, rational discourse and emerges into awareness only through a din of diffraction that may assume many forms – in the case of dream, for example, the form of a rebus. By saying, then, that Lacan's work, in terms of its substance, is a rebus, we mean to suggest that it is dealing with a theme that of its very nature escapes the constriction of rational exposition. (Muller and Richardson 1982, pp. 2-3)

Some Lacanians have exploited this rebus metaphor to develop a theory-internal rationale for Lacan's obscurity. Madan Sarup develops the *Imitation Argument* as follows:

> Lacan's writings are a rebus because his style mimics the subject matter. He not only explicates the unconscious but strives to imitate it. The unconscious becomes not only the subject matter but, in the grammatical sense, the subject, the speaker of the discourse. Lacan believes that language speaks the subject, that the speaker is subjected to language rather than master of it. (Sarup 1992, p. 80)

Sarup appeals to Lacan's theory of the subject as constituted by language or discourse: 'language speaks the subject' and 'the subject is not the master of its discourse'. As the Lacanian psychoanalyst Paul Verhaeghe put it:

> From an analytic point of view [...] the subject does not speak, it is spoken. As a result, the subject floats on top of spoken words. Indeed, when 'I' speak, I do not know what 'I' am about to say, unless I am reading it or have learnt it by heart. In all other cases 'I' is spoken by a desire outside my consciousness that drives me, sometimes with approval, sometimes without. And what I do say always comes, in the final analysis, from the Other. (Verhaeghe 2004, p. 56)

Lacan, we are told, serves as mouthpiece of the unconscious; his discourse is a perfect (and therefore instructive) *imitation* of the unconscious.[60] Dany Nobus seems to confirm this view when he writes that:

> [Lacan] modeled his own discourse on the very rhetoric of the unconscious which he believed to have discerned in Freud's foundational accounts of dreams, slips of the tongue and jokes (Nobus 2004b, p. 196)

Muller and Richardson held that Lacan's *Ecrits* and the *Séminaires* are "essentially a concrete demonstration in verbal locution of the perverse ways of the unconscious as he experiences it" (Muller and Richardson 1982, p. 3). Analogous remarks can be found in Benvenuto & Kennedy (1986, p. 13) and Caudill (1997, p. 5). In other words, the obscurity of Lacan's writings reflects the obscure and symbolical nature of the unconscious, and its pervasive influence on our conscious speech acts. As we will see in the following section, one of the central tenets of Lacanian psychoanalysis is precisely that humans are trapped in a web of signifiers pregnant with unconscious associations, and that the real meaning of words can never be attained. Lacan's writings illustrate that communication can never succeed, because there is always an *objet petit a* (see below) eluding us and resisting further explication (Verhaeghe 2004). There is,

[60] The verbal puns and wordplays in Lacan's discourse are justified by Stanley Leavy: "The theoretical basis of (his) playing with words is found in Lacan's dictum that 'the unconscious is structured like a language'. In his playful punning this claim is concretized, embodied. The unconscious can speak truthfully, revealing the identity of the logically unrelated, cognitively distorted, affectively confused experiences" (Leavy 1983, p. 13).

therefore, an important sense in which Lacan's own theorizing and his expository style are a *performative expression* of the subject matter of the theory itself.

These Lacanian pronouncements seem to defy any sensible notion of intentional speech acts and authorship[61], but for many followers they constitute a suggestive and intriguing explanation of Lacan's obscurity. Ironically, the allegedly inscrutable character of a theory's subject matter entails that every effort to present Lacan's conception of the unconscious in a more or less streamlined fashion will eventually end up as a fatal distortion of the theory's subject matter. Trying to present his thoughts in a systematic and orderly way, as countless interpreters of Lacan have been doing, will fatally *misinterpret* him (and the unconscious) if the Imitation Argument were sound. Thus, Lacanian discourse would forever be condemned to obscurity.

If we are to understand Lacan, we must assume that his discourse is the expression of his unconscious, which is itself structured like a language.[62] So, if the Imitation Argument is sound, readers are taken hostage: Lacanian insights reveal themselves *only if* (not: if) one takes the theory to be a correct imitation of the unconscious. By contraposition: if one denies that Lacan's simulation is faithful to the nature of the unconscious, the content of the theory will not reveal itself to the reader. The hermeneutic circle drawn by Sarup and others is so tight that there is no space left for reasonable dissent, and the whole theory is conveniently insulated against criticism.

9.4 Momentary lapses of reason?

Many critics of Lacan have taken issue with the conceptual incoherence of the different versions of his theory (Sokal and Bricmont 1997; Buekens 2006b; Borch-Jacobsen 1991). Lacan's pronouncements are couched in a number of highly abstract and complex concepts – the Other, the Symbolic, the objet *petit a*, jouissance, the Phallus etc. – which are notoriously difficult to understand. Indeed, even among Lacanians there is no

[61] By definition, a theorist *intends* to say something when he says something. Indeed, the choice of words, the structure of arguments, the precise formulation of ideas constitute the very core of the *intentional* activity called 'theorizing'.

[62] "Le symptôme psychanalysable... est soutenu par une structure qui est identique à la structure du langage... la structure du langage telle qu'elle se manifeste dans les langues que j'appellerai positives, celles qui sont effectivement parlées par les masses humaines.' And: 'L'inconscient est structuré comme un langage" (Lacan 1966, p. 444).

consensus about their meaning (Nobus 1998) and deep theoretical divides continue to exist within the Lacanian community.

An instructive example is the pivotal concept of the big Other, which Lacan characterizes as an abstract 'locus' in the psychic structure of the subject that can be 'occupied' by a range of different signifiers/objects. These are said to *represent* or *function as* the Other for the subject. If one consults and compares different works by Lacanian interpreters, it turns out that the Other can stand for other individuals, society, the law, moral order, the mother-figure, the psychoanalyst himself, the opposite sex, a person's own body, Language, images or even – according to Slavoj Zizek – the simulated reality in the movie the Matrix.

All these representations are said to 'occupy' the 'position' of the Other, but no meaningful conceptual unity can be discerned. A close analysis of the Lacanian literature reveals that a concept such as the Other functions as *a container term* into which any pseudo-theoretical insight about human psychology can be couched. In a paper entirely devoted to the concept, Derek Hook characterizes "the Other" as a "vanishing-point of inter-subjectivity", as "simultaneously 'inside' and 'outside'", and he writes that the Other is "both embodiment of the social substance and yet also the site of the unconscious" (Hook 2008).

Again, however, the followers of Lacan have drawn inspiration from his own arcane views on language and meaning to account for this conceptual befuddlement. The justifications exploit Lacan's idiosyncratic version of Saussurian linguistics, which emphasizes the primacy of the signifier, and Lacan's conception of the unconscious that is structured as a language. "*Of course* the concepts of the Other and the Real are difficult to explain coherently" we are told by the Lacanian theorist, "because meaning can never be fully grasped. As soon as we human beings enter the domain of the Symbolic, we are forever trapped in a web of signifiers." Signifiers can never refer to something real out there, but merely to other signifiers. Because the matrix of signifiers constituting our language is a closed system from which we cannot escape, we are unable to grasp meaning and comprehend each other. When we speak, we are not aware of what we are saying. Lacan believes that language speaks the subject (Lacan 1966), that the speaker is *subjected* to language rather than master of it (Sarup 1992, p. 80). "'I' is spoken by a desire outside my consciousness that drives me" (Verhaeghe 2004, p. 56), and thus we *are being spoken* ("ça parle"). Indeed, some Lacanians believe that the very nature of the subject matter of Lacanian theory escapes rational discourse and scientific evaluation (Leguil-Badal 2006). Lacan himself wrote: "Le réel est, il faut bien le dire, sans loi. Le vrai réel implique l'absence de loi. Le réel n'a pas d'ordre" (Lacan 2005, pp. 137-138). In sum, Lacan's theory reflects profound theoretical insights into the nature of the unconscious and the structure of language.

And what about the apparent contradictions and paradoxes in Lacan's work? We are told that these reflect the divisiveness and structural 'lack' characterizing the human

subject. When we are being introduced in the Symbolic order, our psychic structure develops into certain "knots" that are irreducible to theoretical formulation. Lacan has designated this ineluctable lack as the locus of the *objet petit a*, a concept which, as Lacanian interpreter Bruce Fink himself acknowledges, "can take on many different guises" (Fink 1997, p. 52). According to Lacan, the *objet petit a* is that aspect of the Real that cannot be represented, that forms a structural break in the chain of signifiers. "The [object] a is what remains irreducible in the advent of the subject at the locus of the other, and it is from this that it is going to take on its function" (Lacan 2004, p. 189). The later Lacan coined the term 'sinthome' for that which is beyond meaning and unanalyzable in the so-called topology of the human mind. Other Lacanian concepts fulfill similar roles:

> The Borromean knot marks the outer limit of Lacanian theory, the point where the formalising ambition of the matheme finally collapses into the non-theorizable, the untraslatable real of the symptom. (Thurston 1998, p. 158)

In the same vein, some Lacanians have even tried to explain (and justify) the institutional crisis of their own discipline in theory-internal terms. Reflecting on the many theoretical schisms following the death of Lacan, and the feuds over his intellectual legacy, Nobus writes that knowledge is always "in a state of continuous dispossession", and this has (of course) something to do with the mysterious Other:

> If psychoanalytic knowledge is by definition a knowledge in failure, isn't the crisis of legitimacy a necessary precondition for the discourse of the analyst to sustain itself? Perhaps the only agency that could ever be in the position of owning psychoanalysis is the (unconscious of the) analysand, the Other of psychoanalytic discourse [...] (Nobus 2004a, p. 222)

Indeed, one can construe Lacanian arguments to the effect that the very act of questioning the truth value of Lacanian psychology becomes deeply misguided, because according to Lacan truth itself has a "fictional structure". As Fink wrote in relation to the question of the scientific value of Lacanian theory:

> The fact remains that *science is a discourse*. [...] it implies a dethroning of Science and a reassessment of science as *one* discourse among many. [...] Lacan's discourse theory suggests that there are as many different claims to rationality as there are different discourses. (Fink 1995, p. 138)

We now understand why, at least from a Lacanian point of view, any attempt to criticize the seemingly willful obscurity of Lacan is to reveal a complete lack of understanding of the fundamental Lacanian insights about language, truth and reality that are at stake. We also begin to see why Lacanian interpreters are hardly impressed by critics objecting to the conceptual vacuity or contradictions in his writings. From the theory-internal

perspective, the critic of Lacanian psychoanalysis can only be construed as one who fails to appreciate the divisive dimension of subjectivity, who ignores the elusive *objet petit a* of all our strivings, and who clings on to illusions of objectivity and comprehension. The fundamental tenets of Lacanian theory about the *objet petit a*, the primacy of the signifier and the divisive nature of the human subject are perfect instantiations of epistemic defense mechanisms, as explored in Chapter 5. They ensure that a critical analysis of Lacanian theory never succeeds in affecting the system.

There is a deeply self-fulfilling dimension to Lacanian theory. For those impressed by Lacan's charismatic authority, his pronouncements about language, meaning and reality are compelling partly because they constitute such a remarkably apt account of the *Lacanian edifice itself*. After all, what better illustration of the primacy of the signifier over the signified and the elusiveness of meaning than Lacan's ever-shifting and abstruse conceptual apparatus? What better way to appreciate the controversial claim that meaning always escapes us than to spend a few hours reading Lacan's *Écrits*?

Fundamental to the working of this epistemic defense mechanism is the inflation of Lacan's pronouncements in the restricted context of psychoanalytic therapy to profound insight into the structure of language, human communication, the status of knowledge, and even the nature of reality ('the real') etc. (to be sure, these extrapolations of psychoanalytic insights were already instigated by Lacan himself, particularly in his later years). By projecting the theoretical flaws and paradoxes of their own 'discourse' to *any* other theoretical endeavor and elevating them to deep linguistic and epistemological insights, Lacanians have tried to evade any accountability to traditional standards of conceptual clarity, consistency and explanatory power. Unfortunately, the Lacanian psychoanalyst has confused the predicament of his own belief system for that of the rest of the world. Paraphrasing Karl Kraus, one of Freud's earliest critics, one could say that Lacanian psychoanalysis is itself the disease of which it claims to be the cure.

9.5 The declarative fallacy

Having unraveled a deeply circular defense strategy, it remains a legitimate question what explains the fact that so many convinced themselves that there ever was any substance to the master's arcane pronouncements. For example, what exactly is it that makes Lacanians like Derek Hook wonder why the concept of the *big Other* has been "curiously neglected by critical social psychology" (Hook 2008)?

We propose that the persuasive effect of using these 'signifiers' in interpreting common experiences and the vicissitudes of life lurks in the declarative force of the speech acts in which they function (Searle 1969, 2010) and the meta-representational character of the key Lacanian concepts figuring in the sentences that are used as declaratives. Declaratives are a class of speech acts that combine the word-to-world direction of fit that characterizes orders, for example, with the world-to-word direction of fit that characterizes descriptive statements like assertions. Performatives are probably the best known examples of such speech acts, for they *make* something the case by *declaring* that it is the case (Searle 2010, p. 12). However, the phenomenon has wider application. When a key Lacanian concept is used in a declarative speech act, the speaker simultaneously declares that the sentence is true, thereby assigning to the Lacanian meta-representational concept that figures in it an extension *that is intended to make the sentence true*. It is because Lacan (or Zizek, or...) *declares* that phenomenon X is (an exemplification of) psychoanalytic phenomenon Y that X becomes (illustrates, exemplifies) phenomenon Y. When the truth of the ensuing pronouncement is duly accepted by the intended audience, the term Y is *ipso facto* accepted as having an extension that makes the statement true. The central concepts thus function as "meta-representations" (Atran 2002, p. 276; see also Sperber 1997), in the sense that accepting the statement as true does not presuppose that one has *understood* the sentence in question – by accepting the sentence, the intended audience cooperates with the speaker in making the sentence true. Thus, the project of 'understanding' Lacan must proceed from the assumption that Lacan declares that certain propositions are true, for abandoning that assumption immediately exposes the empty character of his pronouncements. According to Sperber et al., this epistemic charity is typical for the deferential attitude towards religious authorities, gurus and other persons with inflated reputations:

> If [people] were to check the pronouncements of these sources (for instance, „Mary was and remained a virgin when she gave birth" or Lacan's „There is no such thing as a sexual relationship") for coherence with their existing beliefs, they would reject them. But this would in turn bring into question their acceptance of the authority of the source. A common solution to this predicament is to engage in a variant of Davidsonian „charitable interpretation", and to „optimize agreement" not by providing a clear and acceptable interpretation of these pronouncements, but by deferring to the authorities (or their authorised interpreters) for the proper interpretation, and thus accepting a half-understood or „semi-propositional" idea. (Sperber, Clément et al. 2010, p. 382)

But this is a travesty of normal communication: while every interpretation requires a certain amount of charity on behalf of the intended audience, it cannot be allowed that, in order to understand what is being said, one must first *accept* them as declarative

pronouncements that create the truth of what they declare. This, we suggest, explains why almost every interpretation of Lacan begins with an explanation of his obscurantism, and then goes on trying to understand him as if everything he says (but in fact declares to be true) is true.

It should be obvious that it is difficult to criticize a discourse that imposes on its key concepts ('signifiants') an *ad hoc* meaning such that whatever it means on the occasion of use will make the declarative pronouncement true. One is reminded here of the dialogue between Alice and Humpty Dumpty in Lewis Carroll's *Through the Looking Glass*:

> `When I use a word,' Humpty Dumpty said, in rather a scornful tone, `it means just what I choose it to mean -- neither more nor less.'
> `The question is,' said Alice, `whether you can make words mean so many different things.'
> `The question is,' said Humpty Dumpty, `which is to be master -- that's all.' (Carroll 1993, p. 223)

We propose to coin this the Declarative Fallacy: 'When Lacan (or someone speaking on behalf of Lacan) declares that *p*, then *p*'.[63] The Declarative Fallacy exploits a familiar and extremely useful feature of language, viz. the fact that we are always free to introduce new terms or concepts via reference-fixing stipulations or baptisms (Kripke 1980). But if one introduces a term *t*, one must fix its reference in such a way that its future users are able to use it with the very same extension or referent its original user introduced (Kripke 1980); this assumes that the reference fixer remains constant and is open to all or that, as Wittgenstein taught us in the *Philosophical Investigations*, its meaning is "public". In normal contexts, what fixes the reference of a term (often introduced via a declarative speech act) remains constant and later uses of a term *t* (in descriptive speech acts) respect the initial stipulation or baptism. The Declarative Fallacy allows users of Lacanian signifiers to employ a term or concept at will *at any occasion* in declaratives that give it an *ad hoc* meaning that makes the statement true, and thereby achieve immunity from criticism. Moreover, the fallacy requires a curious form of cooperation between speaker and intended reader or audience: the speaker must be taken to have the authority to *declare* that *p*, while the intended audience must accept those pronouncements as truths. This puts them in a position where they have to seek interpretations of the suspect concepts that make the sentences true and to seek justifications for why they do that – the defense strategies explained earlier. But this is a parody of hermeneutics and it explains why, as philosopher of language François Recanati once suggested, interpreters of Lacan "accept everything", but "constantly

[63] In Buekens and Boudry (in press) we have analysed further, and more disturbing effects of the Declarative Fallacy and its social acceptance conditions.

dispute what was said" (Recanati 2000, p. 263; see also Buekens 2005; Buekens 2006b). That is the real hermeneutic nightmare behind sustained conceptual obscurantism: everyone accepts what is declared to be true, but nobody understands what the statements mean. It might well be that this rather awkward situations characterizes what happens on the Lacanian couch, but if one wants to develop a theory about the human mind and its vicissitudes, the practice yields disastrous results. The effects of the declarative fallacy are pernicious exactly because the honest reader or interpreter is forced to manipulate his or her natural sense of charity: while a legitimate presumption of truth is a natural attitude in hermeneutics, in the case of Lacan the reader is forced in a position such that if she does not *accept* as true what is being said, she will never understand what Lacan intended to communicate. The reader is thus taken hostage by Lacan, and the latter inevitably becomes an invulnerable guru whose obscurantism must be legitimized by all available means if the intention to understand is to be maintained.

9.6 Conclusion

The pervasiveness of epistemic defense mechanisms and the declarative fallacy sheds light on the continuing appeal of Lacanian theory in the humanities and 'Cultural Studies'. Lacan's 'open discourse' – itself often imitated by his followers – creates limitless possibilities for applying Lacanian theory to works of art, movies, politics, culture etc. In other words, it leaves an academic interpreter "without even a mathematical chance of having nothing to say" (Crews 2006, p. 61). The declarative fallacy suggests to the gullible reader that a profound insight is lurking right behind the corner, as something that can be glimpsed yet not fully grasped. At the same time, the theory-internal resources of Lacanian theory allow interpreters to explain away conceptual incoherence and obscurantism, and to protect the theory from critical outsiders.

Consequently, we may expect that the Lacanian interpreter will perfectly 'understand' our objections. Surely we are in a futile quest for the *objet petit a* of our desire to understand. Or maybe we cling to the *sujet supposé savoir*, and we are "divided subjects" (*sujets barrés* $) who are unable to confront their "master signifier" (significant maître S1). How could anything be meaningless in the eyes of the Lacanian beholder?

Chapter 10.
The Epistemic Predicament of a Pseudoscience: Social Constructivism Confronts Freudian Psychoanalysis

Apart from the lack of agreement with reality it is in any case a superb intellectual performance. – Albert Einstein on Hermann Weyl's unified field theory

Reality is what refuses to go away when I stop believing in it. – Philip K. Dick

Abstract.[64] Social constructivist approaches to science have often been dismissed as inaccurate accounts of scientific knowledge. In this paper, we take the claims of robust social constructivism seriously and attempt to find a theory which *does* instantiate the epistemic predicament as described by SC. We argue that Freudian psychoanalysis, in virtue of some of its well known epistemic complications and conceptual confusions, provides a perfect illustration of what SC claims is actually going on in science. In other words, the features SC mistakenly ascribes to science in general *correctly* characterize the epistemic status of Freudian psychoanalysis. This sheds some light on the internal disputes in the field of psychoanalysis, on the sociology of psychoanalytic movement, and on the "war" that has been waged over Freud's legacy with his critics. In addition,

[64] This chapter is due to appear in *Theoria* (Boudry and Buekens 2011).

our analysis offers an indirect and *independent* argument against SC as an account of *bona fide* science, by illustrating what science would look like if it *were* to function as SC claims it does.

10.1 Close encounter of a strange kind

In recent controversies over psychoanalysis, which for their unremitting vehemence became known as the "Freud Wars", the scientific status of Freud's legacy has been hotly disputed. Almost simultaneously, another war was being waged over the status of science itself. In the so-called "Science Wars", social constructivist approaches to science were put against more traditional epistemic approaches to science and scientific rationality. Our aim in this paper is to show that classical Freudian psychoanalysis, in virtue of some of its well known epistemic deficiencies, methodological complications and conceptual confusions, provides an intriguing illustration of what science *would* look like if it *were* to function as SC claims it does.

The analysis developed here yields an independent and indirect argument against SC as an account of bona fide science, and reveals a fascinating connection between the two "wars" (for previous encounters between SC and psychoanalysis, see Moore 1999; Gillett 1998; Borch-Jacobsen and Shamdasani 2008). In particular, we show how the framework of SC sheds some light on the internal disputes in the field of psychoanalysis, on the sociology of psychoanalytic movement, and on the "war" that has been waged over Freud's legacy with his critics. Our own approach partly builds on earlier work on the reconstruction of psychoanalytic hermeneutics as a system of unintended institutional facts (Buekens and Boudry 2010; Buekens 2006a), based on Searle's theory of institutional facts (1995).

10.2 Social constructivism (SC)

10.2.1 Epistemological SC

Social Constructivism (SC) comes in many flavors, and the term covers a broad variety of interrelated theories. This ranges from commonsense views about how artifacts and social institutions are constructed by human agents, to controversial and radical theories about the status of knowledge, science and the nature of reality. Noretta Koertge holds that terms like "social construction" or "constructivism", "while they signal a certain sympathy towards nouveau ideas, have no precise referent" (Koertge 1996, p. 269; compare with Haslanger 2003).[65] In this article, we will be concerned with a robust version of SC as applied to scientific knowledge.[66] We will not be concerned with normative and evaluative connotations often associated with constructivist approaches (Hacking 1999, p. 6), as these will not be directly relevant for our epistemological perspective.

In traditional accounts of knowledge, it is taken for granted that "[u]nder the appropriate circumstances, our exposure to the evidence alone is capable of explaining why we believe what we believe" (Boghossian 2006, p. 22). By contrast, robust SC assumes, as Harry Collins put it in his seminal article, that "the natural world in no way constrains what is believed to be" (Collins 1981a, p. 54) and that "the natural world has a small or non-existent role in the construction of scientific knowledge" (Collins 1981b, p. 3). According to SC, the natural world has no significant bearing on theory choice in science. In explaining the beliefs of scientists, the proponent of SC will not appeal to their exposure to the relevant evidence, but rather to non-epistemic factors, such as sociological context, ideological influences and material conditions. This approach is embodied in the so-called 'Strong Programme' for the sociology of scientific knowledge, as developed by David Bloor. More specifically, Bloor's methodological "symmetry postulate" for the explanation of scientific beliefs states the following:

> [The sociology of scientific knowledge] would be symmetrical in its style of explanation. The same types of cause would explain say, true and false beliefs. (Bloor 1991, p. 7)

---

[65] Haslanger (2003, p. 301) remarks that "the variety of different uses of the term has made it increasingly difficult to determine what claim authors are using it to assert or deny".

[66] Rather than exposing particular entities as socially constructed against a background of natural facts, we see SC as being involved in the revision of central concepts like knowledge, truth and reality, coined "elevator terms" by Hacking (1999, p. 21).

> The symmetry postulate ... enjoins us to seek the same kind of causes for both true and false, [and] rational and irrational, beliefs .... (Bloor 1991, p. 175)[67]

This downplaying of epistemic considerations has important consequences for the question of theory change and paradigm shifts in science, as was pointed out by Golinski:

> Given sufficient creativity and resourcefulness on behalf of its defenders, any existing paradigm could be maintained indefinitely. [...] Why, then, should a paradigm ever change? (2005, p. 25)

Following Bloor, Golinski argues that, in order to explain a paradigm shifts, we should pay attention to the "social characteristics of the paradigm community and the balance of forces within it". As the natural world does not constrain theory choice or paradigm shifts, a scientific consensus merely reflects the contingent social and cultural preferences of the scientific community. Bloor does acknowledge that there "will be other types of causes apart from social ones" (Bloor 1991, p. 7), but presumably, the presence of such non-social causes does not account for the difference between true vs. false or rational vs. irrational beliefs. After all, the symmetry postulate invites us to look for the same type of causes in all cases.

It is often noted that the revisionist proposals of robust SC entail bold or prima facie absurd claims, and one might wonder whether any self-proclaimed proponent of SC sincerely endorses them (see also footnote 67). Are the critics of SC attacking a straw man? According to Kukla, the situation is more complicated, as the field of SC abounds in what he terms "reverse switcheroos":

> [Y]ou put forth a strong version of the hypothesis, and when it gets into trouble, you retreat to a weaker version, pretending that it was the weaker thesis that you had in mind all along. (Kukla 2000, p. x)

Thus, while many proponents of SC try to disown these radical interpretations when pressed on the issue, their writings often do imply them. In any case, as we will see, the more robust version of SC turns out to be most interesting from an epistemological perspective.

---

[67] Some adherents of the Strong Programme claim that the 'symmetry postulate' is a purely methodological precept that does not say anything about the *actual* causes of scientific belief, but this seems incoherent. If the natural world plays an important role in the construction of scientific knowledge after all, why would a good sociologist of science ignore this role in practice?

10.2.2 Standard objections to SC

Excellent philosophical critiques of robust SC have been formulated by Kitcher (1998), Koertge (2000) and Boghossian (2006). The standard problems associated with SC can be briefly rehearsed. First, critics claim that SC is incapable of explaining scientific success. If the natural world does not significantly affect the course of scientific development, how are we to explain the successful applications of scientific knowledge? Second, SC is haunted by the specter of Epimenedes, the ancient Cretan who proclaimed that "all Cretans are liars". If epistemic reasons never explain belief in a particular theory, what about the epistemic warrant for SC itself? The theory seems to be self-defeating.[68] Third, critics of SC maintain that the theory rests on implausible generalizations and misconceptions of certain post-positivist developments in philosophy of science (Koertge 2000). For example, the well-known problem of theory-ladenness of observation (TLO) is often misconstrued as lending support to the SC claim that nature does not in any way put constraints on theory formation. However, TLO does not apply uniformly and with equal force in every scientific context, and it is premature to conclude that it is fatal to the epistemic warrant of scientific claims. By applying TLO indiscriminately in every context, a subtle and valuable insight in philosophy of science has been lost.[69] The underdetermination of a theory by evidence (Quine 1953; Duhem 1954) is another thesis that is often confused with the stronger SC thesis of the radical contingency of scientific knowledge. Whereas W.V. Quine made a purely logical point about the possibility to preserve a hypothesis in the face of conflicting evidence, leaving room for pragmatic considerations in theory choice (elegance, explanatory power, etc.), defenders of SC think that underdetermination makes science completely impervious to evidence (see Hacking 1999, pp. 71-75).[70]

---

[68] David Bloor has tried to solve this problem by including the postulate of symmetry in his Strong Programme, stating that the principles of the Strong Programme should in principle be applicable to itself (Bloor 1991, p. 7). As it stands, however, Bloor has merely restated the problem, without discharging the objection of self-refutation.

[69] Once again, the danger of a vicious self-reference lurks, as Hacking observed: "To see all observations as equally loaded with theory is in itself to practice theory-laden observation, that is, observation loaded with a theory derived from Hanson the philosopher" (Hacking 1999, p. 200). In any case, scientists do not always have to decide in favour of one theory, but can also – and often do – suspend judgement, awaiting further empirical or conceptual developments (Slezak 1994, p. 281).

[70] Moreover, SC exaggerates the factors of individual epistemic distortions in scientific practice for example, in confirmation bias, prejudice, selective use of evidence or ideological bias. These distortions do occur in scientific practice, as in any other human activity, but in many cases the effects are sooner or later neutralized, because science is an eminently self-correcting social activity.

10.2.3 Science according to SC

In this paper, we will approach SC from a different perspective and pose the following question: supposing that scientific inquiry and practice were to function as SC claims they do, what exactly would they look like? Can we find a theory which *does* actually instantiate the epistemic predicaments as described by SC? In the following sections, we show that Freudian psychoanalysis, in virtue of well known and extensively documented defects in its epistemological foundations, methodology and conceptual apparatus, indeed exemplifies the general characteristics that SC (mistakenly) ascribes to serious science and scientific practice. Our analysis yields a new and independent strategy for challenging SC as a viable account of *bona fide* science, by offering a *reductio* of the SC program. In section 10.3 we briefly review a number of major conceptual and methodological flaws in Freudian psychoanalysis that have been extensively discussed elsewhere (Esterson 1993; Macmillan 1997; Cioffi 1998; Crews 1986). In section 10.4 we show how SC provides an excellent account of how psychoanalytic theories are developed and accepted. Philosophical worries about circularity in our argument will be addressed in section 10.5.1.

10.3 Major defects in Freudian psychoanalysis

10.3.1 The dynamic unconscious

Freudian psychoanalysis consists of both a complex dynamic psychology, a method for investigating the human mind and a framework for interpreting human behavior (in this paper we are not concerned with the effects of psychoanalytic therapy). Psychoanalytic doctrine revolves around the notion of the dynamic unconscious, an imperceptible realm of the human mind full of repressed mental contents, mostly sexual fantasies and desires stemming from early childhood.

The *explananda* of psychoanalytic interpretations cover a wide range of mental phenomena and their products, including neurotic symptoms, irrational thoughts and behavior, dream contents, slips of the tongue, works of art, social phenomena like

religions, etc.[71] The structure of a psychoanalytic interpretation typically takes the form "X is/counts as Y", where the Y-position is occupied by a psychoanalytic concept, and the X-position by an empirical description of observational source material, to which the psychoanalytic concept is assigned. According to the Freudian psychoanalyst, human thoughts and actions display certain anomalies, quirks and inconsistencies which betray the working of unconscious motives and fantasies. These psychological phenomena are supposed to reveal, through a sometimes complex chain of associations, hidden unconscious processes and meanings. Typical instances of psychoanalytic interpretations include: "your compulsive behavior *is in fact* an enactment of perverse childhood fantasies", "the stranger in the dream *represents* your father" or "your emotional insecurity *is a manifestation* of infantile castration anxiety".

Although we deliberately restrict our analysis to classical Freudian psychoanalysis, it should be noted that the divergent psychoanalytic schools that followed Freud's seminal theory have typically retained much of the problematic methodology and epistemology instituted by Freud: the existence of the psychodynamic unconscious, the notion of repression, the method of free association and symbolic interpretation as the gateway to the unconscious, etc. To the extent that these psychoanalytic schools have relied on the same defective aspects of the theory, our arguments apply with equal force (Macmillan 1997; Cioffi 1998).

10.3.2 Antagonistic subsystems

According to Freudian psychoanalysis, the human mind is the playground of a constant struggle between different mental subsystems. On the one hand, Freud often describes these mental systems as possessing intentional content, characterizing them in terms of personal-level concepts borrowed from folk psychology. In his early work, Freud analyzes the purposeful interaction between the unconscious and a mental entity called the censoring mechanism, which attempts to keep repressed mental contents from entering consciousness by means of distortion and disguise. This dynamic framework of antagonistic subsystems was later developed into the tripartite division of Ego, Super-Ego and Id. On the other hand, however, Freud also describes these different mental subsystems and their mutual interactions in purely mechanical terms, for example as being in the business of discharging and distributing a form of mental energy called

[71] In Buekens and Boudry (2010) we defend that, in virtue of its pseudohermeneutical character, psychoanalysis is capable of understanding virtually every human phenomenon.

libido. This tension between personal-level psychology and an impersonal libidinal economy has persisted throughout psychoanalytic literature (Gardner 2000).

Freud developed a complex set of technical concepts to describe dynamical interactions between different mental subsystems. For example, the concept of "negation" describes the transformation of an unconscious wish into a negative form (its denial) upon entering consciousness; "substitution" denotes the replacement of mental content by a substitute through a chain of unconscious association, in which the libido of the first content is transferred to the second, a process of displacement called "cathexis". "Condensation" denotes the bringing together of libidinal energy in one link connecting two associative chains. "Inversion" and "repression" similarly represent mental mechanisms for transforming mental contents through an invisible libidinal economy.

The vague and open-ended character of these concepts extends the possibility of drawing analytical inferences from observable source material (i.e. the explanandum in a psychoanalytic interpretation) to unconscious mental states and processes almost infinitely. In addition, Freud's psychology allows a single element in the empirical source material to have multiple unconscious determinants, a phenomenon often referred to as "overdetermination". These conceptual resources enable the psychoanalyst to make creative use of different sorts of symbolic associations, linguistic connections, double-entendres and homonyms, creating multiple layers of psychoanalytic interpretation. Importantly, what critics perceived as methodological extravagance was for the Freudians themselves inextricably connected with the very nature of the object of inquiry: we are dealing with a dynamic and intentional unconscious after all, which is in constant struggle with the mechanism of censorship, and which seeks ingenious and deceitful ways to provide an outlet for amassed libidinal energy.

Consider, for example, Freud's use of the concept of "inversion". Frank Cioffi convincingly argued that, although Freud theorized that neurosis develops when perverse desires remain unsatisfied, he did not recognize that patients who overtly indulged in their perverse desires but were neurotic nonetheless, constituted a refutation of his theory (Cioffi 1998, pp. 119-121). In the case of patients who suffered from neurosis without displaying overtly perverse behavior, Freud explained the symptoms as an outlet for libidinal energy amassed in response to repressed perversities. In the case of overtly perverse neurotic patients, Freud maintained that the symptoms expressed a repressed *aversion* against their indulging in perversities, and constituted a case of "inversion".

Another example of typical psychoanalytic reasoning is the way apparent falsifications of the Oedipus complex were handled. For Freud, the affectionate behavior of little boys towards their mother was a manifestation of incestuous desire. On the other hand, if a boy showed affection towards the father and was cold or hostile towards

the mother, as in the famous case of Little Hans, Freud explained the behavior as a reaction formation against the actual incestuous desires, which he thought were partially repressed (Freud 1955a; Van Rillaer 1980, pp. 141-155). Thus, the concepts of inversion and reaction formation allowed Freud to account for virtually every observation.

10.3.3 Conspiracy thinking

The Freudian unconscious is an entity that actively *resists* interpretation, and that will always try to deceive us in unexpected and cunning ways (Gellner 1985).[72] Thus, when Freud was unable to find traces of a pathological complex or unconscious desire to account for a patient's behavior, he was undeterred and treated this as a token of unconscious resistance. The more the material offered by a patient resisted interpretation, the more it counted *in favor* of the theory. This characteristic pattern of reasoning in psychoanalysis bears a striking resemblance to conspiracy theorizing (Farrell 1996). For example, consistent with his account of the unconscious, Freud believed that his patients (and his critics) harbored a secret and unconscious wish to see his theories and interpretations proven wrong, and so never to see their own unconscious desires exposed. For instance, one of Freud's patients dreamt that she had to spend her holidays with her despised mother-in-law. This seemed to belie Freud's claim that *every* dream is an unconscious wish-fulfillment, but within the framework of psychoanalytic thinking it could be turned into a confirmation of the theory. As Freud himself explained,

> The dream showed that I was wrong. *Thus it was her wish that I might be wrong, and her dream showed that wish fulfilled*. [italics in original] (Freud 1953a, p. 151)

Freud argued that "these dreams appear regularly in the course of my treatments when a patient is in a state of resistance to me" and he predicted that the same would happen to his readers (Freud 1953a, pp. 157-158).[73] Indeed, Freud and his followers became infamous for explaining away criticism from their opponents as tokens of unconscious resistance to the theory, thus further attesting to the truth of psychoanalysis:

---

[72] Or more precisely, in both simple and far-fetched ways. The degree of complexity in psychoanalytic interpretations varies greatly, from very straightforward symbols for genitals on the basis of superficial visual similarities, to intricate, multi-layered and multi-lingual analyses.

[73] The quotation continues: "Indeed, it is to be expected that the same thing will happen to some of the readers of the present book: they will be quite ready to have one of their wishes frustrated in a dream if only their wish that I may be wrong can be fulfilled."

> They [the critics] are therefore bound to call up the same resistance in him as in our patients; and that resistance finds it easy to disguise itself as an intellectual rejection and to bring up arguments like those which we ward off in our patients by means of the fundamental rule of psycho-analysis. (Freud 1957, p. 39)

As we argued in Chapter 5 and in section 6.4.3, such moves are not merely immunizing gambits which can be neatly disentangled from the theory, but are instead perfectly legitimate, explanatory moves within the psychoanalytic framework, and instantly recognizable as genuine psychoanalytic interpretations (Boudry and Braeckman 2010). This pattern of reasoning, which bears a striking resemblance to conspiracy thinking, is pervasive throughout psychoanalytic literature, and it follows directly from the characterization of the unconscious as an intentional and deceitful mental entity.

10.3.4 The quantitative factor

Freud treated a patient's explicit denial of his hypotheses – for example in his use of the concept of penis envy – as yet further confirmation of his claims, but that didn't mean that he was prepared to accept cases where patients readily *accepted* his interpretations as refuting his theory. Indeed, if the patient's dreams seemed to *confirm* Freud's notions, they could be explained as an example of "compliance towards the analyst", and thus again be relegated to unconscious motives (Freud 1961, p. 117) Thus, neither the denial or the (belated) acceptance of an interpretation posed a problem from the perspective of Freudian theory.

The difference between both forms of behavior could be explained by the analyst as the result of unobservable variations in the strength of unconscious resistance on the one hand and the intensity of libidinal energy on the other hand. This "quantitative factor" in the patient's mental economy had the effect of forestalling the falsification of what initially looked like testable predictions. As Freud himself made clear in a remarkably candid passage, it could always be invoked *post factum* to account for the unexpected presence or absence of any given symptom:

> We cannot measure the amount of libido essential to produce pathological effects. We can only postulate it after the effects of the illness have manifested themselves. (Freud 1924, p. 119)

10.3.5 Conceptual double lives

Another important feature of Freudian psychoanalysis, which further contributes to its epistemic predicament, is that its concepts lead what may be called a "double life"

(Cioffi 1998, p. 118): sometimes they seem to be semantically-rich and clearly-delineated, but on other occasions they are inflated so as to become almost indefinite and meaningless. This conceptual double life makes central psychoanalytic concepts virtually immune to refutation. Cioffi (1998, p. 15) mentions Freud's "disingenuous alternation" in the scope of the libido-concept, in which he switches between an explicitly sexual libido on the one hand and a general kind of love and affection on the other hand.[74]

As pointed out earlier, many concepts in Freudian psychoanalysis (e.g. repression, projection, wish-fulfillment) alternate between personal-level psychology and blind libidinal economy, a form of equivocation that makes psychoanalytic interpretations particularly ambiguous and elusive (Gardner 2000). Esterson (1993, p. 230) has concluded that the functions of the central concepts in Freud's ego-psychology (Ich, Über-Ich, Es) "are so imprecisely delineated that they can be employed in almost arbitrary fashion to provide support for virtually any theoretical formulation." Elsewhere (Buekens and Boudry 2010) we have argued that the extension of many psychoanalytic terms is not fixed until applied in interpretations, which goes some way to explaining their open and indeterminate meaning (Borch-Jacobsen and Shamdasani 2006).

10.3.6 Inseparability of theory, methodology and practice

A number of critics of psychoanalysis (Grünbaum 1984, 2008; Eagle 1988; 1996, 2002) have insisted on a clear distinction between the theory-as-such and the tendency of its advocates to use immunizing gambits and other methodological tricks in the face of falsifying material. These authors maintain that the theoretical problems we reviewed above have nothing to do with psychoanalysis properly speaking, but should be laid at the door of individual analysts.

In this paper, however, we follow critics like Cioffi (1998), Crews (1986) and Macmillan (1997), who have meticulously demonstrated that, in practice, it is all but impossible to indicate a point where the orthodox version of the theory ends and where immunizing strategies and methodological obfuscations begin (Boudry and Braeckman 2010). This is because what Grünbaum and Erwin designate as "dubious" methodological

[74] This semantically double life of concepts like 'libido' fulfilled two conflicting demands: on the one hand, only the narrow, carnal interpretation could explain why fathers threatened their sons with penile amputation and why libidinal drives were so inadmissible for our moral sensibilities that they had to be repressed; on the other hand, only the wider interpretation allowed Freud to maintain that beneath the surface of each and every symptom the libidinal factor is lurking.

practices and immunizing tactic emerge *directly* from theory-internal epistemic properties. As Cioffi wrote:

> (W)e have no canonical statement of the theory: no agreement on what constitutes modifications of the theory rather than *post hoc* elucidations of it. [...] What we have in Freudian theory is a combination of epistemically ambiguous utterances with methodologically suspect practices. (Cioffi 1998, p. 300)

The way in which the concept of *resistance* has been put to use by Freud and his acolytes, for example, has been rightly dismissed by critics as a specimen of heads-I-win-tails-you-lose reasoning. Nevertheless, it proves difficult to disentangle such fallacious reasoning from psychoanalytic theory itself, because it is effectively supported by the way the unconscious is conceptualized in Freudian theory. If Freud's model of the human mind is correct, and if the unconscious really is some sort of trickster in disguise, then *indeed* it becomes natural to label counter-arguments and criticisms as manifestations of unconscious resistance to psychoanalytic 'truths' and 'interpretations'.

10.3.7 A cumulative effect

The remarkably versatile and multi-directional methodology of Freudian psychoanalysis (Timpanaro 1976; Macmillan 1997; Cioffi 1998), which has long been noted by its critics, is the natural outcome of dividing the mind into intentional and antagonistic substructures. As we have seen, Freud's particular dynamic conception of the human mind creates an abundance of inferential possibilities when applied in hermeneutical practice, enabling the analyst to turn any psychological phenomenon into the symptomatic outcome of a hidden psychodynamic conflict. In addition, the conceptual equivocations in the theory render Freud's hermeneutic machinery even more versatile.[75] As Frederick Crews wrote:

[75] The therapeutic methods of free association and transference analysis, although demonstrably unavailing for probing another person's mind (Grünbaum 1984) and based on placebo effects (Jopling 2008), are not discussed here as one of the central methodological pitfalls of psychoanalysis, because these methods are not employed in psychoanalytic interpretation of human phenomena like works of art, literary texts or religion, and thus do not constitute the *central* epistemic problem of psychoanalysis.

> Each posited subset of 'the unconscious' permits another strand of contrary motivation to be added to the already tangled explanatory skein, leaving us, if we are sufficiently gullible, so impressed by the psychoanalytic interpreter's diagnostic acumen that we think we are witnessing elegant and validated feats of deduction instead of being told a self-serving detective story in which the mystery itself [...] is an artefact of question-begging manoeuvres. (Crews 2006, p. 56)

The cumulative effect of these methodological and conceptual problems is that, if the psychoanalytic unconscious exists, it is deprived of any capacity to put epistemic constraints on theoretical claims and psychoanalytic interpretations. Indeed, any guarantee for interpretive congruency in Freudian psychoanalysis is frustrated by the methodological flexibility and conceptual deficiencies inherent within the theory (Van Rillaer 1980, pp. 87-92; Esterson 1993, p. 242; Macmillan 1997). This was the verdict reached by Malcolm Macmillan in his *Freud Evaluated*:

> [T]he so-called discoveries are dependent upon methods of enquiry and interpretation so defective that even practitioners trained in their use are unable to reach vaguely congruent conclusions about such things as the interpretation of a dream or symptom [...] (Macmillan 1997, p. 516)

Indeed, the internal feuds and factions characterizing post-Freudian psychoanalysis bear witness to the epistemological problems described by critics as Macmillan and Cioffi (Borch-Jacobsen and Shamdasani 2006, 2008). Already in 1962, the psychoanalyst Judd Marmor observed (with understandable disquietude) that, by means of the psychoanalytic method, confirmations could be found as easily for Freud's Oedipus complex, as for Adler's inferiority complex, or for Lacan's symbolic Father, or for Jung's anima and persona:

> [D]ependent on the view of the analysts the patients of each school generate precisely those data that support the theories and interpretations of their analysts. (Marmor 1962, p. 289)

Returning to the main issue, we will now show that these findings resonate with the SC tenet that "what is believed to be" (Collins 1981a, p. 54) is in no way constrained by "nature" (in this case, what is going on in our minds).

10.4 A constructivist redescription of psychoanalysis

10.4.1 SC and Freudian psychoanalysis: no constraints on evidence

Our central claim is that Freudian psychoanalysis perfectly illustrates the epistemological predicament described by proponents of SC. Indeed, the critique of psychoanalysis presented in the previous sections resonates with the "debunking project" that is central to the program of SC. As Boghossian aptly noted:

> A social construction claim is interesting only insofar as it purports to expose construction where none had been suspected, where something constitutively social had come to masquerade as natural. (Boghossian 2006, p. 18)

This is exactly what critics of Freudian psychoanalysis have been engaged in all along: exposing as "constructions" what Freudian theorists presented as genuine natural facts out there, waiting to be discovered.[76] The received critical view of Freudian theory could be rephrased as follows: what psychoanalysts present as 'insights' and 'findings' are merely artifacts of the theory itself and of its deficient methodology; only those who have already embraced Freudian theory 'see' the described psychoanalytic phenomena (e.g. the phallic meaning of a dream symbol; the child's erotic pleasure in thumb sucking).

The epistemic predicament of Freudian psychoanalysis can now be redescribed within the framework of SC. Take the critical observation that, in Freudian hermeneutics, any guarantee for interpretive congruency is frustrated by the methodological flexibility and conceptual versatility of the system. As Harry Collins would have it, "what is believed to be" in psychoanalytic hermeneutics is "in no way constrained" (1981a, p. 54) by what is actually going on in our minds. This perfectly illustrates two key tenets of SC, viz. that evidential considerations play an insignificant role in theoretical developments, and that the resolution of theoretical debates and conflicts in psychoanalysis is not driven by epistemic considerations.

The critical observation that psychoanalysts always succeed – sometimes with considerable ingenuity – in molding seemingly adverse evidence into accordance with their theory, illustrates another tenet of SC: the ability to preserve the theory in the face of recalcitrant material. As Golinski put it, "given sufficient creativity and

[76] After all, to be meaningful and coherent, Freudian theory has to presuppose that there is an independent mental reality and that the analytic method yields knowledge about it, this in spite of later attempts at a hermeneutical (Ricœur 1970; Habermas and Shapiro 1981) or constructivist reconstruction of psychoanalysis (Spence 1982; Schafer 1992; Moore 1999).

resourcefulness on behalf of its defenders, the existing paradigm could be maintained indefinitely" (Golinski 2005, p. 25). Most philosophers of science would agree that this is a rather implausible claim when applied to *bona fide* science, but in the case of psychoanalysis we claim that the description is perfectly accurate. Fully in line with Judd Marmor's observation, critics of psychoanalysis have often noted that there is no rational method to resolve the persistent disputes between the followers of Sigmund Freud, Otto Rank, Alfred Adler, Jacques Lacan, Daniel Winnicott, Anna Freud, Melanie Klein and many others. For example, when his disciple Otto Rank introduced the concept of *birth trauma* as a pre-oedipal source of neurosis (a Freudian heresy), Freud could only reply that he was unable to confirm the remnants of such a trauma in his clinical work, and he urged a reinterpretation of the material in terms of oedipal desires. Both schools have since then continued to find 'confirmations' for their own theoretical framework, and have failed to 'see' the mental phenomena described by their adversaries (Cioffi 1998, pp. 17-19).

10.4.2 Historiography and sociology of psychoanalysis

For the historiographer of psychoanalysis, David Bloor's "symmetry postulate" seems an appropriate methodological tool. Since the theoretical choice for Oedipus complex, death wish, inferiority complex or birth trauma does not depend on the nature of empirical evidence, one must resort to deeply non-epistemic factors to explain theoretical disputes, developments, and schisms in the history of psychoanalysis (Borch-Jacobsen and Shamdasani 2006). As the Strong Programme recommends, the historiographer of sociologist or psychoanalysis should not be concerned with the "truth" of such or such psychoanalytic doctrine, and in all cases he should look for the "same kind of causes", viz. ideological background, social networks, personal animosity, and cultural context.

To give just one example, in the second half of the 20th century the doctrine of universal *penis envy* in women was progressively abandoned in many psychoanalytic schools. It is quite implausible that this theoretical change was driven by evidential considerations, for Freudians theorists had 'confirmed' the doctrine countless times in the past and touted it as one of the cornerstones of analytic theory (Cioffi 1998, pp. 27-28). Moreover, the method of investigation remained the same, so what could account for this theoretical development? Instead of looking for epistemic reasons, historians of psychiatry and psychoanalysis would be better advised to investigate the changing social and cultural sensibilities of the time, which began to regard the concept as patriarchal and misogynist. The development of the psychoanalytic concepts of *breast envy* (Melanie Klein) and *womb and vagina envy* (Karen Horney) to compensate for this "phallocentrism" on Freud's part must be seen in the same light (Sayers 1987). Indeed,

19th century preconceptions about female submissiveness and inferiority constitute a good candidate for explaining the genesis of the concept of penis envy in the first place. This focus on sociological and ideological causes is precisely what sociologists of scientific knowledge like Bloor would recommend. The idea that the concept of universal penis envy was gaining acceptance in Freud's time because "Nature had spoken" and psychoanalysts had paid heed to Her is quite implausible in light of a critical assessment of Freud's theory and methodology. As David Bloor would have it, "[w]hat function does truth, or talk of truth, play in all this? It is difficult to see that much would be lost by its absence" (Bloor 1991, p. 40). In the absence of epistemic constraints on theory change, the explanatory *vacuum* is filled in with various sociological, ideological and psychological factors.

10.4.3 The construction of meaning and the archaeological metaphor

It is interesting to note that Freud himself acknowledged that the work of the analyst closely resembles the practice of (re)construction. About his analysis of the Wolf Man, Freud wrote:

> All I mean to say is this: scenes, like this one in my present patient's case, which date from such an early period and exhibit such a content, and which further lay claim to such an extraordinary significance for the history of the case, are as a rule not reproduced as recollections, but have to be divined – constructed – gradually and laboriously from an aggregate of indications. (Freud 1955b, p. 51)

Freud also compared the psychoanalyst to the archaeologist, carefully excavating the buried remnants of the past and uncovering layer after layer of unconscious meaning (Freud 1953c, 1955a). But of course, the context of psychoanalytic inquiry does not resemble the direct accessibility of archaeological excavations at all, as Freud could only tenuously *infer* the existence of alleged unconscious phenomena on the basis of certain cues (dreams, associations, patterns of behavior) (Moore 1999). In psychoanalytic therapy, the archaeology metaphor reinforces the misconception that the psychoanalyst merely brings to the surface that which was present all along in the patient's mind.[77] Freud used other realism-inducing metaphors to describe the inquiries of the analyst. For example, he likened the manifest dream content to a "rebus" or "picture puzzle", in which symbols have to be deciphered to reveal hidden meanings

[77] In that sense, the hypothesis of a dynamic unconscious full of forbidden wishes and desires, combined with the concepts of repression and denial, creates favorable psychological conditions for the creation of false insight and successful suggestion (Jopling 2008).

(Freud 1953a, pp. 277-278). These powerful metaphors conveyed the image of an inquiry into an objective mental reality out there, and they have paved the way for a misinterpretation of Freudian social constructions as natural, empirically-detectable facts.

10.4.4 An independent argument against SC

The first part of our central argument is that the framework of SC offers an accurate account of the epistemic predicament of psychoanalysis, in particular the classical Freudian version. We will now argue that that SC's applicability to psychoanalysis yields that it must embody a bad account of how *bona fide* science works.

To secure this conclusion, we need to show that a successful description of Freudian psychoanalysis in terms of a social-constructivist account succeeds *in virtue* of methodological and conceptual problems that are *not* manifest, or at least not to the same extent present in *bona fide* science. There is some circumstantial evidence for this claim: many of the features we discussed are widely recognized as *distinctively* psychoanalytic, even by Freud's contemporaries (Borch-Jacobsen and Shamdasani 2006), and they were criticized on independent factual and theoretical grounds (Esterson 1993; Macmillan 1997; Cioffi 1998). Indeed, for many scientists and philosophers, they served as the basis to question the theory's scientific credentials (Derksen 1993; Cioffi 1998; Popper 2002).

The central question remains, however, what it takes for a theory to exemplify the epistemic situation as described by SC. The proponent of SC can retort that our argument is superfluous: "Even if so-called bona fide science doesn't exhibit the *specific* problems you correctly identified in psychoanalysis, it is no less socially constructed". Our opponent may maintain that the viability of a constructivist redescription of Freudian psychoanalysis is *not* due to the characteristic problems we mentioned, but due to *general* features that are also manifest in *bona fide* science. In other words, the specific problems we discussed are superfluous, and *bona fide* science is just as much "socially constructed".

However, this would entail that the various theoretical complications and loopholes of Freudian psychoanalysis, which have fascinated critics and defenders alike, are *epistemically gratuitous*. The proponent of SC is then committed to denying that these characteristic problems are responsible for what critics have condemned as the epistemic vacuousness of Freudian psychoanalysis, and this is quite implausible. First, it is simply incoherent to claim that the conceptual and methodological flaws have no epistemic consequences, *i.e.* are epistemically inert. We have documented in detail that their cumulative effect consists of *reducing* epistemic constraints on hermeneutic practice and theory formation, to produce spurious evidence about an object that is

actually an artifact of the theory (the dynamic unconscious), and to distract attention from this very process.[78] Even defenders of psychoanalysis have occasionally acknowledged that these specific problems compromise the theory's epistemic status and its aspiration to be recognized as a genuine science (Eagle 1993). Second, there are good reasons to assume that in other respects they *hindered* the success of the theory, since they undeniably compromised its credibility in the eyes of countless scientists and philosophers.

Note that our dialectical position is not committed to the implausible claim that the methodological and conceptual problems that plague Freudian psychoanalysis are entirely absent from *bona fide* science. Imre Lakatos maintained that every scientific theory builds a "protective belt" (Lakatos and Musgrave 1970) of auxiliary hypotheses around its core claims. Likewise, the theory-ladenness of observation is a genuine problem that affects *bona fide* science as well. But this is a far cry from the stronger picture outlined by SC. Every theory can withstand a certain amount of anomalies using face-saving auxiliaries, but that does not mean that they are all on equal footing. The strength and imperviousness of the protective belt around a core theory is a matter of degree, not of absolute difference. One does not need to be a naive falsificationist to recognize that, in *bona fide* science at least, a sufficient number of empirical anomalies may indeed threaten a theory, especially if far-fetched modifications are in order to get the theory in accordance with reality. In Freudian psychoanalysis, by contrast, the particular methodological problems and conceptual resources we discussed ensure that no amount of empirical observations can seriously endanger central psychoanalytic propositions. No matter what nature says, psychoanalysts always hear the same voices. Because of its specific epistemic predicament, Freudian psychoanalysis, in contrast with *bona fide* science, exemplifies the picture of science SC defends and promotes. In that sense, we have uncovered an example of what science *would* look like if it *were* to function as SC claims it does.

[78] See for example Freud's explicitly empiricist and objectivist rhetoric (Buekens and Boudry 2010).

10.5 Discussion & objections

10.5.1 The circularity objection and the demarcation problem

"Your argument begs the question," the critic may retort. "On the one hand you claim that Freudian psychoanalysis is a pseudoscience, because it displays the features described by SC. On the other hand, you hold that SC is a bad account of science because it correctly describes a pseudoscience (i.e. Freudian psychoanalysis). This is a circular argument."

This objection allows us to clarify the overall structure of our argument. First, our empirical argument against SC as an account of scientific practice is independent from traditional conceptual and philosophical objections (they were briefly rehearsed at the beginning of this article). Second, we do not reject Freudian psychoanalysis *just because* it fits the SC framework. We do *not* propose our analysis as a new demarcation criterion, and we do not think that our framework is capable of capturing all the features that characterize pseudosciences. Third, it may well be that some of the methodological and conceptual problems we have discussed figure in proposed solutions to the demarcation problem (Derksen 1993), but it does not follow that we thereby also *endorse* that demarcation criterion, nor that our argument depends on the rejection of Freudian psychoanalysis as pseudoscience under that supposed demarcation criterion. The only relevant premises in our argument are (i) the *presence* of these characteristics and defects[79] in Freudian psychoanalysis, and (ii) their *cumulative effect* on the epistemic status of the theory. It may well be possible – although this was not a line we intended to take – to turn the case of Freudian psychoanalysis against SC on the *presupposition* that there is a correct demarcation criterion according to which psychoanalysis should be dismissed as a pseudoscience. However, this will not impress the defender of SC, since the possibility of a viable demarcation criterion separating science from pseudoscience is precisely what is problematic from the point of view of SC. In this article we tried to give an independent argument against SC, not an argument that simply presupposes the very characterization of science SC objects to.

[79] The fact that we talk about 'defects' may seem tendentious. However, our argument does not depend on this terminology but stands in its own right. In any case, with this choice of term, we are not suggesting that we endorse a demarcation criterion on the basis of these characteristics.

10.5.2 Social constructivism and psychoanalysis as natural allies

It is interesting to note that some contemporary psychoanalysts have themselves embraced some version of SC (Spence 1982; Stern 1992; Schafer 1992; Moore 1999). In this constructivist school, the empirical ambitions of Freud are largely abandoned, and it is argued that "narrative truth" or "interpretive construction" is everything a psychological theory (or indeed any theory) can offer. Eagle (2003) has conclusively argued that, by renouncing any claim at empirical insight, this postmodern version of psychoanalysis inevitably leads to relativism. However, one might wonder why SC has attracted the attention of so many psychoanalysts. It seems plausible that constructivist readings of psychoanalysis help neutralizing scientific and philosophical objections to Freud's theory by extrapolating the epistemological problems from which psychoanalysis suffers to science in general. In that very specific sense, psychoanalysts find a natural ally in constructivists, some of whom have been found to openly embrace relativism (Bloor 2007).

10.5.3 Institutional facts

In Buekens & Boudry (2010) we develop an account of psychoanalysis as a system of "institutional facts" along the lines of John Searle's theory of social institutions (Searle 1995; Lagerspetz 2006) and propose this analysis as the best explanation of what really happens in the hermeneutic practice of psychoanalysis. The upshot of that analysis is that psychoanalysis is based on a confusion between natural facts and institutional facts: what Freud thought of as descriptions of natural facts are, to a large extent, declaratives that create and, when accepted by others, sustain institutional facts. In contrast with global SC, the reconstruction in Buekens & Boudry (2010) assumes a firm distinction between natural facts, which exist independently of human intentionality, and institutional facts. This differs from the implausible anti-realist credo of SC that all scientific facts are socially constructed. The argument in Buekens & Boudry (2010) resonates with the conclusion in this paper that Freudian theories and interpretations are "social constructions" rather than verifiable natural facts about the vicissitudes of the human mind.

10.5.4 Conclusion

SC approaches to science have often been dismissed as inaccurate accounts of scientific knowledge. Our aim in this paper was to take the claims of radical SC seriously and to

find out whether we can uncover theories which *do* instantiate the epistemic predicament as described by SC. As we have shown, (Freudian) psychoanalysis fits the bill, in virtue of its well-known conceptual problems, its peculiar epistemic structure and its methodological flexibility. The combination of what psychoanalysis *pretends* to be (i.e. a science), and the way the theory *really* functions (i.e. a system that produces arbitrary constructions), makes a redescription in the framework of SC particularly apt. By showing in some detail what it takes for a theory to 'create' its own object, our analysis yields an independent argument against SC as a global account of science.

Part 3 – Epidemiology of Weird Beliefs

Chapter 11.
How Convenient! The Epistemic Rationale of Self-validating Belief Systems

When men wish to construct or support a theory, how they torture facts into their service! – Charles Mackay (1974 [1841], p. 459)

The trouble with having an open mind, of course, is that people will insist on coming along and trying to put things in it. – Terry Pratchett

Abstract.[80] This paper offers an epistemological discussion of self-validating belief systems and the recurrence of "epistemic defense mechanisms" and "immunizing strategies" across widely different domains of knowledge. We challenge the idea that typical "weird" belief systems are inherently fragile, and we argue that, instead, they exhibit a surprising degree of resilience in the face of adverse evidence and criticism. Borrowing from the psychological research on belief perseverance, rationalization and motivated reasoning, we argue that the human mind is particularly susceptible to belief systems that are structurally self-validating. On this cognitive-psychological basis, we construct an epidemiology of beliefs, arguing that the apparent convenience of escape clauses and other defensive 'tactics' used by believers may well derive not from

[80] This chapter is based on a forthcoming paper in *Philosophical Psychology* (Boudry and Braeckman forthcoming). Parts of it were presented at the Fourth Conference of the Dutch-Flemish Association for Analytic Philosophy at the University of Leuven (2010).

conscious deliberation on their part, but from more subtle mechanisms of cultural selection.

11.1 Introduction

The satirist H. L. Mencken once wrote that: "the most common of all follies is to believe passionately in the palpably not true. It is the chief occupation of mankind". Mencken was clearly engaging in hyperbole, but he has some statistics on his side (Irwin 2009; Hines 2003): according to a 1999 Gallup poll, 45% of all Americans believe that the earth is less than 10,000 years old and that all living species were created by God in their present form. A vast number of people believe that the murder of John F. Kennedy was a conspiracy, that the moon landing in 1969 never happened, or that the Bush administration was involved in the 9/11 attacks (a conspiracy theory that is also popular in Europe). The number of people who put faith in bogus medical treatments, or even the number of different bogus therapies, is alarmingly high. A 2005 Gallup poll conducted in the United States, Canada, and Great Britain found that ca. 25% of all persons questioned believe in astrology and around 20% believe that extraterrestrials have visited the earth recently. Even higher numbers were found for the belief in haunted houses and communication with the dead. Similarly, all over the world, sects and religious cults continue to attract followers.

In this paper, we discuss belief systems across widely different domains and focus on their self-validating nature as part of an explanation of their wide appeal and enduring popularity. After an introductory section on the received view about weird beliefs and irrationality (section 11.2), we review the relevant literature on belief perseverance, cognitive dissonance and motivated reasoning (section 11.3). We proceed by analyzing the recurrence of certain epistemic defense mechanisms and immunizing strategies in many of the most widespread "irrational" belief systems. By demonstrating that these belief structures exhibit a remarkable degree of *natural resistance or resilience* in the face of disconfirmation and criticism, we challenge the received idea that they are intrinsically vulnerable or fragile (sections 11.4 and 11.5). In line with the emerging literature on *cognitive constraints* on the formation and distribution of beliefs (Boyer 1994; Atran 2002; Barrett 2007), we explain the success of these belief systems in terms of ordinary modes of human cognition, and within the framework of an epidemiology of beliefs (section 11.6) (Boyer 1994; Sperber 1996; Lienard and Boyer 2006). Our epistemological approach is informed by, but not reducible to, the cognitive research on

motivated reasoning and cognitive dissonance, and makes for a level of explanation in its own right. Finally, we show that the epistemic "engineering" of certain belief systems may well derive not from conscious deliberation on the part of believers, but from more subtle mechanisms of cultural selection (section 11.7).

11.2 Pseudoscience and the Paranormal

11.2.1 The tenacity of weird beliefs

The tenacity of belief systems that are highly implausible, or whose content contradicts well-established scientific knowledge, has often exasperated skeptical scientists and philosophers alike. Dyed-in-the-wool skeptics, however, have long come to realize that firm believers are very difficult to convince with evidence and rational arguments. To believe otherwise is to commit the "rationalistic fallacy" (Pigliucci 2002, pp. 234-236), the idea that all one has to do to make people abandon their pseudoscientific beliefs is to explain things a little more clearly. Experience with debating "believers" has conclusively shattered this illusion, and the phenomenon of belief perseverance in the face of disconfirming evidence has been well researched in psychology (Anderson, Lepper et al. 1980; Carretta and Moreland 1982). But why is it that rational arguments are generally to little avail in the domain of pseudoscience and the paranormal?

Among skeptics engaged in the scientific evaluation of these claims, there is a widespread conviction that believers in the paranormal and pseudoscience are obviously irrational, if not plain stupid. Carl Sagan once noted that, to his regret, many of his skeptical colleagues find that "those [...] who believe in all these stupid doctrines are morons" (Sagan 1995, pp. 29-30). For example, Richard Dawkins, annoyed by the continuing opposition to evolution by religious fundamentalists, once remarked that "[i]t is absolutely safe to say that if you meet somebody who claims not to believe in evolution, that person is ignorant, stupid or insane" (Dawkins 1989). After all, so the argument goes, pseudoscientists are not capable of rational argumentation, they simply ignore or fail to understand evidence that does not fit their dogmatic ideas, and they keep committing fallacies of reasoning that have long been laid to rest (e.g. Godfrey 1979; Estling 2005). The sentiment that some belief systems are so obviously wrong or absurd that no sane person would ever come to accept them, has also pervaded academic research on for example superstition and religious cults. In an attempt to explain why seemingly normal people succumb to pseudoscience and cults, some

researchers have argued that such belief commitments can only be sustained by elaborate organizational devices, psychological indoctrination and other techniques for distorting normal reasoning functions. Other researchers have speculated on the existence of a special mode of "magical thinking" (Kurtz 1986), which is disconnected from normal reasoning faculties, or which constitutes a pre-rational stage in the development of the human mind (for an overview, see Atran 2002, p. 141).

Modern research in cognitive psychology belies these speculations and suggests that many of our false beliefs stem from mundane reasoning errors and biases which are inherent in the way the human mind processes information (Nisbett and Ross 1980; Gilovich 1991). For example, instead of attributing superstition to a special mode of "magical thinking", researchers have documented the role of our flawed understanding of randomness and coincidence in the formation and persistence of superstitious beliefs (Gilovich 1991; Vyse 1997; Talmont-Kaminski 2008). Thus, the persistence of superstition emerges as a side-effect of our natural ability for pattern recognition and causal inference making (e.g. Foster and Kokko 2009). In evolutionary terms, the cost of overlooking causal relations is higher than that of occasional false positives, which explains the high sensitivity of our cognitive faculties to correlations in the environment. Nesse (2001) has termed this the "smoking detector principle", because smoking alarms are designed to err on the side of caution for the same reason. In a similar vein, researchers are starting to approach the human penchant for magic and the paranormal as a natural by-product of the way our brain employs ontological categories (e.g. physical, biological, mental) to make sense of the world (Lindeman and Aarnio 2007). Paranormal beliefs are then regarded as the result of a confusion between the core attributes of these ontological categories. In general, cognitive psychologists have argued that local irrationality emerges as the inevitable by-product of our brain's ability to efficiently gather and process information (Talmont-Kaminski 2008, 2009).

As we will see, evidence against the sheer "irrationality" or "stupidity" of believers has also come from sociological studies of cults and sects, which have shown that even the most outlandish belief systems have a form of internal logic and coherence, and that believers are often not the simple-minded fanatics they are taken to be (e.g. Dein 2001; Tumminia 1998; Lukes 2007). Moreover, cults and sects who did rely on crude brainwashing and mind control devices to recruit new members have generally met with little success (Mercier in press; Streatfeild 2007). It seems that attempting to impair people's normal reasoning abilities is an inefficient way of winning new converts for a "weird" belief system.

In this paper, we want to focus on one particular misconception underlying the sentiments of skeptical outsiders of weird belief systems. Echoing the remarks of H. L. Mencken, people often express puzzlement about what they perceive as the "palpable" falsity of these belief systems: "How could anyone in his right mind believe such nonsense?" In other words, people assume that pseudoscientific belief systems are

"inherently fragile" (Snow and Machalek 1982) – that believers are constantly faced with overwhelming adverse evidence, which would compel any reasonable person to immediately give up such beliefs. We argue that this assumption is largely misguided, and that a closer look at the structure of weird belief systems reveals that believers are in fact well-prepared to withstand such difficulties. More specifically, once believers accept the central premises of the belief system in question, they have ample explanatory resources at their disposal to reason their way through apparent disconfirmations, and to withstand criticism from outsiders. Before delving into this discussion, however, we need to cover some psychological ground about irrationality, motivated reasoning and belief perseverance.

11.2.2 The case of creationism and "blind faith"

Almost half of the population in the United States believe that all living species were created in their present form by God, and that the earth is less than 10,000 years old. Adhering to a literalist reading of Scripture, these people also believe in the Garden of Eden, Noah's Ark and the Great Flood. It is difficult to find a belief that flies more in the face of modern science. Consequently, it is tempting to argue that, since the creationist belief system is so "palpably not true", those who endorse it surely must be completely impervious to rational arguments and evidence.

On the other hand, it is surprising to note how often creationists go to considerable lengths to "massage" scientific evidence into their preconceived Biblical framework. Moreover, they have – often with amusing results – taken great pains to overcome exegetical inconsistencies and other difficulties in their belief systems to answer questions like: How did Noah manage to get all the animals into the Ark? If Adam and Eve were the only humans around, did Cain marry his own sister? What did T. Rex use his claws and teeth for in the Garden of Eden, before death and carnivore diets entered the world? Why does the book of Genesis contain two apparently inconsistent accounts of creation?

Of course, the arguments invented by creationists to address these problems are selective, distorted and heavily biased (Kitcher 1982; Pennock 1999; Pigliucci 2002), but there is a more interesting point to be made: for people who are blindly and "irrationally" committed to religious faith, many creationists have a surprising concern with inconsistencies and adverse evidence. Why do they not just ignore anything that does not fit their ideas, instead of bothering themselves with elaborate rationalizations and *ad hoc* explanations? One might object that creationists merely need these elaborate arguments to attract new converts, but that explanation pushes the question ahead. If religious fundamentalism is simply about committing oneself blindly to the truth of a

holy book, why do people need arguments to be convinced in the first place? In other words, why are creationists not *more* irrational?

11.3 The Psychology of Belief Perseverance

11.3.1 Motivated reasoning and confirmation bias

Psychologists have found some truth in the received psychological wisdom about irrationality. For example, research on "confirmation bias" suggests that people actively avoid being confronted with disconfirming evidence (Nickerson 1998), and they sometimes simply fail to notice discrepancies between their beliefs and the available evidence (Benassi, Singer et al. 1980). This literature on "confirmation bias" is sometimes taken to imply that people simply "forget" failures and ignore adverse evidence, but the actual psychological mechanism may be more interesting: for example, researchers found that people betting on sport games remember their losses *better* than their wins. Rather than just "forgetting" their apparent failures, people take the time to scrutinize them intensely in search of elements that allow them to rationalize these failures away. Typically, they accept their wins at face value but rely on *ad hoc* explanations to turn their losses into "near wins" (Gilovich 1983). Researchers have shown that humans are remarkably creative in inventing such *ad hoc* explanations for events (Ross, Lepper et al. 1977) and in explaining away adverse evidence to rescue cherished beliefs from refutation (Gilovich 1991; Tumminia 1998). In a classical experiment, Lord and his colleagues (Lord, Ross et al. 1979) asked defenders and opponents of capital punishment to read two studies, one of which suggested that the death penalty deterred people from committing crimes, whereas the other suggested that it was not an effective deterrent. Both groups detected more methodological problems in the study disfavoring their own beliefs, and hence rated this study lower, while they took the study in favor of their own beliefs at face value.

Kunda (1990) explains that, ironically, this pervasiveness of *ad hoc* reasoning and special pleading suggests that there are limits to the extent to which people are engaged in "motivated reasoning": "[T]he biasing role of goals is [...] constrained by one's ability to construct a justification for the desired conclusion: People will come to believe what they want to believe only to the extent that reason permits" (Kunda 1990, p. 483). This is because people like to think of themselves as objective and unbiased reasoners. In psychological terms, they place a high premium on consistency and impartiality (Kunda

1990; Tavris and Aronson 2008; von Hippel and Trivers in press). When people are motivated to cling to a belief, they do not feel comfortable with blithely ignoring adverse evidence or simply shutting their ears to anyone who opposes their views. Instead, they engage in more subtle forms of *ad hoc* reasoning, rationalization and special pleading to arrive at their desired conclusions and to justify their beliefs to others, e.g. reinterpreting the facts, weighing them against background knowledge, finding some reason to discredit the source, etc. (Gilovich 1991, pp. 54-56). This practice allows them to uphold an "illusion of objectivity concerning the manner in which [...] inferences were derived" (Pyszczynski and Greenberg 1987, p. 302).

The conventional wisdom on so-called "wishful thinking" is that, as Francis Bacon put it, man "prefers to believe what he prefers to be true". The psychological evidence, however, suggests that there are constraints on the ways in which people let their desires and goals guide their beliefs (Ditto and Lopez 1992). The cognitive premium on a flattering self-image of being unbiased and reasonable explains why many people, even defenders of weird belief systems, will typically scrutinize adverse evidence until they find some justification for rejecting or ignoring it (see the use of immunizing strategies in section 11.5).

Apart from that, it is clear that the power of "wishful thinking" is limited in yet another respect. Often enough, people are firmly committed to weird beliefs that show no signs of wishful thinking at all, but in fact confirm their worst fears. For example, many among the most widespread irrational beliefs – eternal damnation, witchcraft, evil conspiracies – are positively frightening and menacing (Guthrie 1993; Atran 2002, pp. 75-78). If it were true that people believe what they prefer to be true, why do they not just prefer *not* to believe in these phantoms, *a fortiori* when the evidence for them is so scant? Certainly, as psychological research has shown, if people have *already* committed themselves to belief in witchcraft or hell, and have acted accordingly over a period of time, they may be motivated to persevere in that belief. That brings us to the next section.

11.3.2 Cognitive dissonance

In many everyday situations, believers just accept adverse evidence and revise their beliefs accordingly (Sperber 1990). If I think the capital of Ghana is Abidjan and I find in an atlas that is in fact Accra, I do not write a letter of complaint to the publisher of that atlas. If I thought I left my keys in the drawer and I cannot find them there, I usually revise my belief (still, if I am really confident or just stubborn, I can persist that someone else must have taken them away). Under what circumstances can we expect people to persevere in their beliefs and to explain away such evidence? Cognitive

dissonance theory suggests that, from a psychological and motivational point of view, there has to be something *at stake*.

According to cognitive dissonance theory (Festinger, Schachter et al. 1964; Aronson 1992; Tavris and Aronson 2008), when people are presented with new evidence that conflicts with their previously held beliefs, this results in a form of cognitive tension called "dissonance". Importantly, the strength of this uncomfortable tension depends on the degree to which people have invested in their beliefs, for example by way of public commitment, or by the time and effort spent acting in accordance with these beliefs (Batson 1975). If the psychological investment in a belief is high, people are more motivated to reduce dissonance by rationalizing away disconfirming data. In the refined version of dissonance theory, dissonance arises not so much because of two conflicting cognitions, but because adverse evidence conflicts with one's self-esteem as a competent and reasonable person.[81] This accords with our earlier observation that, when people explain away unwelcome evidence, they do so in a way that allows them to uphold an illusion of objectivity. For example, if a psychic has publicly professed his powers and risks losing his credibility, he is unlikely to be put off his balance by blatant failure. Or if a believer has spent a substantial amount of time and money on astrology consults, typically no amount of rational argumentation and debunking efforts will make him renounce his beliefs. As Nicholas Humphrey noted: "psychic phenomena can, it seems, survive almost any amount of subsequent disgrace" (Humphrey 1996, p. 150). By contrast, if the psychological stakes are low, as in the everyday situations we mentioned above, the motivation for belief perseverance will be greatly reduced. Consider another example related to paranormal beliefs: suppose that Anna and Paul both start to suspect that they have psychic powers, but their level of confidence is not very high. While Paul hastens to tell his friends that he may be psychic and even performs some psychic readings, Anna decides to conduct an experiment on herself at an early point, when her beliefs are still privately held. All other things being equal, it is much more likely that Anna will abandon her beliefs silently when she discovers that they do not pan out (Humphrey 1996, p. 105), while Paul will rationalize his failures because he has already made a public commitment. Thus, we would predict that people with an inquisitive and cautious mindset are more likely to put their hunches to the test early on, and are less likely to be sucked into commitment to wrong beliefs like these. By contrast, people who rush to conclusions and start spreading the news right away

[81] Aronson distinguishes three components of a positive self-image that are shared by most people: a consistent and stable self-image; a sense of self as a competent person; and a sense of self as a morally good person. In this paper, we are mainly concerned with people's self-image of being competent and reasonable.

will more often find themselves in a situation where they obstinately refuse to abandon a false belief.[82]

A classic illustration of cognitive dissonance can be found in the landmark study by Leon Festinger and his colleagues, who infiltrated a doomsday cult and observed the behavior of the followers when the prophesized end of the world failed to come true (Festinger, Schachter et al. 1964). The followers who had resigned from their jobs, given away their material belongings and were present at the arranged place and time with full conviction in their imminent salvation, became even more ardent believers after the prophecy failed, and started to proselytize even more actively for the cult. However, those for whom the cognitive stakes were lower (e.g. those who kept their belongings and stayed home in fearful expectation of what was supposedly to come), were more likely to abandon their beliefs afterwards.

Early cognitive dissonance theory required that prophecies be sufficiently specific and unequivocal, and that believers themselves acknowledge them to be roundly refuted. This aspect of Festinger's theory was belied by more recent studies of millennial cults (for an overview, see Dawson 1999), which suggest that believers afterwards rarely if ever recognize that the prophecy which they issued has not been borne out. Indeed, the *denial* of failure is "the common mode of adaptation of millennial groups" (Melton 1985, p. 21). Instead of recognizing failure and proceeding to ignore it completely, as early cognitive dissonance theory suggested, committed believers explain away apparent failure by means of semi-plausible *post hoc* rationalizations (Dein 2001; Dawson 1999), consistent with the psychological findings on the illusion of objectivity.

Cognitive dissonance theory is concerned with the conditions that *give rise* to dissonance and that *motivate* for dissonance reduction, but as such it does not describe the way in which this tension is resolved. In section 11.3.1, we have already shown that belief perseverance is typically achieved by *ad hoc* reasoning and rationalization rather than blunt denial, and in the next section, we consider what conditions are *conducive* to this practice.

[82] The James Randi Educational Foundation offers a $1,000,000 prize to anyone who is able to demonstrate, under controlled observing conditions, evidence of paranormal or supernatural powers. Claimants for the challenge are always asked to conduct a private experiment on themselves before coming to the official test. Interestingly, after being instructed as to how a proper self-test can be performed, many of them are never to be heard of again. By contrast, according to Randi's experience, those who do turn up for the real test, because they failed to conduct such a self-test (or did not do it properly), *always* make recourse to rationalizations to explain away their failure.

11.4 The structure of self-validating belief systems

Cervantes' classic novel Don Quixote tells the tale of an elderly gentleman who is obsessed with books of chivalry, and has succumbed to the delusion that he is an errant-knight on an epic mission to restore the golden age of chivalry. Although he is confronted with a series of tragic defeats and humiliations, Don Quixote is able to persevere in his grand delusion by invoking invisible malicious wizards thwarting his every action. But Quixote is not stupid. When the canon, one of the characters in the novel, is confronted with the "extraordinary nature of Don Quixote's madness", he marvels that "in all his remarks and replies he should show such excellent sense, and only lose his stirrups [...] when the subject of chivalry was broached" (Cervantes 2008, p. 644).

If an ardent believer is confronted with what outsiders perceive as clearly disconfirming evidence of his belief system[83], he will tend to resist belief revision to the extent that he is able to come up with plausible rationalizations and excuses in the face of difficulties. As we saw, these rationalizations allow the believer to uphold an "illusion of objectivity", thus reducing the level of cognitive dissonance. If we now raise the question as to what non-motivational factors facilitate this ability to rationalize away apparent failures, we can take up two different perspectives, one cognitive-psychological and one epistemological.

From a psychological point of view, one may ask in what way intelligence is related to belief perseverance. Although the received view holds that intelligent people are less likely to accept wrong beliefs, we submit that, once intelligent people become highly committed to a belief, it will prove more difficult to put them off their balance with adverse evidence and criticism. Just as Don Quixote had no difficulties in explaining away his failures and in brushing aside counterarguments, skilled reasoners are more proficient at inventing and constructing rationalizations in the face of difficulties, and they will be more prone to belief perseverance when they experience cognitive dissonance. Even worse, intelligent people may be more vulnerable to wrong beliefs in the first place. Mercier (in press) has argued that the more people rely on reasoning through communication to achieve epistemic improvement, the more they are likely to accept a number of wrong beliefs amidst the bulk of true beliefs. According to Mercier, this explains why highly intelligent people have the habit of endorsing some quite weird beliefs. For example, Michael Shermer has discussed the cases of such scientific

[83] As we will see, the very notion of "clearly disconfirming evidence" will become problematic once we take into account the defense mechanisms and escape clauses inherent in the system.

luminaries as the renowned cosmologist Frank Tipler, who has tried to demonstrate the dogmas of Christianity with intricate mathematical and physical equations, and Harvard psychiatrist John Mack, who is a firm believer in alien abductions. As George Orwell put it in a different context: "One has to belong to the intelligentsia to believe things like that: no ordinary man could be such a fool" (Orwell 1945). Shermer argues that precisely the intellectual skills of these individuals make them highly proficient in defending implausible beliefs, and hence prone to perseverance in the face of overwhelming adverse evidence: "Smart people believe weird things because they are skilled at defending beliefs they arrived at for non-smart reasons" (Shermer 2003, p. 64).

In addition to this psychological take on the issue, we argue that it is fruitful to approach the problem of irrationality and belief perseverance from an epistemological perspective as well. What if some systems of beliefs are more resilient in the face of adverse evidence and criticism? What if some of them provide more explanatory and conceptual resources for believers to draw upon in the face of difficulties? In that case, the belief system itself would be conducive to belief perseverance and rationalization on the part of individual believers. If we accept this possibility, we can develop a straightforward epidemiological model of beliefs (Sperber 1990, 1996; Dawkins 1993; Boyer 1994, 1998): over the course of the history of human culture, millions upon millions of beliefs have been tried out and entertained by as many different persons, and only some of these are remembered, acquired and transmitted. If we accept that there is constant cultural and interpersonal variation in the generation and transmission of beliefs, it is inevitable that, for a variety of complex reasons, some ideas will be more successfully remembered, recalled and propagated. As Pascal Boyer noted: "certain features are recurrent because they are more likely to be entertained, acquired, and transmitted by human minds" (Boyer 1994, p. ix, 1998). Importantly, as Scott Atran suggested, considerations of "cognitive optimality might be at work not only at the level of individual beliefs but at the level of belief structures as well" (Atran 2002, p. 101). The claim we want to develop, is that one of the constraints that channel beliefs and belief structures is the degree of structural resilience they exhibit to adverse evidence and critical arguments. We argue that this epistemological consideration partly accounts for the puzzling popularity of certain "weird" beliefs systems. Note that our account must be distinguished from straightforward selectionist approaches of culture, typically in terms of memes or culturgens (Blackmore 2000; Dawkins 1976; Richerson and Boyd 2005). Following the critiques of Boyer, Atran and Sperber, we think that memetics and similar approaches take a too simplistic view of the notion of replication, and they largely obscure the shaping role of our cognitive architecture (Sperber 2000; Boyer 1994). Nevertheless, the seminal work on memes and mind viruses by Dawkins (1976, 1993) and Dennett (1991, 1996) contains valuable insights into the ways self-validating belief structures coalesce and are rendered impervious to criticism and adverse evidence.

In Chapter 5, we have documented how many "weird" belief systems exhibit certain internal, structural features that render them invulnerable to adverse evidence and critical arguments. We have termed these "epistemic defense mechanisms" and distinguished them from "immunizing strategies", which are defined as arguments brought forward in support of a belief system. In contrast to epistemic defense mechanisms, immunizing strategies are independent from the belief system at hand. The distinction between two types is not always clear-cut, as immunizing strategies may loosely draw on theory-internal resources, or develop into an integral part of the theory over time.

In the following section, we will briefly rehearse the different types of the epistemic defense mechanisms explored in Chapter 5, which will suffice to demonstrate that the alleged fragility of typical weird belief systems rests on a superficial analysis.

11.5 Epistemic Defense Mechanisms

11.5.1 Multiple endpoints and moving targets

In astrology and in prophetical works such as those of Nostradamus, it is typical to be presented with a series of ambiguous statements having what psychologists call "multiple endpoints" (Gilovich 1991, pp. 58-59; Hines 2003), as in the parody prediction by Woody Allen: "Two nations will go to war, but only one will win". In fact, typical astrological descriptions are amenable both to a specific interpretation and a range of broader and more metaphorical ones, e.g. "a father-figure stands behind you". This creates an asymmetry between what will count as hits and misses of the predictions in question, allowing the astrologer – or gullible believer – to switch back and forth between specific and broad interpretations. In a variation on this theme, a belief system consists of statements that are specific and exciting on first inspection, but when running into trouble, they are belatedly modified so as to make them trivial or uninteresting. The belated re-interpretation of a failed doomsday prediction on a spiritual level is a standard example of such a move (for specific examples, see 5.3.1.1). More generally still, the indeterminate and mysterious nature of many religious and pseudoscientific propositions ensures that they are closed to normal epistemic evaluation (Sperber 1990), and that contradictions and adverse evidence will go largely unnoticed to the believers (Sperber 1996, pp. 91-92; see also the discussion of "quasi-propositions" in religion, Atran 2002).

11.5.2 Postdiction of invisible causes

In certain belief systems, invisible causes are postulated to account for a range of phenomena, in such a way that their working can only be inferred ex post facto from the observed effects. If the causal relations and conditions in the belief system are not sufficiently specified, and allow for all sorts of secondary elaborations, believers can get entangled in subtle feedback loops between theory and observations, which keep the belief system forever outside the reach of empirical refutation (for examples, see 5.3.2). This pattern of spurious postdiction is also apparent in the way parapsychologists explain away null results and cherry pick data to "determine" where and when psi forces were active (Wiseman 2010). Likewise, cults groups often draw upon a range of unfalsifiable concepts and events to avert disconfirmation (e.g. Dein 2001).

11.5.3 Conspiracy thinking

(a) Conspiracy theorists typically believe, against the received view of a historical event, that a group of interested agents have been secretly pulling the strings to bring about the event in question, all the while carefully covering up their actions. They argue for this view on the basis of "errant data" (Keeley 1999, p. 118), i.e. anomalies, unexplained details and inconsistencies in the official story. On the other hand, when investigations fail to reveal the conspiracy or even flatly contradict it, believers typically turn the evidence on its head, arguing that this is exactly what would be predicted by their view. After all, conspirators, being who they are, can be reasonably expected to erase all traces of evidence leading to their plot, and to lead the rest of us astray with forged evidence (Clarke 2002).

(b) Furthermore, the conspiratorial pattern of reasoning allows believers to explain away the motives for disbelief and criticism within their own belief system, for example by accusing the skeptics of being somehow implicated in the conspiracy themselves. In this way, criticism of any sort is immediately deflected and transformed into further confirmation of the belief system (for examples, see 5.3.3).

11.5.4 Invisible escape clauses

In many pseudoscientific belief systems, we are confronted with an imponderable force or cause that, when push comes to shove, confounds the expectations initially engendered by the theory, and conveniently explains away apparent failure. A host of such escape clauses can be found in the field of parapsychology: e.g. the idea that the

presence of inquisitive minds tends to disturb psychical phenomena, which is known as "negative psi vibration" or "catapsi" (for a sceptical discussion, see Humphrey 1996; Wiseman 2010), or the argument that psi is "actively evasive" because its primary function is to "induce a sense of mystery and wonder" (Kennedy 2003, p. 67) (for further examples, see 5.3.5). In some cases, we are dealing with an immunizing strategy that is independent from the belief system at hand. In other cases (e.g. parapsychology), these escape clauses may develop into fully-fledged epistemic defense mechanisms, forming an integral part of the theory.

11.6 Epidemiology of Beliefs

11.6.1 The development of resilient belief structures

The central claim developed in this paper is that beliefs that develop into systems which are more successful in withstanding empirical failures and in "surviving" the onslaught of critical arguments, be it from the inquisitive believers themselves or from skeptical outsiders, will be more readily acquired, remembered and selected among their competitors.

Of course, beliefs do not "develop" into self-validating structures all by themselves. Beliefs are entertained by individual agents, and they are modified and revised by individual agents. Over time, the problems these agents encounter within their system of beliefs will inspire solutions in the form of modifications, reinterpretations and elaborations. Not all of these changes will be equally *successful* from a psychological point of view, and hence not all of them will tend to survive. We claim that certain "successful" configurations of beliefs may be expected to become recurrent in widely different domains, despite huge cultural and interpersonal variation.[84] Thus, rather than turning into full-blown self-validating structures all at once, we maintain that beliefs *crystallize* into such systems after a number of successive modifications and elaborations, which result from attempts to resolve inconsistencies and to rescue the belief system from apparent refutation.

[84] This effect of increased epistemic resilience can be relatively small, compared to other cultural and cognitive constraints. However, as Liénard and Boyer noted: "In cultural transmission [...] very small effects aggregated over many cycles of transmission are sufficient to create massive trends" (Lienard and Boyer 2006, p. 824).

Consider again the case of doomsday cults, which are literally confronted with the problem of surviving the day on which prophecy fails. If the day of truth arrives and the predicted events are not borne out, the belief system is faced with serious institutional crisis (Zygmunt 1970). If, on the other hand, the system is flexible enough to cope with eventual failure, by allowing for some convenient escape clauses, excuses or reinterpretations, it may withstand the impact of reality (Balch, Domitrovich et al. 1997). For example, the cult of Jehovah's Witnesses has a long history of what outsiders perceive as blatant prophetic failures, but the movement does not show any signs of disappearing. This is partly because, as Zygmunt's study on Jehovah's Witnesses makes clear, the prophecies of the cult were phrased "in a manner that made them only partially open to disconfirmation" (Zygmunt 1970, p. 944). As they allow for enough "wiggling room", the failed prophecies can always be retrospectively related to real historical events, and thus be "converted into partial successes" (Zygmunt 1970, pp. 944-945), strengthening the conviction of the followers and renewing their proselytizing efforts. In other words, the belief system of Jehovah's Witnesses has made use of the defense mechanism of multiple endpoints and deflationary revisions which we described above.

Of course, these reinterpretations do not present themselves spontaneously, but that does not mean that they are deliberately constructed by believers with strategic purposes in mind (see section 11.7). In the doomsday cult, a plausible *post hoc* rationalization of prophetic failure is typically suggested by the group leader and taken up by the other members (Dawson 1999, p. 65). Alternatively, group members may entertain different rationalizations and reinterpretations, in the full conviction that the prophecy must be true in *some* sense, and the solutions that emerge as cognitively "optimal" are taken up by other believers. In this way the belief system may slowly develop an increasing resilience in the face of adverse data.

Thus, if the believers succeed in constructing elaborations on or carrying out reinterpretations of their belief system that make it impervious to empirical failure (to which they will often be strongly motivated), the belief system will survive the day on which the prophecy fails, and live on in this more resistant form. All other things being equal, the weird belief systems that reach a cultural level of dissemination tend to be the ones that have stabilized on a form that is immune to the empirical refutation and criticism from outsiders. Those that are too fragile wither away and are simply no longer there for us to observe.[85]

---

[85] Of course, it is quite possible that the success of such a belief system is compromised by other factors offsetting the gain in epistemic resistance. For example, too many elaborations and defense mechanisms may render the belief system cumbersome and/or too complicated, hampering recall and transmission to other believers. In this way, the very features that account for its epistemic immunity may make the belief system

In contrast to doomsday prophecies, most belief system do not hinge on a single moment of truth sometime in the (distant) future.[86] Most weird belief systems – conspiracy theories, homeopathy, magic healing, parapsychology, ufology, etc. – involve claims that are supposed to derive support from currently available evidence. Essentially, however, the epistemological predicament of these belief systems is no different, as they too have to provide for the resources to cope with unwelcome facts and with disbelievers. Conspiracy theories, which are equipped with built-in protection against empirical failure and hostile criticism from outsiders, have "succeeded" in this regard. Their epistemic structure guarantees that believers will always have some way of explaining away difficulties. As we noticed above, Don Quixote's delusional belief system was fundamentally conspiratorial in nature (Farrell 1996). Anything that was in apparent conflict with Quixote's fantasy world of knights, castles, maidens and dragons, was interpreted by him in terms of malicious sorcerers who make things appear different than they are. As with any conspiracy thinker, nothing and nobody could convince Quixote that the world of chivalrous knights existed only in his imagination.

In many conspiracy theories that are currently popular on the Internet, apparently disconfirming evidence is interpreted as forged evidence and false information spread by the conspirators, and detractors are suspected of being part of the conspiracy itself, having been bribed by the government, or having merely been misled by the cunning of the evil plotters. As in the case of doomsday prophecies, the social dynamic of a group of believers may further facilitate this process. If a new piece of evidence turns up that seems to be in conflict with the conspiracy hypothesis, or a new argument is voiced by critics, different ways of explaining away these difficulties may be tried out, and the ones that are most "successful" from a psychological perspective, in virtue of their allowing believers to preserve an illusion of objectivity, are taken up by other members to become part of the belief system.

As we saw, the conspiracy template turns up in a variety of different belief systems, as it is such a convenient way of dealing with problems. For example, creationists in the second half of the 20th century have cultivated the idea that evolutionary theory and all the evidence supporting it is nothing less than a satanic ploy to lure the faithful into disbelief (Morris 1963). Similarly, UFO believers have proclaimed for over several decades that there is a vast government conspiracy to obscure the real evidence for extraterrestrial visits to the earth (Hines 2003, pp. 257-298; Park 2002, Ch. 9). In general, if a group of people is firmly committed to a system of beliefs, which is then

---

less successful in other respects. As in many biological adaptations (e.g. the peacock's tail), it is plausible that, in such cases, a trade-off will take place.

[86] And even in the case of Jehovah's Witnesses, the prophecy is embedded in a complex network of beliefs and practices.

increasingly threatened by mounting adverse evidence, the community of (remaining) believers will often settle on a form of conspiracy defense. The reader may object that this is not what usually happens in disputes between scientists, even though many of them are also highly committed to a cherished theory or hypothesis. Be that as it may, resistance to change and belief perseverance are certainly not alien to scientific disputes, even though science is valued as a self-corrective enterprise that depends on the relentless correction and overthrow of old theories. Scientists too can be unduly conservative in their beliefs, but typically they resort to more sophisticated ways for rescuing a theory from falsification, as witnessed by the repertoire of *ad hoc* moves we explored in Chapter 6 (Boudry in press). Nevertheless, in some heated scientific controversies, the losing party does resort to conspiracy theorizing: see for example the downfall of cold fusion, the Duesberg hypothesis about the non-infectious nature of AIDS, and more recently, the small pockets of continuing scientific resistance to theories of anthropogenic climate change (Pigliucci 2010).

11.6.2 An epistemological or cognitive approach?

The epidemiological argument outlined in this paper emerges in light of persistent *cognitive features* that conspire to make us vulnerable to self-validating belief systems: (a) our proficiency at *ad hoc* reasoning and rationalization; (b) the motivation to reduce cognitive dissonance; (c) the persistence of the confirmation bias; and (d) the psychological premium placed on being rational and free from bias.

Note that an analysis pertaining to the epistemic structure of belief systems accounts for only one factor in a more general epidemiological model of culture. In Sperber's epidemiological model of representations, a host of cognitive, psychological, ecological and cultural factors channel the formation and dissemination of beliefs.[87] In particular, Sperber has focused on the "relevance" (Sperber 1985; Sperber and Wilson 1995) of representations to domain-specific cognitive modules. Briefly, according to Sperber, representations that are relevant independently of a local context will stand a greater chance of cultural success: "independence of an immediate context means that relevance will be maintained in spite of changes of local circumstances – that is, it will be maintained on a social scale" (Sperber 1996, p. 140).

[87] In the newly emerging field of Cognitive Science of Religion (CSR) (Boyer 1994; Barrett 2000; Boyer 2001; Atran 2002; Barrett 2004; for a recent overview, see Barrett 2007), researchers have brought these insights from cognitive psychology to bear on the study of religion: "much of what is typically called religion may be understood as the natural product of aggregated *ordinary* cognitive processes" (Barrett 2000, p. 29).

By way of illustrating how these relevance considerations apply likewise to the belief systems discussed here, consider again the case of conspiracy theories. Our susceptibility to evil conspiracy theories is not only a result of their self-validating epistemic structure, but is arguably also a function of at least two specific cognitive modules: a mechanism for agency detection (Guthrie 1993; Barrett 2000; Atran 2002; Barrett 2004) that is biased towards over-attribution of agency in our environment[88], and a "hazard-precaution system" (Lienard and Boyer 2006) geared towards detecting danger and acting in dangerous situation.[89] Seeing that conspiracies involve the secret and potentially threatening actions of hidden agents, we realize why they tend to activate the cognitive processes mentioned above, and hence why they never fail to command our attention. In the words of Liénard and Boyer, cognitive modules such as these are liable to "cognitive capture" (Lienard and Boyer 2006, p. 821) by specific representations that meet their input conditions. Our epistemological analysis further contributes to an understanding of how evil conspiracies of all stripes – cover-ups by the government, secret plans of the Illuminati or the Elders of Zion, etc. – often reach a level of wide cultural dissemination and why they are so resistant to adverse evidence.

Although our epistemological argument is informed by research on motivated reasoning and cognitive dissonance, we view it as a level of explanation in its own right, which allows for the identification of patterns and trends that are not visible from the lower level of cognition. Our susceptibility to self-validating belief systems becomes only transparent when we connect *different* cognitive and psychological findings and pursue an epistemological approach to the problem.

11.7 Questions of Sincerity

Epistemic defense mechanisms and immunizing strategies may appear *convenient* for the believer who is motivated to cling to his beliefs, but in fact to think in strategic and intentional terms may be misleading here. If our analysis of the epistemic rationale of

[88] According to cognitive and evolutionary psychologists, the hypersensitivity of these cognitive mechanisms makes good evolutionary sense: traces in the grass or a rustle of leaves may signal the presence of enemies, sickness may be the result of poisoning by a rival, etc. The cost incurred by a false negative (failing to detect agents) is significantly greater than the cost of a false positive (detecting agency where there is none).

[89] Another example, which we have already mentioned, is our inclination towards superstitious beliefs, which may be partly explained by our cognitive proficiency at pattern detection and our difficulties with evaluating random coincidences (Gilovich 1991).

self-validating belief systems is accurate, this appearance of strategic "convenience" may well be the outcome of cultural-selection processes rather than straightforward and conscious deliberation. In this respect, it is interesting to have a brief look at the suspicions which skeptics of pseudoscience have often voiced regarding the sincerity of believers. We rehearse two often-heard arguments to that effect and proceed to show why they are largely misguided (or at least inconclusive).

11.7.1 Avoiding tricky situations

Many pseudoscientists seem to carefully avoid situations that would put them at a risk of empirical refutation. For example, mediums and clairvoyants have all sorts of excuses for refusing to participate in the type of controlled test that is bound to expose their lack of powers all too clearly (Hines 2003). As we noted, believers consider proper scientific investigation of the paranormal inappropriate or impossible, arguing that the phenomena in question are, in the words of a skeptic, "unpredictable, unrepeatable, shy, highly context-dependent, droll, evanescent, dreamlike" (Humphrey 1996, p. 73). Likewise, not only do successful astrologers and soothsayers avoid making claims that are too vulnerable to refutation, but they are particularly reluctant to do so when questioned by skeptics. Some pseudoscientists even expressly warn against giving demonstrations in the vicinity of inquisitive minds. As in the words of one of Franz Anton Mesmer's followers: "Never magnetize before inquisitive persons!" (quoted in Mackay 1974 [1841], p. 290). For many skeptics, this suspiciously evasive behavior on the part of believers is a telltale sign of insincerity.

11.7.2 When push comes to shove ...

Many people proclaim to believe in supernatural or paranormal causation, all the while relying on more mundane courses of action when push comes to shove. For example, as Nicholas Humphrey noted, many people profess to believe in telepathy, but "when they themselves want to communicate to a distant friend, they play safe and write or call them" (Humphrey 1996, p. 55). Interestingly, anthropologists have noted that, in many cultures where supernatural spells and magic are used for achieving a certain material goal (rainfall, a good harvest, victory in a battle), people always make sure to rely on more down-to-earth methods as well, which suggests that they are not all too confident in supernatural causation as they claim to be. In a paper discussing magic and religion, anthropologist Eli Sagan noted:

> A people going to war may sing over their spears in order to make them more effective. If there ever have been people who felt they could defeat an enemy in war merely by singing and who therefore dispensed with spears we have not heard of them. (Sagan 1979, p. 93)

Similarly, many believers suddenly lose their professed faith in the paranormal and supernatural when their own lives are at stake. For example, when the chakra healer himself falls seriously ill, he will make sure to consult a regular doctor. In other cases, supernatural faith suddenly becomes somewhat half-hearted. For example, when Pope John Paul II was shot and critically wounded in an assassination attempt in 1981, he asked the surgeons not to remove his Brown Scapular during the operation, stating that Our Lady of Fátima would help him recover. But why did the Pope rely on scientific medicine and surgery in the first place, instead of putting faith in supernatural help? On the basis of similar examples, Humphrey and other skeptics have suggested "that most people know only too well how things stand" (Humphrey 1996, p. 56).

11.7.3 No Need for Deliberation

Against the two arguments presented in section 11.7, we submit that believers' suspicious behavior may well result from more subtle mechanisms of self-deception and rationalization (von Hippel and Trivers in press). We briefly show how the epidemiological argument presented here supports that conclusion. As for the first argument, it is instructive to imagine the fate of a psychic who is *not* so careful to avoid tricky situations or who is not equipped with a bag of excuses for doing so. For example, an astrologer who is confident enough to make very risky predictions is bound to have a hard time explaining his failure after the fact. A self-proclaimed psychic who recklessly accepts the invitation for a scientific experiment, unprepared for refutation, will sorely disappoint himself, not to mention the followers who witness the failure.[90] By contrast, psychic healers and mediums who happen to come across an argument that suggests to them that scientific investigation of psi is impossible or inappropriate, will be less likely to be confronted with cases of blatant failure. Those who have learnt and cultivated successful excuses for shying away from tricky tests are precisely those who are still in the game to defend their beliefs and to convince others.

Of course, this sketchy explanation leaves open many unanswered psychological questions. Further discussion may bear on how selectionist explanations translate on

[90] Interestingly, the skeptical literature does contain a few cases of psychics who made a "reckless" claim, agreed to be put to the test, and afterwards did accept the negative verdict, or at least started to doubt their powers.

the psychological level, and how they relate to issues of intentionality. If an agent starts to rationalize away adverse evidence for her beliefs, either through arguments of her own or through ones that she has picked up elsewhere, does this not suggest a certain level of awareness on her part? It seems that, at least, the agent has come to realize that a certain situation may prove threatening to her belief. Moreover, in section 11.3.1 we suggested that agents *consciously* scrutinize conflicting evidence in search for reasons to dismiss it. Still, we think this is not incompatible with an agent sincerely holding a weird belief: precisely the agent's solid conviction of being right arguably motivates her search for flaws in unwelcome evidence ("something *must* be wrong with this data"). In any case, a thorough discussion of intentionality, self-deception and motivated reasoning falls outside the scope of this paper (see for example Mele 1994; von Hippel and Trivers in press).

In relation to the second argument, we can apply a similar reasoning. In Chapter 12, we will see that people who expect the sort of supernatural or paranormal causation that would make more mundane courses of action to the same effect superfluous, are bound to be disappointed by the results. The person who asks a friend for dinner by telepathic means alone will surely be spending a lonely night. In general, those who expect tangible results from psi powers will be forced either to abandon their belief in psi, or to correct – by way of rationalization – their expectations on the causal power of psi in a way that does *not* make ordinary modes of action superfluous.

As for the cases of faltering faith when life is at stake, an even more obvious selection process is at work. For example, it is not difficult to imagine what *will* happen to a people who relies wholeheartedly on magical spellbinding and dispensed with weaponry. As Sagan (1979, p. 93) dryly noted, they will probably be "all dispatched in the midst of their spellbinding". Indeed, there are many documented cases of people who go to battle virtually unarmed because they believe they enjoy supernatural protection. Similarly, up until this day many religious people who are inflicted by a lethal disease relinquish any form of medical treatment, convinced as they are that faith alone – or "alternative" therapy – can save them (e.g. Peters 2008; Edgerton 1992). If anything, the anthropological evidence suggests that people are perfectly capable of sincere conviction in highly dangerous beliefs. In any case, to return to our present argument, even when self-proclaimed supernaturalists pursue more mundane courses of action as soon as life is at stake, this does not *necessarily* mean that their beliefs are insincere. Again, the apparent design may be "authorless", resulting from a process of cultural selection.

In summary, we cannot take the "convenience" of believers' suspicious and evasive behavior at face value, i.e. as a token of strategic deliberation. Rather than being the result of conscious deliberation on the part of individual believers, epistemic defense mechanisms and evasive behavior in general may exhibit what Daniel Dennett has termed a "free-floating" rationale (Dennett 1996, p. 78, see also pp. 164-165).

Note that our epidemiological argument does not deny the existence of conscious deliberation to the same effect. Even if the question of sincerity may often be difficult to resolve, it is obvious that there are a lot of conscious impostors among pseudoscientists and paranormal mediums (Wiseman 1997, p. 12), and that religious leaders need not always be sincere in their beliefs (Dennett and LaScola 2010). For example, mediums such as Uri Geller and faith healers like Peter Popoff have been caught cheating several times during performances. Thus, the decision to protect one's self-proclaimed paranormal powers from exposure can be perfectly deliberate. However, as Dennett notes:

> A tactic that works *can* be used deliberately and viciously, but it can also work – sometimes better – in the hands of an innocent enthusiast who would never dream of doing anything duplicitous. (Dennett 2006, p. 365)

Consider magician Robert-Houdin's advice to performers "never [to] announce beforehand the nature of the effect which you intend to produce" (quoted in Wiseman 1997, p. 42). In case something goes wrong, this golden rule allows the magician to finish the trick in another way, without having failed in the eyes of the public. But the same rationale can work without intentional action, for example in the hands of a psychic who has "learnt" not to announce the psi effects he intends to produce.

11.8 Conclusion

In this paper, we challenged the common assumption that pseudosciences and other forms of weird belief systems are inherently fragile. Instead, they exhibit a surprising degree of resilience in the face of apparently adverse evidence and criticism from outsiders. Based on a number of findings in cognitive psychology, we argued that this invulnerability of belief systems may in part explain their unabated popularity. All other things being equal, belief systems that allow the believer to remain outside the reach of refutations, or that provide some convenient ways of coping with difficulties, will be more likely to be selected among competing beliefs and belief systems, and more likely to be disseminated. In this way, our argument is intended as a contribution to the general question about human culture set forth by, among others, Dan Sperber: "[W]hy are some representations more successful in a human population, more "catching" than others?" (Sperber 1996, p. 58). We also noted that the use of epistemic defense mechanisms and immunizing strategies, together with the generally evasive behavior of

pseudoscientists, often strike the outsider as suspiciously convenient. However, rather than being the outcome of conscious deliberation on the part of believers, this strategic convenience may well be authorless – resulting from mechanisms of cultural selection.

Our susceptibility to self-validating belief systems is a function of several aspects of the way our human "belief engine" works: its inclination towards confirmation bias, its proficiency at rationalization and *ad hoc* reasoning, its valuation of an appearance of objectivity, and its motivation for cognitive dissonance reduction. If we view these insights from cognitive psychology in an epistemological light, and if we insert them in an *epidemiological* model of beliefs, then the enduring popularity of self-validating belief systems and the recurrence of defense mechanisms and immunizing strategies is hardly surprising.

Chapter 12.
In Mysterious Ways: On the Modus Operandi of Supernatural Beings

God moves in a mysterious way
his wonders to perform
(William Cowper)

'And don't tell me God works in mysterious ways'
Yossarian continued, hurtling on over her objection.
'There's nothing so mysterious about it. He's not working at all. He's playing. Or else He's forgotten all about us. That's the kind of God you people talk about – country bumpkin, a clumsy, bungling, brainless, conceited, uncouth hayseed.' – Joseph Heller, Catch22 *(1994 [1962], p. 206)*

God is silent, now if we can only get Man to shut up.
(Woody Allen)

Abstract.[91] The psychology of prayer and supernatural causation has received surprisingly little attention from empirical researchers. This paper discusses implicit belief patterns about the *modus operandi* of gods and other supernatural agents. We review the available empirical literature on petitionary prayer and modes of supernatural causation. Building further on the concept of *theological incorrectness*,

[91] This paper is currently under review at the journal *Religion.*

which has emerged in the burgeoning field of cognitive science of religion, we propose that religious believers "prefer" modes of divine action that are subtle and indistinguishable from the natural course of events. This psychological account is based on mechanisms of psychological self-correction and rationalization, confirmation bias, and intuitive *folk physics.* We proceed to extend our argument into an epidemiology of religious representations concerning supernatural causation, taking into account the different cognitive and external constraints channeling the dissemination of religious beliefs.

12.1 Introduction

Petitionary prayer is one of the most widespread expressions of religious behavior, but surprisingly little is known with regard to its psychology.[92] Religious believers all over the world have attempted to engage in interactions with gods, spirits, witches, dead ancestors and other supernatural beings (Zaleski and Zaleski 2005), asking them to intervene on their behalf and bring safety, good fortune (or bad fortune for others), cure from illness, and many other goods. The psychological dimension of this interaction with supernatural beings is a relatively unexplored domain. According to believers, how does God go about answering prayers? Under which circumstances may one expect supernatural beings to intervene in the natural world? In the book of Mark, we read that "What things soever you desire, when you pray, believe that you receive them, and you shall have them" (Mark 11:24). However, even the devout will admit that these are rather high hopes, and often the book of Psalms is more on the mark: "Why do you stand afar off, O Lord? Why do you hide yourself in times of trouble?" (Psalm 10:1) In the Christian tradition, the problem of God's silence in times of trouble has been pondered by countless theologians and ordinary believers, and as with the classical problem of theodicy, many ingenious rationalizations have been put forward (e.g. Murray 1993; Howard-Snyder and Moser 2002; Swinburne 2004; but see Schellenberg 2006).

In this paper we discuss a related but equally pertinent question: what explicit or implicit beliefs do people hold about the nature and mechanism of supernatural intervention in the natural world? How does God bring about things in the natural

[92] Naturally, praying practices are complex and can take a diversity of forms (Zaleski and Zaleski 2005). In this paper, however, we confine ourselves to petitionary prayer and similar practices in other religious tradition.

world, of which, presumably, he is not part (for a philosophical account, see Fales 2010) By what means does he fulfill the needs of the person soliciting help? In other words, what is the psychology of the *modus operandi* of supernatural beings? We develop an epidemiology of religious representations concerning petitionary prayer and supernatural causation, on the basis of the concept of *theological correctness* and cognitive constraints on the formations and dissemination of religious beliefs. As a point of departure, we discuss a series of original experiments on the psychology of prayer conducted by Justin L. Barrett, an anthropologist and leading figure in the cognitive science of religion (Barrett 2001).

12.2 Theological correctness

In the newly emerged field of Cognitive Science of Religion (CSR) (for a recent overview, see Barrett 2007; Barrett and Lanman 2008), researchers have brought findings from cognitive and developmental psychology to bear on the explanation of religious ideas and practices. CSR researchers argue that the formation and dissemination of religious representations is channeled by a number of domain-specific cognitive systems that are stable across different cultures and are noticeable in early childhood (e.g. Spelke and Kinzler 2007). This basic cognitive architecture makes us predisposed towards believing in supernatural entities such as invisible ancestors, immaterial spirits, animals that can change shape, ghosts, holy mountains, etc. In a review article of the field, Barrett takes as the main tenet of CSR the thesis that "much of what is typically called religion may be understood as the natural product of aggregated *ordinary* cognitive processes" (Barrett 2000, p. 29).[93]

One interesting concept emerging from this cognitive take on religion is that of *theological correctness* (Barrett and Keil 1996; Barrett 1999; Slone 2004), which refers to the discrepancy often observed between the official theology of a religion and the actual beliefs and practices of believers. Barrett observed that orthodox theology typically

[93] Although we have framed our discussion in terms of religion, it need not be confined to traditional religious categories. In some cases, the practice of directing requests to supernatural entities may be better classified as instances of 'magic' or 'superstition' rather than (institutionalized) religion (see below). This need not be problematic, as one of the tenets of CSR's naturalistic approach is precisely a dissolution of sharp boundaries between such categories as religion, superstition, magic and pseudoscience. By explaining religious beliefs and practices as predictable by-products of our basic cognitive architecture, the CSR approach automatically brings religion closer to other expressions of human nature.

dictates properties of supernatural beings that are highly counterintuitive and that strain our cognitive resources, e.g. omnipotence, omniscience, eternal existence. When questioned about their opinions and given some time to reflect, people profess to accept official theology, but when they are engaged in "online" tasks, applying religious concepts in practice, they make tacit assumptions that violate official theology. Instead, believers tend to fall back on more intuitive and anthropomorphic versions of supernatural beings. Boyer (2001, p. 285) has termed this the "tragedy of the theologian".

For example, while Christian doctrine demands that God can attend any number of events at the same time, people are caught reasoning as if God answers one prayer and then shifts his attention to the next. To take another example, Jason Slone (2004) has argued that the Calvinist doctrine of predetermination is "maximally counterintuitive", because it leaves no room at all for human free will, and hence has over time yielded to a conception of God that is more consonant with our intuitions.

The practice of prayer faces a number of other theological paradoxes that are shared by all religious traditions entertaining the notion of supernatural omniscience (Johnson 2005; Bering and Johnson 2005), in particular the three great monotheistic religions. First, praying to an all-knowing being such as God appears rather pointless, because God is supposed to be aware of my problems in any case. Second, if we make the additional assumption that God is morally perfect, and assuming that I request something morally good, God would have already done what I demand from Him anyway. But then in what sense can praying have any effect (Stump 1979)? If people would consistently pay heed to these theological doctrines, they would not be engaged in petitionary prayer. Nevertheless, religious believers who assent to this official theology cannot help but inform God about their problems, just as they would do with ordinary social actors who have no full epistemic access to their inner mental life. According to Barrett,

> [t]he simplification of concepts from the theological to the religious level appears to consist of a systematic distortion of features such that they more closely resemble intuitive ontological assumptions [...]. (Barrett 2000, p. 30)

In this paper, we focus on another form of cognitive tension that arises from traditional theology with respect to prayer. If God is an all-powerful Being, *prima facie* He may perform actions in any way He chooses.[94] Religious believers are generally not

[94] Of course, God cannot perform any actions that either involve logical contradictions (creating a round triangle), violate his other attributes (committing a sin, since He is morally perfect), or are self-defeating (creating a stone that He cannot lift). Further limitations may apply, but they are not relevant to our purposes, as people rarely pray for divine actions that involve such outright contradictions. In this context, we merely want to argue that there is no intuitive reason to expect an omnipotent (or very powerful) being to perform only certain *types* of action as opposed to others.

instructed regarding the mode of operation of the divine being to which they direct their requests. A similar problem applies to any supernatural being who is conceived as very powerful and capable of action in the natural world. As the theory of theological correctness suggests, however, the praying habits of ordinary believers need not reflect the theological doctrine of omnipotence.

12.3 Modes of supernatural causation

12.3.1 Ranges of action

In an interesting paper on petitionary prayer among North American Protestants, Barrett (2001) applies the theory of theological correctness to divine omnipotence in Christianity. As far as Christian theology is concerned, he argues, God may act either mechanically (materializing or removing physical objects, influencing physical processes), biologically (affecting the health of living beings, e.g. healing a person or striking him with disease), or psychosocially (influencing psychological states, e.g. relieving my pain, giving me the strength to face an ordeal, make someone fall in love with me, etc.). Barrett predicts that the silence of orthodox theology with respect to God's preferred *modus operandi* leaves a "theological vacuum" (Barrett 2001, p. 268) that is filled by ordinary intuitions about what one may reasonably expect from social agents. Past research has established that believers intuitively conceive of God as being located in some distant place (viz. Heaven), even if their official theology dictates otherwise. Furthermore, believers know from experience that normal social actors are bad at mechanical "action at a distance", whereas they are good at "affecting psychological states [...] at a distance" (Barrett 2001, p. 260). Because religious believers imagine God as a "human-like agent far away", they think He can be expected to act psychosocially rather than mechanically, thus ignoring the doctrine of divine omnipotence in practice. Indeed, in a series of experiments with Protestant subjects, Barrett found a praying preference among his subjects for psychosocial instead of mechanical or biological acts. Subjects were presented with a number of fictitious scenarios describing a troublesome predicament in which divine help would be welcome. For every scenario, they were presented with a mechanistic, a psychological, and a biological solution, and asked to rate how likely they would pray for that particular solution. Barrett found that subjects preferred the psychological (M = 5.77)

solution over either the biological (M = 4.93) or mechanistic one (M = 4.23), with a significance level of 0,001.

Barrett correctly points out that psychosocial action is less constrained in the sense that it does not require physical contact, and that the intuitions of even young infants are sensitive to this fact (Woodward, Phillips et al. 1993; Spelke, Phillips et al. 1995). However, psychosocial action reaches its own limits at distances not far beyond those of mechanical action. Leaving aside modern telecommunication technology, psychosocial action by way of speech and body language is limited to a few (tens of) meters. Admittedly, some methods exist for extending the radius of psychosocial action – waiving from a distance, sending a courier, emitting smoke signals – but these are limited in speed and efficiency. In any case, similar extensions are available for physical and biological action as well, e.g. throwing a missile, sending a drug or poison. At small and medium distances, psychosocial action is very efficient indeed, but beyond that point it soon reaches limits comparable to those of mechanical action (which, of course, is related to the fact that psychosocial action is really a special case of mechanical action). In other words, at great distance, where God presumably resides, a normal social actor would not be any more efficient in effecting psychosocial action than mechanical action. On the other hand, we already possess indirect evidence that people can conceive of psychosocial action at a large distance. After all, prayer typically consists of a silent mental act directed at God. If people believe that God listens to their prayers, then they believe that God can be reached by purely psychosocial means. Maybe people just expect that God will return psychosocial requests with psychosocial actions?

The questions remains if, as Barrett supposes, mechanical action at a large distance strains our causal intuitions more than psychosocial action does, and is therefore intuitively less preferable. Although theoretically plausible, we think this hypothesis sits uncomfortably with the available empirical data on religious prayer practice and belief in supernatural causation. As for Christianity, one of the most common reasons for prayer are health issues (Schmied 2007; ap Siôn 2009), and Christian believers all over the world are firmly convinced that God can answer prayer by effecting a miraculous healing – a belief that has often met with dire consequences (Peters 2008).[95] According to a 2004 survey by the National Center for Health Statistics, 43 percent of the American adult population had prayed for their own health in the previous year

[95] It might be argued that people think of disease as a *psychosocial* affliction, e.g. possession by an evil spirit. Although this may apply in particular cases, it certainly does not work as a general solution. Nowadays, in contrast with pre-scientific times, religious believers are often well aware of the biological nature of their illness (e.g. tumor, infection), but that does not stop them from praying to God and attributing their healing to supernatural intervention. In any case, the psychologization of supernatural causation does not work at all in the examples of straightforward physical causation which we will discuss next.

(Barnes, Powell-Griner et al. 2004). As is well known, the official procedure for beatification in the Roman Catholic Church requires a miraculous healing "from beyond the grave".[96] The reliability of such miracle reports has stirred controversy for over a long time (Hume 2000 [1748]; Peters 2008; Dawkins 2006), and ever since Francis Galton sparked the great prayer-gauge debate at the end of the 19th century (Galton 1872; Zaleski and Zaleski 2005; Mullin 2008; Park 2008), the alleged therapeutic effects of petitionary prayer have received attention from medical researchers (for a recent example, see Benson, Dusek et al. 2006), and even the effect of prayer on plant growth has been investigated (for an overview, see Francis and Evans 1995).

Belief in supernatural causation of a biological and mechanical sort is a recurrent feature of religions across the world, as is the practice of appeasing, petitioning and asking favors from supernatural beings (Zaleski and Zaleski 2005). In his worldwide survey of 186 different cultures, Murdock (1980) reports that, in every single one of them, illness and misfortune are attributed to the actions of supernatural beings (Johnson 2005). Typical requests for non-psychosocial action include praying – or performing rituals – for rainfall or good harvest, for averting natural disasters, for protection on a battlefield, for pregnancy, etc.

This solid tradition of mechanical and biological interventions is also reflected in virtually all religious scriptures, where supernatural beings are portrayed as capable of performing all kinds of non-psychological miraculous feats (next to psychosocial ones). Many examples from the Old and New Testament are well-known: God parting the Red Sea, tearing down the walls of Jericho, appearing in a burning bush, healing king Hezekiah and the prophetess Miriam, and of course, impregnating a virgin and physically coming down to earth as Jesus Christ. If people view God as someone from whom they may intuitively expect psychosocial action, why are accounts of non-psychosocial forms of divine intervention so pervasive?

12.3.2 The causal closure thesis

In view of the pervasiveness of belief in both mechanical and biological actions, one may wonder whether people have any intuitive difficulties with imagining such modes of divine causation, as opposed to psychosocial action. Research suggests that believers do not show any particular interest in the causal mechanisms by which supernatural beings effect changes in our world, and instead like to focus on the motivations and

[96] The alleged miraculous healing from beyond the grave (by the candidate saint) has to be investigated by the Consulta Medica, a board of physicians appointed by the Vatican, to determine whether the recovery was sudden and permanent, and to rule out scientific explanations for the healing.

intentions of these entities (Boyer and Bergstrom 2008, p. 121). We propose that it is instructive to approach this problem by considering the laws of the natural environment in which religious believers go about performing their prayers and doing their business. Researchers in the theological correctness framework have correctly outlined the ways in which ordinary cognition makes people stray from official theological doctrines, but they have paid relatively little attention to the way the world out there can effect similar constraints. In his book on theological (in)correctness, D. Jason Slone even writes that CSR researchers should not be much concerned with whether religious representation "refer to external realities":

> The content-claims of religions are peripheral to the actual object of study in the cognitive science of religion. [...] Whether or not gods exist makes little or no difference at all to the study of brain mechanisms that are involved in the production of religious thought and the performance of religious actions. (Slone 2004, p. 47)

Slone is right that external reality does not affect the innate brain mechanisms responsible for religious beliefs (although of course it does so indirectly via the working of natural selection), but we think it is self-evident that external reality must in some way influence the specific representations that these cognitive mechanisms give rise to. If we want to develop a realistic epidemiological model of religious representations (Sperber 1996), we have to take into account *both* our cognitive architecture and the external stimuli on the basis of which our cognitive apparatus is operating.

In the case of petitionary prayer, our epidemiology of representations must be informed by the relevant scientific knowledge on supernatural causation: (i) there is no credible evidence for the efficacy of petitionary prayer or other forms of supernatural intervention in the natural world (mechanistic, psychological or otherwise), and the most extensive and careful studies of petitionary prayer have yielded negative results (Benson, Dusek et al. 2006, pp. 378-381; Park 2008; Hines 2003; Matthews, Conti et al. 2000), (ii) in light of foundational scientific principles such as the law of energy conservation (Fales 2010), interference of supernatural agents in the natural world is highly improbable and comes at a steep cost (viz. revising the whole framework of physics, see section 1.4.2).

If we lived in a world in which prayer would regularly be answered (or some prayers, maybe only those of some religious creeds), this would clearly make a difference on the formation and dissemination of prayer practices and beliefs. For example, if the prayers of a certain religious creed were to work reliably for curing disease, we could imagine the news to spread like wildfire, and surely that religion would rapidly win new converts. For the sake of the argument, therefore, let us take for granted the causal closure thesis of nature, which holds that no physical event can have a non-physical

cause. In that case, soliciting the help of supernatural beings for bringing about natural effects anywhere in this world can never be efficacious.

To see how this state of affairs is bound to affect the pattern of prayer beliefs and practices, consider the following scenario.[97] Suppose I am cast up on a desert island with no food, no drinking water and no prospect for help. I know that ships are passing by occasionally, but I have no idea when the next ship is due. What if I direct a prayer to God and ask Him to materialize food and water before my eyes? If the causal closure thesis is correct, reality will soon catch up with me and I am bound to be disappointed. By contrast, if I pray to God to grant me strength to endure my ordeal until the next ship arrives, my predicament is a little different. If indeed I manage to survive until the next ship comes by, I might be tempted to attribute my rescue to God. We propose that people prefer the latter form of prayer not because of the psychosocial nature of the request, but rather because the effects asked for are indistinguishable from the natural course of events. Note that the two characterizations need not concur. If I pray for rainfall (for drinking water) or for a ship to come by, I am clearly requesting physical acts from God, but this time of a far more subtle sort than in the case of food materialization. A ship might have been on its course to my island in any case, and maybe clouds were already packing together. In these cases, no less than in the psychosocial example, it is impossible to distinguish divine action from the contingent and natural unfolding of events. Has God really intervened on my behalf, for example by subtly steering the boat's course, or influencing ocean currents, or manipulating the captain into taking a different route? There is no way to ascertain this, as there is no way to know for sure whether I would have survived a few days of starvation in any

[97] The scenario is similar to one that has been used in Barrett's experiment (Barrett 2001, p. 261).

case.

Figure 6 After drinking a fake magic potion brewed by Getafix, Centurion Crismus Bonus believes he possesses superhuman strength, but he is soon forced to lower his expectations (from *Asterix the Gaul* by Goscinny & Uderzo).

12.3.3 Subtle divine action

In the conclusion of his prayer study, Barrett suggests that, in addition to psychological causation, people may also prefer modes of divine action that are "ambiguous", which aligns well with our proposal. However, Barrett contrasts this with forms of "mechanistic causation", whereas we think that mechanistic causation can be effected

in subtle ways as well. We can distinguish at least three kinds of settings in which (alleged) supernatural causality is rendered subtle and unascertainable.

(1) supernatural agents may interfere with or manipulate complex, stochastic processes in which causal relations are difficult to assess, e.g. weather phenomena, natural disasters, luck on a battlefield, success in chance and sport games.

(2) they may influence natural processes that are either invisible or difficult to observe directly, and whose causal determinants are poorly understood, e.g. being cured from or stricken with disease, becoming pregnant.

(3) they may act as partial causes in conjunction with natural causes (Lupfer, Tolliver et al. 1996, pp. 388-389), in a way that makes it difficult to disentangle the respective contributions, e.g. giving me the strength to win in a duel, helping me finish an exam, supporting a bridge that is on the brink of collapse.

In all these cases, we have no full epistemic access to the causal relations and causal antecedents responsible for the effect, which *allows* our minds to (partly) attribute the event to supernatural agency. To put it in another way, explanation in terms of supernatural causation are parasitic upon events whose causal nexus is at least partly mysterious, i.e. not (fully) open to epistemic access. Note that, although the supernatural *mode of action* may be subtle and unascertainable, the alleged *effect* attributed to supernatural intervention need not be. After all, being cured from a lethal disease or surviving on a battlefield are quite tangible results.

When pope John Paul II was shot and severely wounded in an assassination attempt in 1981, he believed that the "motherly hand" of Our Lady of Fátima "guided the bullet's path", enabling the pope to stop "at the threshold of death" (Stanley 2000). Any form of robust physical or biological intervention would have sufficed to protect the pope, but Catholics believe that the Virgin Mary interfered in this particular, subtle way: not by preventing the gunman to shoot in the first place, or by directing the bullet away from the pope's body altogether, but by steering its course ever so slightly so that the Pope, though severely hurt, *just* managed to survive the assassination attempt.

Similarly, people have no difficulties with praying for biological interventions *per se*, such as cure from arthritis or cancer, but they are unlikely to pray for an amputated limb to grow back. Likewise, praying for rainfall to ensure good harvest seems an acceptable praying habit, but asking God to materialize full-grown crops before your very eyes seems a little unusual, even if God is omnipotent. The point is summoned well in "Le Jardin d'Épicure" by Anatole France (1894), in which the French poet visits Lourdes with a companion who, upon seeing all the braces and crutches hanging there

as evidence of healing, remarks: "A single wooden leg would have been quite more convincing!"[98]

By its very nature, psychosocial causation (e.g. possessing courage or mental strength, being relieved from anger or depression, etc.) is typically more complex and less observable compared with many forms of biological and physical causation. For that reason, it is plausible that the 'preference' for subtle modes of causation often translates into requests for psychosocial causation, as Barrett's research suggests. However, whereas Barrett's (2000) account entails that forms of biological and physical causation are counterintuitive *in general* and hence less preferred, our hypothesis predicts that people have no cognitive difficulty in conceiving such acts and hence are not hesitant to request them, provided that their mode of causality takes on subtle and unascertainable forms. Note that, while we argue that people regularly pray for supernatural interventions that *de facto* involve physical and biological causation, we do not claim that people bother to cognitively represent the precise *modus operandi*, either when making their requests (through prayer or ritual) or when attributing some subsequent event to a supernatural actor answering their prayer. As Boyer and Bergstrom wrote, people are less interested in the precise causal mechanisms of supernatural intervention than they are in its effects and in the agents responsible for them.

> People assume that the ancestors or gods are involved in various occurrences (bad crops, illness, death, etc.) but generally do not bother to represent in what way they bring about those states of affairs. (Boyer and Bergstrom 2008, p. 121)

In superstition and magic in particular, we find that people are capable of belief in the causal powers of certain objects and events without bothering to represent the specific causal mechanisms involved (see for example Evans-Pritchard 1965, pp. 82-83). To give a few examples from Western culture, it is completely unclear how throwing a horse-shoe or walking under a ladder could bring about luck or misfortune, although this does not stop people from finding such causal beliefs perfectly credible (Slone 2004, pp. 103-120). This lack of interest in causal understanding aligns well with the idea developed by Dan Sperber and Pascal Boyer that religious explanations are "relevant mysteries" (Sperber 1996, p. 73; Boyer 2001, p. 14): they do not so much explain events in terms of more simple and familiar processes, but they make use of salient and evocative mysteries instead.

[98] The remark is often wrongly attributed to Anatole France himself, who in fact disagreed with his friend and argued that wooden legs would not have impressed him.

12.4 The psychology of supernatural causation

12.4.1 Psychological self-correction

How does this preference for subtle divine action develop and play out psychologically? Janssen et al. (1990) and Barrett (2004) have drawn attention to the rather indeterminate and abstract way in which religious people often describe the effects of their prayers (support, blessing, trust, etc.), appearing in stark contrast to the concrete needs which typically occasion them (e.g. serious illness). An interesting suggestion for this discrepancy is hinted at by Janssen et al., but not further developed in their paper: "It could be argued that people adapt the intended effects to the experienced effects, accepting a principal discrepancy between needs and effects" (Janssen, De Hart et al. 1990, p. 105). Indeed, even if the intended effect of my prayer is something like sudden and full recovery from illness, it is plausible that people will eventually lower their expectations. Consider the famous Serenity prayer written by the Protestant theologian Reinhold Niebuhr:

> God grant me the serenity / To accept the things I cannot change / Courage to change the things I can / And wisdom to know the difference.

Instead of asking God to bring about a certain state of affairs directly, as one may expect an omnipotent being would be capable of, I ask him to influence my psychological attitude regarding my predicament. As we argued above, this mode of action (condition (3), as a partial cause cooperating to a natural effect) is more difficult to ascertain than straightforward intervention on God's part. Moreover, note that the very structure of the Serenity prayer can accommodate a certain amount of failure. If I succeed in changing whatever it is that I wanted to change, this is because God granted me the courage to do so. On the other hand, if I fail to change whatever it is that I tried to change and eventually I have to give up, this is because God granted me the serenity to realize that I could not change it anyway (and of course to know the difference).

In general, religious believers who request miracles from God that are subtle and indistinguishable from the natural course of events, will have a better chance of finding themselves in a situation in which they can attribute the events in questions to God answering their prayers. A better chance, that is, if compared with those who request firework displays by God. As Nicholas Humphrey wrote, people must have known all along that full-fledged and palpable miracles just don't happen: "They must have known these sobering truths because time and again they and their fellow human beings must have come slap up against the evidence for them" (Humphrey 1995, p. 54). Mechanisms of psychological self-correction will inevitably steer believers away from demanding

divine acts of the latter kind, and instead steer in the direction of modes of action that are more or less indistinguishable from the natural course of events.

This is not to say that religious believers who have 'learnt' not to pray for robust divine action will never find themselves in a situation in which their prayers remain apparently 'unanswered'. If the causal closure thesis of nature is correct, even prayers for subtle modes of divine action will appear to be answered at most occasionally. Indeed, the history of prayer itself bears witness to the simmering doubts about its efficacy, as Zaleski & Zaleski noted:

> [T]he sheer abundance of devout tracts exhorting the faithful to pray often, pray fervently, and pray with confidence in achieving desired results suggests that belief in the efficacy of prayer has always needed some degree of shoring up. (Zaleski and Zaleski 2005, p. 333)

The frequency of apparent confirmation will depend on the probability function of the event type in question (e.g. spontaneous remission of a disease, rainfall in a dry season), but in any case, psychological research has established that beliefs may become entrenched even on the basis of a small number of apparent confirmations. People are prone to *confirmation bias* (Nickerson 1998), which means that they pay attention to and remember confirmations of a favored hypothesis, while they quickly discard or explain away adverse evidence. More specifically, the psychological literature on cognitive dissonance and motivated reasoning (Kunda 1990; Aronson 1992; Tavris and Aronson 2008; Melton 1985) suggests that, when firmly held beliefs are confronted with apparent failure, people rely on a repertoire of rationalizations. As Barrett noted in a later discussion of his prayer study (Barrett 2004), the whole idea of addressing a request to a whimsical supernatural agent automatically suggests some ways of accommodating failure: "prayer commonly assumes the possibility that a request could be approved, denied, or put off until a later date" (Barrett 2004, p. 71). In such a setting, a prayer that appears to be failed may in fact be granted "on a different timetable" or just be denied for some good reason that we mortals cannot fathom. In this way, as Barrett notes, "negative evidence rarely threatens belief in God" (2001, p. 74). Other rationalizations are possible for occasional failure: "My prayer did not fit into God's plan" – "Too much answered prayers would spoil us" –– "My praying ritual was not performed correctly" – "God cannot be coerced through prayer" – "the purpose of prayer is 'to construct the soul, not to instruct God'" (Augustine). If believers are highly committed to their faith, we can expect them to be motivated to maintain belief in the goodness and

omnipotence of God (e.g. Kushner 1981), and to explain away failure by what Evans-Pritchard has termed "secondary elaborations" (see below).[99]

The persistence of the confirmation bias ensures that even a small number of apparent successes may succeed in outweighing the instances of apparent failure (Barrett 2004, p. 74). In the case of prayers for robust divine action, however, the causal closure thesis leaves little or no room for apparent confirmations. Thus, if we confine ourselves to personal experiences regarding the efficacy of prayer, the point is not that disappointment and cognitive dissonance will never arise for those who prefer subtle modes of action on God's part, but that those who demand robust actions will *always* encounter failures that occasion the need for some rationalization, and will never encounter occasional personal successes to compensate for the failures. The confirmation bias needs *some events* to be biased towards.

12.4.2 Folk physics

Besides this form of psychological adjustment, there may be purely cognitive factors at play in the preference for subtle modes of causation. Research in developmental psychology suggests that, from early infancy, humans possess an intuitive core knowledge about spatio-temporal objects, which is sometimes designated as *folk physics*. Looking time experiments with children reveal a number of such implicit assumptions: 1) objects move as bounded and discrete wholes (*cohesion* principle), 2) objects move along continuous and connected paths (*continuity* principle), 3) objects do not interact at a distance (*contact* principle) (Spelke 1994; Spelke and Kinzler 2007). It is not surprising that natural selection has endowed us with an intuitive grasp of these basic spatio-temporal principles, because the physical environment on which our ancestors depended for survival really does obey them, at least at the scale of medium-sized objects.

Is it plausible that people find robust forms of supernatural causation counterintuitive because of their violating deeply-engrained folk physics? For example, the instant materialization or displacement of desired objects before my very eyes would violate the principle of continuity. Or, to give another example, if a physical obstacle would suddenly disintegrate as a result of my prayer to remove it, this would surely violate the principle of coherence. On the other hand, we see that precisely such breaches in the

[99] In the wake of scientific evidence on the inefficacy of prayer, many liberal theologians have relinquished the idea of divine interventions in the natural world altogether. Notably, since the prayer-gauge debate in the 19th century, theologians have begun reinterpreting the value of prayer on a purely moral and spiritual level, purging it from any form of miraculous interventionism. (Mullin 2008)

fabric of our natural world by supernatural agents often form the subject of religious narratives, legends and holy writ. The Old Testament stories about Yahweh parting the red sea or tearing down the walls of Jericho are clearly violations of the basic principles of folk physics. Furthermore, many religious traditions contain stories that seem to exploit violations of precisely the intuitive principles mentioned above. For example, Christian monks and saints are often believed to possess the ability of bilocation, i.e. being physically present at two different places at the same time (Nickell 1993, pp. 216-219). Similar stories are to be found in Buddhism, Hinduism, paganism, shamanism and many other religious traditions. Likewise, popular stories about psychokinesis and action at a distance in religion and magic (e.g. transmuting water into wine) are clear violations of both the contact and coherence principle.

Indeed, researchers in the Cognitive Science of Religion have argued that religious representations are successful precisely because they violate intuitive expectations about the ontological categories to which they belong. To be more precise, religious representations have to be "minimally counterintuitive" (Boyer 1994; Atran 2002) to achieve a cognitive optimum, which means that they display a small number of violations against a background of intuitively expected properties. Supernatural concepts that violate our intuitions on many different fronts typically strain our cognitive resources too much and are more difficult to remember and process.

If intuitive violations are precisely what makes religious representations salient and attention-grabbing, it seems unlikely that robust supernatural causation *per se* is cognitively burdensome. Firework displays by God is all right as long as it is in the context of distant hearsay and religious mythology, but it is hardly ever the subject of personal experience or first-person eyewitness testimony (try multiplying bread or parting a sea) (see also for example Evans-Pritchard 1965, pp. 195-201). Religious believers are well aware that such supernatural feats are exceptional, and cannot be expected to occur on a regular basis. Although believers have no difficulty in conceiving of such past events, it is plausible that folk intuitions lower their expectations when it comes to personal prayer in ordinary situations. The mechanisms of psychological self-correction which we have described can then be viewed as further enforcing this cognitive preference.

12.5 Epidemiology of representations

Up to this point, we have described the formation of prayer beliefs and practices on the level of individual psychological mechanisms. We can now take this approach one step further and outline the ways in which larger cultural trends emerge from these psychological processes and intuitions. In Sperber's epidemiological model of culture (Sperber 1990, 1996), which has had a formative influence on the cognitive science of religion, our shared cognitive make-up acts as so many constraints through which the dissemination of representations is channeled. Small selection pressures in the transmission of beliefs, aggregated over many transmissions, will give rise to larger cultural trends. Liénard and Boyer (2006) and McCauley and Lawson (2002) have applied this approach to religious rituals, and we can now apply it to belief patterns about supernatural causation. Assuming that the facts of nature are the same for all, it is reasonable to expect that the disillusion with robust modes of divine causation will become part of the collective experience of religious communities (see also Barrett 2004, pp. 70-74). Belief in the present-day feasibility and reliability of robust supernatural feats, as opposed to distant hearsay and historical narratives, are unlikely to be thriving if they are to compete with beliefs in subtle divine causation.

There are at least four different ways in which representations about supernatural causation may be passed on and disseminated in a religious community: (i) new members are explicitly instructed by older members regarding the things they may reasonably expect from supernatural beings, (ii) believers take familiar accounts of answered and unanswered prayers as templates for the ways people engage in interaction with supernatural beings, (iii) new members assimilate theological rationalizations from religious doctrine for why supernatural beings act in certain ways but not in others, (iv) they are initiated in collective prayer practices and adopt the habits of older members.

Naturally, this sketchy and general outline of pathways of cultural assimilation leaves many questions to be answered. In the context of this paper, however, we merely want to draw attention to the point that not every new member of a religious community need go through the same stages of personal disappointment of the kind described by Humphrey above. By means of explicit or implicit religious instruction, cultural selection forces may supplement and reinforce intrapersonal selection to the same effect. To put it bluntly, I may refrain from asking spectacular displays by God either because I have myself experienced the disappointment following such requests, or because others have instructed me not to make them.

To illustrate this effect of accumulated experience on religious beliefs, consider a second study by Barrett, in which he uses the same eight scenarios but substituted God with a "comparably endowed super-agent" (Barrett 2001, p. 264). Subjects were asked to

consider a futuristic supercomputer (Uncomp) with comparable God-like powers, but "components physically located all over the earth" The term "praying" was replaced by "asking for". Barrett found that subjects in the Uncomp group did not have any preference for either mode of action and explains this in terms of the non-locality of Uncomp. Our account suggests a different explanation: subjects had no experience with the fictitious Uncomp, whereas of course practicing believers had plenty of experience with directing prayers to God. Whereas religious believers have had ample opportunity to adjust their conception of God's *modus operandi* on the basis of experience, and have also been exposed to a long tradition of religious believers with similar experiences, none of this holds for Uncomp. A futuristic computer with God-like powers is simply a fictitious character invented by the experimenter, which is quite a different matter from God (at least for those who believe that God is not a fictitious character). Thus, subjects have no reasons not to accept the stipulated omnipotence of Uncomp at face value, which explains their relative lack of preference for either mode of action.[100]

Note that there may be still other ways in which certain praying practices (or patterns of belief) may be conducive to a process of self-validation, and thus possess an "advantage" in terms of cultural dissemination. First, belief in the biological healing powers of supernatural beings may achieve cultural success in virtue of the fact that genuine faith on the part of the person afflicted may engender a placebo effect, the result of which may afterwards be attributed to God's help. Second, if I pray to God to give me the strength to face a difficult ordeal, my act of praying and my faith in God may itself increase my self-confidence and reduce stress levels, resulting in a form of self-fulfilling prophecy. In addition, we have to take into account a self-validating selection effect regarding the people who, other things being equal, are still around to recount their miraculous healing or rescue, as Nicholas Humphrey's Law of the Efficacy of Prayer makes clear: "In a dangerous world there will always be more people around whose prayers for their own safety have been answered than those whose prayers have not."[101]

---

[100] Another and even more simple explanation is that Uncomp is explicitly described as having components "*physically* located all over the earth" (Barrett 2001, p. 264, our italics), whereas God presumably is not to be thought of as a physical entity. This difference by itself may account for the greater preference for physical causation in the Uncomp group.

[101] The "law" was posted on the EDGE Question Center, 2004, retrieved from http://www.edge.org/q2004/q04_print.html

12.6 Different forms of supernatural causation

Thus far, we have mainly focused on the Christian tradition, describing how the psychology of supernatural causation plays out in the practice of petitionary prayer directed to a single divine being. In cultures with very different supernatural beliefs and religious practices, however, we might expect to find similar preferences for subtle modes of action. An interesting case study is provided by Evans-Pritchard's landmark anthropological investigation of magic and witchcraft among the Azande in Sudan. The Azande believe that some members of their community are witches and possess the mystical power to injure and kill other individuals, to harm their crops, to make houses collapse etc. Indeed, the Azande invariably attribute death, disease and other forms of misfortune to the malignant action of witches, thus making no distinction between different modes of causation (psychosocial vs. biological or mechanical). As Evans-Pritchard noted, however, the Zande people are certainly not unaware of the natural causes leading to such events. Interestingly, they believe that the mystical cause of witchcraft acts *through* a chain of natural causes, making a subtle contribution as a "co-operating cause" (Evans-Pritchard 1965, p. 72).

For example, when the Azande attribute the collapse of a mud house to witchcraft, they know very well that, as it happens, a colony of termites has been gnawing through the pillars of the house and undermining its foundations. Although they accept this natural cause of the event, they insist that only witchcraft explains why *this* particular house collapsed at *that* particular moment. Again, we see that the actions of the supernatural agent – in this case, a human being endowed with supernatural powers – are believed to contribute as a partial cause to some event, which renders the supernatural *modus operandi* subtle and virtually imperceptible. As Evans-Pritchard notes, "[t]he attribution of misfortune to witchcraft does not exclude what we call its real causes but is superimposed on them and gives to social events their moral value" (Evans-Pritchard 1965, p. 73). Why don't witches just make the house collapse at one stroke, instead of acting indirectly and through the efforts of termites? On the other hand, how would the belief that witches may destroy houses at a single stroke fare in the Zande community, compared with the belief that they consistently employ the services of ants or other seemingly natural causes?

The Azande also regularly consult oracles about the various threats of witchcraft and about the courses of action to take in their life. In one of the most respected Zande oracles, a poison is administered to a fowl, following a number of elaborate preparations, and a question is put to it. The oracle is believed to provide a yes/no answer depending on whether the fowl survives or dies. After taking note of a series of such divinations, Evans-Pritchard notes that there is no objective way to predict whether or not the fowl will die, given the amount of poison or the size of the fowls. To

all intents and purposes, the fate of the fowls is a matter of chance.[102] As we have argued before, precisely such stochastic processes, to which human beings have no epistemic access, are psychologically optimal for belief in supernatural causation. Because they allow for regular confirmations, they are conducive to self-validation, a process that is augmented by the kinds of questions and the way they are typically phrased. Among the Azande, as we saw in section 6.3.2, apparent oracular failures are explained away or reinterpreted by "evasive secondary elaborations" (Evans-Pritchard 1965, p. 319) that are provided for by the belief system itself: improper preparation of the poison, the violations of taboos, interference of witchcraft or evil magic, refusal of the oracle to give the right answer. In section 5.3.2 we have argued that, when the causal relations in a belief system are underspecified and the effects are ambiguous, this engenders subtle forms of inferential circularity, rendering the belief system impervious to adverse evidence.

12.7 Discussion

In many religious traditions, in particular the three monotheistic faiths, supernatural beings are conceived as very powerful or even omnipotent agents who can act in any way they like. As the theory of theological (in)correctness has revealed, however, religious believers typically do not respect such counterintuitive theological doctrines when they are engaged in everyday religious practice, even though they may endorse it when explicitly questioned and given some time to reflect. The concept of theological correctness has sparked a renewed interest in the psychology of petitionary prayer, which is still a relatively unexplored domain. Barrett's study on prayer is a very welcome exception, but his account suffers from a conceptual problem: the characteristic distance range of psychosocial as opposed to mechanistic action is largely similar, which complicates his argument about an intuitive preference for supernatural causation of the former kind.

On an empirical level, the hypothesis defended by Barrett sits uncomfortably with the fact that beliefs in supernatural beings acting mechanically and biologically are pervasive all over the world (Johnson 2005; Murdock 1980). We have set up a different

[102] The same point holds for the other oracles in the Zande belief system (the behaviour of termites and the manipulation of some rubbing-board) (Evans-Pritchard 1965, pp. 352-386).

account of *theological incorrectness* and framed it in the context of our best scientific knowledge about the efficacy of prayer and the evidence for the causal closure thesis of nature. Viewed in that light, we saw that people will developed 'preferences' for supernatural interventions that are subtle and indistinguishable from the natural course of events. These different modes of action need not be cognitively represented as such by believers – as indeed believers do not care much about modes of supernatural action – but they may emerge from the psychological mechanisms of self-correction and cultural selection forces.

Our account suggests that people stray from orthodox theology not only because of the way their minds work, but also in virtue of how the world looks like (and the interaction between both). In an epidemiology or religious representation, both our innate cognitive make-up and the structure of external reality impose selective pressure on representations. In particular, we argued that, if the causal closure thesis is correct, and given that the causal structure of our world is partly inscrutable, beliefs in subtle and unascertainable modes of supernatural causation will achieve a *cognitive optimum* (Boyer 1994; Slone 2004; Boyer 2001), because they are more susceptible to occasional 'confirmation' and less vulnerable to repeated disconfirmation. In line with Chapter 11, we have argued that psychological mechanisms of self-correction and basic principles of folk physics will steer believers away from beliefs in robust and palpable forms of divine action, an effect that is further enforced by cultural selection. Further research may extend this approach to attribution of supernatural agency in general (i.e. not related to prayer), for example in terms of divine punishment and retribution for moral transgression. It will be interesting to know whether, in such cases as well, believers are fond of Gods that move in mysterious ways.

Conclusion

'I can't believe that!' said Alice.
'Can't you?' the Queen said in a pitying tone. 'Try again: draw a long breath, and shut your eyes.'
Alice laughed. 'There's not use trying,' she said: 'one can't believe impossible things.'
'I daresay you haven't had much practice,' said the Queen. 'When I was your age, I always did it for half-an-hour a day. Why, sometimes I've believed as many as six impossible things before breakfast.' – Lewis Carroll, Through the Looking-Glass

Where pseudoscience goes wrong

After a long journey through the hinterland of science, we can retrace our steps to draw a number of general conclusions. In the first part of this thesis, we challenged the common philosophical wisdom on the supernatural and the problem of demarcation. Intelligent Design Creationism (IDC) fails to live up to its scientific pretension not because it is involved with a supernatural designer *per se*, but because it exhibits general telltale signs of pseudoscientific discourse: IDC theorists refuse to flesh out their design hypothesis and prefer to make use of convenient immunizations that make the theory impervious to criticism; the concepts devised by IDC advocates are ill-defined and suffer from equivocations in a way that transforms their central argument into a moving target; the theory is too vague to allow for specific predictions and to achieve any form of genuine explanatory unification; the bulk of IDC literature consists of purely negative

arguments, launched with the sole purpose of distorting evolutionary theory and sowing doubt among the public at large; etc.

Because IDC invokes supernatural entities, and because it is guilty of all those pseudoscientific sins, critics of the theory have conflated both issues and have rushed to the conclusion that one is intrinsically connected to the other. But this seems to be an example of what psychologists call the *representativeness bias*, i.e. a tendency to assume that the data which happen to be available are representative of the category to which those data belong. Because we have become so accustomed to supernaturalists committing these pseudoscientific sins, and because we have grown weary of creationist hypotheses that, when push comes to shove, boil down to "God did it and His ways are mysterious", we can hardly imagine that any other supernatural hypothesis would be viable. But this is mistaken. In the world *we* happen to live in, there is simply no way to make sense out of the living world in terms of supernatural design *except* by retreating to the claim that God did it and His ways are mysterious. If there was even the remotest possibility of fleshing out the supernatural design hypothesis in a way that did not run slap up against the facts, creationists (and theists in general) would surely be the first to spread the gospel. Only by highlighting in what ways supernaturalists *might* have succeeded does it transpire how miserably they have failed.

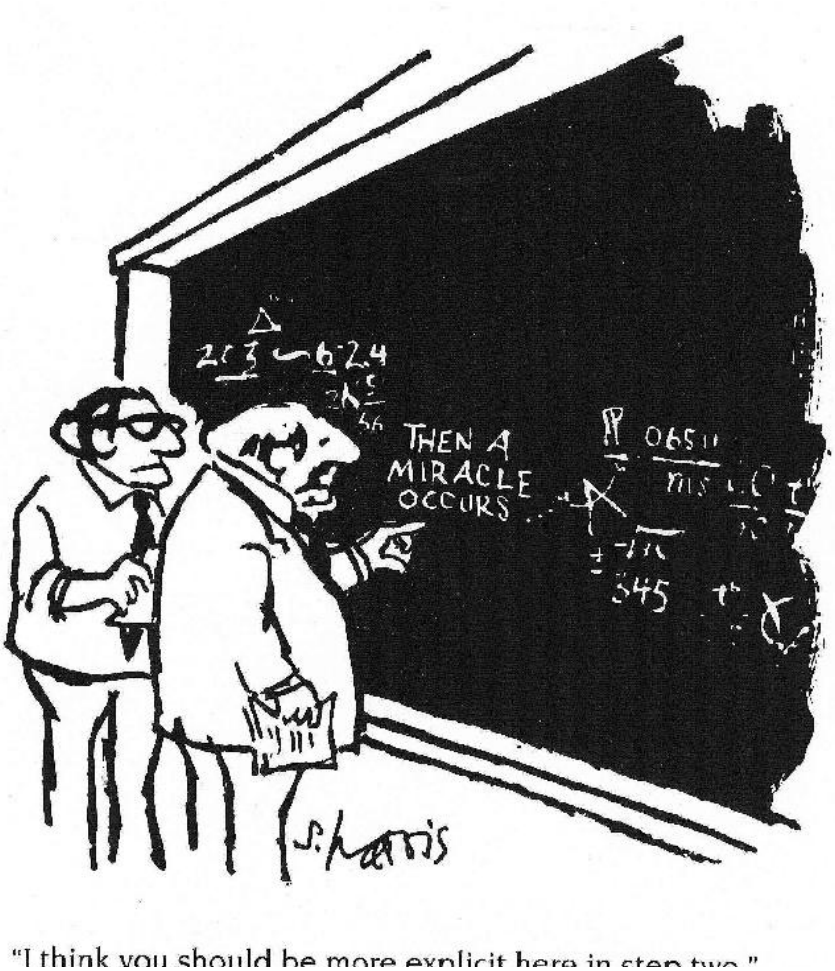


Figure 7 A *status quaestionis* of the IDC literature. Are we dealing with an intrinsic problem associated with supernatural hypotheses?

A corollary of our view is that science is in principle open to extraordinary evidence for the supernatural, but that its current verdict on all supernatural hypotheses proposed thus far has been uniformly negative. In fact, our claim is more modest than what defenders of Intrinsic Methodological Naturalism (IMN) claim: as we see it, there is no compelling reason why supernatural interactions with the natural world, in the sense we have defined those terms, would be closed off to scientific investigation. If anything, the burden of proof is on those who exclude the supernatural from the purview of science *a priori*. They need to come up with an airtight and nontrivial definition of natural and supernatural causes that explains why the latter can never be amenable to scientific investigation. A defender of Provisory or Pragmatic Methodological Naturalism (PMN) can stick to a working definition of the supernatural, and is making no stronger a claim than that at least *some* of the phenomena satisfying that definition are amenable to scientific investigation.

The only viable philosophical route towards discarding supernatural causation from science *a priori* is either to set up an argument establishing that the very idea of entities

beyond space and time is impossible because it rests upon a category mistake, or that the idea of causal interaction between the natural and supernatural realms cannot be made coherent (Fales 2010). Once we accept the metaphysical possibility and possible causal efficacy of the supernatural, however, as most defenders of IMN make an explicit point of emphasizing, the methodological strictures of IMN no longer make any sense.

The principle of IMN is based on a confusion between accidental features of the supernatural hypotheses that are on offer and intrinsic problems with the supernatural, but as we have argued in Chapter 2, it is also for a large part prompted by the political motivation to install a truce between science and religion. If science is perceived as being in irresolvable conflict with religion, so the argument goes, and if a choice between both magisteria is forced upon the public at large, then an overwhelming majority will side with the forces of darkness.

Although we appreciate the political rationale behind this strategy, we have argued that ruling the supernatural out of science by philosophical fiat is counterproductive, philosophically incoherent and misleading. Most importantly, it plays right into the hands of IDC proponents. Always eager to cast evolutionists in the role of closed-minded and dogmatic materialists, IDC advocates have presented the strictures of IMN as a symptom of metaphysical prejudice against supernaturalism. Indeed, detractors of IMN view the principle as some kind of *immunizing strategy* of scientific naturalism, the use of which is the very *antithesis* of the scientific attitude, as we have seen in Chapter 5. Without a commitment to metaphysical naturalism, and in the absence of a sound methodological rationale, the *a priori* rejection of supernaturalism seems unfair indeed. In addition, the accommodationist strategy of dealing with religious dogmatism threatens to backfire on science, making some of the most solid arguments for evolution unintelligible, and fostering a misleading image of major scientific achievements. The position we have defended throughout the first part of this work (PMN) does not suffer from these philosophical and strategic complications:

(1) At least some supernatural hypotheses entail empirically detectable consequences, and are thus open to scientific investigation.

(2) As a matter of principle science is open to paranormal and supernatural claims, although the prospects of such hypotheses being borne out are extremely dim, and no sensible scientist wants to waste much time and resources investigating these claims.

(3) The absence of God (and other supernatural entities) from the corpus of scientific knowledge is not a result of metaphysical prejudice or methodological exclusion, but of contingent scientific failure. The deck was not stacked against the Almighty to begin with, quite to the contrary.

(4) Darwin's empirical case against special creation by supernatural intervention is perfectly respectable and was a major factor in overturning the program of natural theology.

(5) Evolutionary theory does not inexorably lead to atheism, but there is no avoiding that it has undermined one of the most compelling – one is tempted to say the only respectable – argument for theism.

Chapter 7 is an illustration of the way we think the scientific pretensions of theories like IDC should be confronted: not by appealing to the naturalistic strictures of science and settling the matter by analytical definition, but by getting down into the trenches and confronting the conceptual and empirical shortcomings of IDC head-on. At the same time, Chapter 7 clears up the conceptual fog around irreducible complexity and shows why there have been several, apparently contradictory lines of criticism against the argument. In this case, luckily, this is not due to shaky philosophical doctrines on the part of some IDC critics, but it is entirely to blame on Michael Behe's own vacillations in his definition and use of irreducible complexity.

Incidentally, such a direct line of attack has also been pursued by most defenders of IMN in tandem, including by Judge John E. Jones in his ruling on the Kitzmiller vs. Dover case. As we have argued in section 2.4.5, however, this makes their case against IDC open to the charge of incoherence, or at least of superfluousness. Why bother to put forward empirical arguments against IDC if its claims can be safely ruled out from the outset?

For many scientists, philosophers and liberal theologians, IMN embodies the modern *modus vivendi* between science and religion much in the same way as the principle of Non-Overlapping Magisteria did for Stephen Jay Gould and his followers.[103] By keeping scientific inquiry at bay and relegating the supernatural to a wholly separate domain, defenders of IMN have tried to safeguard a place for religion to which science cannot reach, but they have done so at the expense of philosophical integrity.

How to deal with pseudoscience

The first chapters of this doctoral thesis drive home the point that it is one thing to correctly recognize a pseudoscience for what it is, but quite another to find the appropriate philosophical and scientific weapons to confront it. We cannot afford a casual and complacent attitude towards pseudoscience, notably when it comes to the latest spate of creationist protestations against evolutionary theory. The luxury of being "obviously" on the right side of the fence may have the perverse effect that one does not even bother to support one's position with the best rational arguments. But even a theory that is wrong beyond any reasonable doubt deserves to be taken seriously and

[103] Interestingly, in his latest book our colleague Massimo Pigliucci (2010) offers an excellent rebuttal of Gould's NOMA solution (Gould 1999), but he fails to notice that the IMN position adopted in the ruling of Judge Jones is itself more politically convenient than philosophically accurate (Boudry 2010).

confronted with careful arguments, particularly when its defenders exploit every opportunity to cast themselves in a victimized role and complain that they are not being given a fair hearing.

This is not to say that the strategy as we propose it does not face some difficulties of its own. To even engage with creationist opposition at all may leave uninformed bystanders with the impression that there is a legitimate scientific debate going on, or that evolutionary scientists are forced on the defense. It is even more difficult to explain why some hypotheses are no longer seriously investigated at all, except insofar as they still enjoy wide support from the public at large (the same perverse effect is apparent in the case of public scares about, for instance, the link between cell phone radiation and cancer, or between vaccination and autism). Evolutionary theory has been in the indisputable scientific pole position for so many decades now, and any respectable alternative has been discredited since so long ago, that the decisive historical reasons for its victory have faded into the background. The evidence in favor of evolution by natural selection *per se*, and the damaging empirical and conceptual problems facing the different versions of natural theology and vitalism that were entertained until the end of the 19th century, are simply no longer part of the cutting-edge scientific debates nowadays. Were it not for the continuing resistance from the creationist fronts, the need for the current spate of books expressly aimed at didactically laying out the evidence for evolution would be much less pressing (Coyne 2009b; Dawkins 2009; Pigliucci 2002; Scott 2009; Sarkar 2007).

We cannot ignore the politically powerful and well-organized creationist movement, but neither can we afford to use philosophically dubious doctrines for countering their arguments. Creationists are proficient at poking holes in evolution even when such holes exist only in their imagination, so *a fortiori* we should not allow them an opportunity to discover a *real* hole. As Philip Kitcher put it:

> Creationist literature is especially creative in its misunderstandings, in pinning odd views on the opposition and then playing "Gotcha!" [...] creationists can manufacture spurious problems faster than evolutionists can unmask their sophistries. (Kitcher 2007, p. 71)

Likewise, as we have argued in Chapter 3, we should not be using outdated and misleading metaphors that unwittingly fuel design intuitions and play right into the hands of creationists. It is of little use if we succeed in keeping creationism out of the classroom if we still allow the use of covertly creationist metaphors in biology textbooks for no sound scientific reason. Of course, this does not mean that we should do away with design talk in biology altogether, as the notion of design is vital to any discussion of biological adaptation and functionality (Dennett 1996). However, we should be cautious in using detailed mechanistic metaphors that stretch the design metaphor beyond its limits. A careless use of mechanistic and informational metaphors is ill-

advised especially in the light of current research in cognitive and developmental psychology, which suggests that persistent misconceptions about evolution are grounded in the cognitive architecture of our mind. For example, experiments with young children reveal their proneness to what has been termed "promiscuous" teleology: children intuitively prefer explanations in terms of purpose and intentional design over simple causal explanations, even for natural phenomena where no design is involved. Deborah Keleman has even suggested that children are "intuitive theists" (Kelemen 2004; Kelemen and Di Yanni 2005), in the sense that they are intuitively disposed towards viewing the natural world in terms of artificiality and teleofunctional design. If indeed there are deeply-engrained cognitive limitations and biases that hinder the acceptance of evolutionary theory (Blancke, Boudry et al. in press), then *a fortiori* we have to be cautious in employing metaphors that bolster these intuitions about design and artificiality.

Method in madness

This last point about our cognitive susceptibility for design reasoning can be viewed as a prelude to the third part of this dissertation, in which we draw upon a wide array of cognitive research to explain why belief systems that display certain structural features are so pervasive and culturally successful. Before outlining this argument in an epidemiological framework, we have offered a purely theoretical analysis of pseudosciences in terms of immunizing strategies and epistemic defense mechanisms. In addition to the examples offered in Chapter 5 and Chapter 6, which were drawn from a wide range of different domains, we have elaborated on a number of specific case studies in the subsequent chapters: the concept of irreducible complexity in IDC, the theory-internal explanations for obscurity and conceptual incoherence by Lacan and his apologists, and the methodological and conceptual complications in Freudian psychoanalysis.

From a philosophical point of view, epistemic defense mechanisms are particularly interesting, because they lend the system of beliefs in question a kind of self-validating rationale. In Chapter 11, we have drawn upon research on cognitive dissonance, belief perseverance and rationalization to construct an epidemiological framework that explains the cultural success of these self-validating beliefs. By outlining the free-floating rationale underlying these belief structures, we have shown that there need not be much conscious deliberation involved in the construction of their protective shield,

and that cultural selection forces plus mechanisms of psychological self-correction may spontaneously give rise to such constructions. Our analysis has shown that, contrary to the prevailing opinion about the *fragility* of weird belief systems, there is substantial method in madness. Not surprisingly, intelligent people are anything but invulnerable to irrational beliefs, because they above all are proficient at rationalization, *ad hoc* reasoning and belief perseverance in the face of difficulties.

The third part of this thesis was explicitly framed in the epidemiological approach to culture developed by Dan Sperber, and the cognitive and selectionist turn in the scientific study of religious beliefs. The general question put forward by Sperber and researchers in CSR is why certain cultural representations are more successful than others. By figuring out the cognitive and epistemic constraints through which the dissemination of cultural representations is channeled, through small incremental effects, we are able to explain larger cultural trends. As Liénard and Boyer have put it:

> In cultural transmission [...] very small effects aggregated over many cycles of transmission are sufficient to create massive trends. (Lienard and Boyer 2006, p. 824)

One possible objection against our approach is that we stack the deck against certain theories by calling them "weird" to begin with, as if their implausibility has already been established beyond reasonable doubt, and only the question of their pervasiveness remains to be answered. How indeed do we decide that a belief system is "weird"? Is that assessment not based on a subjective notion of weirdness? We do not think that much hinges on this terminology, as the epistemological analysis of these belief systems must stand on its own merits. If our analysis is correct, the beliefs and belief systems in question propagate not in virtue of their standing in certain relations to reality, but rather due to the way in which they are enmeshed with other beliefs and to the way our minds operate. The self-validating rationale exhibited by these beliefs ensures that they are not constrained by the way the world actually is, and this *allows* them to take on weird contents.

The most parsimonious view about such belief structures is that they propagate for no other reasons that the ones just mentioned. In other words, establishing their resilience and self-validating nature does not *equate* with demonstrating that they are wrong, but it is good circumstantial evidence for that thesis. Although it is by no means impossible for such beliefs to be true, their truth would be a matter of sheer luck or coincidence. In Chapter 10, we have belabored that point in the context of Freudian psychoanalysis, arguing that the epistemic structure of Freudian doctrine, together with its flexible methodology, ensures that there are no significant constraints on theory formation in psychoanalysis. Likewise, in Chapter 11, we have argued that conspiracy theories and certain paranormal belief systems, simply by virtue of their internal

structure and their rich conceptual resources, may have a shallow ring of plausibility that is completely disconnected from their relation to any objective facts out there.

An important theme that is developed throughout the second and third part of this work is that the study of pseudoscience and irrational belief systems sheds light on the nature of human rationality itself. Just as optical illusions and visual impairments are informative about the fabric of visual perception, the study of irrational belief systems allows us to gain better access to the intricacies of human cognition. Though he has not often been quoted approvingly in this work, Sigmund Freud has captured this principle with a beautiful metaphor:

> If we throw a crystal to the floor, it breaks; but not into haphazard pieces. It comes apart along its lines of cleavage into fragments whose boundaries, though they were invisible, were predetermined by the crystal´s structure (Freud 1964, p. 59).

The idea that irrationality is the outcome of a special mode of thinking, disconnected from normal reasoning functions, is being progressively abandoned by cognitive scientists and psychologists. Rationality and irrationality are better conceived as opposite sides of the same coin (Talmont-Kaminski 2007, 2008). The present study on pseudoscience vindicates this more sophisticated view on rationality and its dark counterpart.

(1) On closer inspection, many forms of persistent irrationality are predictable side-effects of reasoning heuristics that are quite efficient in the appropriate contexts for which they evolved.

(2) Believers are not the simple-minded fanatics for which they are often mistaken, and intelligent people are anything but invulnerable to irrational belief systems. Quite to the contrary, there are some compelling reasons to think that intelligent people are *more* prone to irrational thinking than others.

(3) Popular pseudoscience and other weird belief systems are more complex and less fragile in the face of adverse evidence and criticism than is commonly assumed. The assumption that they manifest distinct modes of irrational reasoning, disconnected from and opposed to rational reasoning processes, has not been borne out.

In Chapter 12, we have applied the cultural selection framework developed in section 11.6 to belief formation about the *modus operandi* of supernatural beings and the efficacy of petitionary prayer. Assuming the causal closure thesis of nature, which we have argued for on *a posteriori* grounds in Chapter 1, we noted that it is highly unlikely that belief in the efficacy of prayer is pervasive because of the actual occurrence of supernatural intervention, psychosocial or otherwise. A number of psychological, cultural and cognitive explanations are available for tackling the problem of explaining the persistence of belief in "theo-mundane causation" (Fales 2010). We have developed an epidemiological framework that explains not only why such beliefs may nonetheless be thriving, but also why they take on specific, culturally successful forms. On the basis

of a number of cognitive considerations, and the nature of observational confirmation, we have argued that beliefs in modes of divine action that are subtle and indistinguishable from the natural course of events are most attractive, because they are more conducive to occasional confirmation and more resilient to persistent failure.

The demarcation problem

The debate about methodological naturalism, the demarcation problem and appropriate strategies for dealing with pseudoscience is reminiscent of a philosophical controversy in the 1980s between, among others, Michael Ruse, Larry Laudan, Philip Quinn and Barry Gross (reprinted in Ruse 1996a; Pennock and Ruse 2009). In 1981, Michael Ruse appeared on the witness stand of the McLean v. Arkansas trial on the teaching of "scientific creationism" in biology classrooms. In his expert testimony, Ruse offered a set of demarcation criteria for distinguishing real science from pseudoscience such as creationism, which were eventually taken up by judge William Overton in his ruling:

(1) It is guided by natural law.
(2) It has to be explanatory by reference to natural law.
(3) It is testable against the empirical world.
(4) Its conclusions are tentative, i.e. are not necessarily the final word.
(5) It is falsifiable.

In a short article following the trial, Laudan (1982; reprinted as Laudan 1996b) takes issue with Ruse's involvement in the trial and his proposed set of demarcation criteria, arguing that it is full of "woeful fallacies" (1996b, p. 354) and presents a "false stereotype of what science is and how it works" (p. 355). In a second and influential essay (Laudan 1983; reprinted as Laudan 1996a), Laudan takes his critique one step further, proclaiming that *all* past philosophical attempts to separate science from non-science have failed, and in the end declaring the death of the demarcation project (Laudan 1983; reprinted as Laudan 1996a). In the first part of this doctoral thesis, we have defended an approach for dealing with pseudoscience that is in some sense congenial to Laudan's proposal. As Laudan has dismissed the term pseudoscience as a "hollow phrase" which any honest philosopher should get rid of, whereas we have employed it extensively throughout this work, it is instructive to revisit the philosophical aftermath of the McLean v. Arkansas trial and Laudan's critique in particular. This allows us to clarify

where exactly we follow Laudan's lead and where we part ways with him in, as it appears, perfectly diametrical directions.

Laudan is right that Ruse's definition of science leaves a number of loopholes for creationists to exploit, and that it does not reflect a consensus among philosophers of science. In particular, we follow his argument that creationist claims are not intrinsically untestable (see 1.4.5), and that it is important to distinguish between establishing the existence of a phenomenon and explaining it in a lawlike way (see 1.4.2 and 1.4.3). Finally, we sympathize with Laudan's conclusion that, rather than trying to find a silver bullet that will put an end to creationist nonsense for once and for all, we should "confront their claims directly and in piecemeal fashion by asking what evidence and arguments can be marshaled for and against them" (Laudan 1996b, p. 354). However, Laudan's conclusion on Ruse's involvement in the trial and especially his obituary of the demarcation project are premature. As for his indictment of Ruse's testimony, Laudan seems to ignore that Ruse was standing as a witness in a trial on biology curricula and the separation between church and state, not in a academic conference on philosophy of science. Ruse's list is a rough and ready demarcation that is intended to suit specific legal purposes, not an attempt to finally resolve the demarcation problem. Although it is perfectly respectable to quarrel with some of the proposed criteria, Laudan extends remarkably little charity to Ruse and seems largely insensitive to the legal and constitutional issues that were involved in the case (Gross 1983; Pennock 2009; Ruse 1996b). In the context of this work, however, Laudan's sweeping pronouncement on the demarcation problem is more interesting.

First, we should note that a large part of Laudan's critique of demarcationism comes down to nitpicking and pointing out irrelevant complications that have no epistemic import. For example, Laudan lists several forms of knowledge which, although certainly deserving epistemic warrant, are sociologically not recognized as "science" (e.g. singular historical claims, football strategies, literary theory, etc.). Laudan seems to think that this is most damning to the conception of science versus pseudoscience, but we beg to differ. For a number of understandable reasons, disciplines such as history and literary theory are not commonly included under the umbrella of "science" properly speaking, although any reasonable person would acknowledge that there is reliable knowledge to be found in these domains. For example, there is a very interesting philosophical discussion going on about the distinctive nature of historical science, as opposed to experimental science (e.g. Cleland 2002), but this is not what keeps the philosopher of pseudoscience awake at night. The distinction that interests the demarcationist is not that between forms of reliable empirical knowledge that are sociologically and institutionally classified as science and those that are not, but between those forms of empirical knowledge that rightly deserve epistemic credit and those that fail to live up to their pretensions. If one acknowledges that historical inquiry can produce knowledge that is on a par with good science, the question is whether and

how we can distinguish this from what David Aaronovitch has termed "voodoo history" (Aaronovitch 2010). To the extent that the received history of the second world war deserves to be called scientific, Holocaust denial certainly does not. Does anyone want to quarrel over whether Holocaust denial ought to be labeled as either voodoo history or pseudoscience, or maybe pseudo-history or failed history?

A more fundamental problem facing Laudan's diatribe is that, even if it is futile to construct a set of necessary and sufficient conditions for defining science that will survive close philosophical scrutiny, that does not mean that the demarcation problem has ceased to exist. If anything, it is Laudan who clings to an old-fashioned demarcationism, because he is unable to envisage any other way in which the distinction between science and pseudoscience could be rendered intelligible.

First, it may be more fruitful to view the categories of science and pseudoscience in terms of Wittgenstein's notion of *family resemblance* (Vermeir 2009): although there is no unique and essential feature shared by all pseudosciences, we can find a web of overlapping similarities that distinguish individual pseudosciences from bona fide science. The distinction between science and pseudoscience may be "vague" in a technical sense: while there are borderlands in between both domains, populated by specimens of theories that are interesting in their own right, we can readily point to paradigm cases of both *bona fide* science and pseudoscience. For example, in Chapter 5, we acknowledge that scientific theories are tested in bundles, and every scientific research programme has a protective set of auxiliary hypotheses around its core hypotheses. Resilience to falsification is a matter of gradients, but that is not say that we cannot come up with straightforward examples: if a face-saving auxiliary purchases no theoretical progress (in the senses we have discussed) to balance the resulting loss in theoretical simplicity, the move may be rightly dismissed as *ad hoc*. Much philosophical attention has been devoted to interesting borderline cases of adhocness in the history of science, and sometimes this has fostered the opinion that there is no consistent and objective way to explicate the notion of adhocness. In Chapter 6, however, we have drawn upon examples of pseudosciences to illuminate our inchoate notion of adhocness. By revealing a systematic pattern of evasive reasoning in typical pseudoscience, while at the same time acknowledging forms of theoretical protection in genuine science, we have vindicated and at the same time qualified Popper's insistence on the importance of empirical boldness in theory formation. When viewed in that light, Laudan's take on the demarcation project is like for a biologist to dismiss the concept of "mammal" because of such interesting in-betweens as the platypus.

Second, Laudan's argument that Ruse and Overton have "egregiously confuse[d] doctrines with the proponents of those doctrines" (Laudan 1996b, p. 353) shows that he is not acquainted with much pseudoscientific literature. Philip Quinn has echoed Laudan's complaint in his own comment on the Arkansas v. McLean case, and we have already encountered it in section 6.4.3:

> The requirement is that a theory *be* falsifiable by empirical evidence, not that its adherents admit that it has been falsified if and when it has been. (Quinn 1996, p. 381)

However, our discussion in Chapter 5 has shown that it is not always fruitful, nor indeed consistently possible, to follow Laudan's and Quinn's lead and frame the issue of demarcation purely in terms of logical relations between a set of propositions (the doctrine) and empirical observations. No sooner had we introduced the distinction between immunizing strategies and epistemic defense mechanisms, than we had to problematize and partly retract it. Indeed, one of the main purposes for introducing this distinction was to show that, in the most interesting cases, it cannot be rigorously maintained. To the extent that these pseudoscientific moves are already implicit in or provoked by the system of beliefs itself, we have decided to talk about epistemic defense mechanisms. By contrast, if the belief system does have identifiable borders and distinctive content, and if the face-saving move in question falls outside these theoretical bounds, then we are dealing with an immunizing strategy. At the end of Chapter 5 and in section 6.4.3, we have stressed that there is no general and foolproof way to identify the theory-as-such and evaluate its scientific credentials. Inevitably, we had to move our focus to the *attitude* of pseudoscientists themselves, and we arrived at the conclusion that, at least to a certain extent, pseudoscience is simply what pseudoscientists do. Again, by insisting on a demarcation purely on the basis of doctrines and by dismissing any consideration for the attitude of pseudoscientists as "ad hominem" (see also Quinn 1996, p. 381), it is no wonder that Laudan is ready to announce the death of the demarcation project.

The demarcation problem will not yield to an easy and clear-cut solution, as Laudan and others have pointed out, but that makes it all the more interesting and challenging from a philosophical point of view. We concur with Laudan that the problems of Intelligent Design Creationism or "scientific creationism" cannot be reduced to a violation of some basic rule of science that is written in stone, but this is a far cry from arguing that theories like creationism do not deserve to be labeled as *pseudoscience*. How would Laudan call an endeavor which miserably falls short of evidence *for* it while being abundant in evidence *against* it, whose empirical and conceptual problems have been repeatedly exposed by its critics, and which, in spite of all this, continues to be touted as scientific or worthy of credence by its adherents, on the basis of spurious evidence and unsound arguments? Don't we want some term to distinguish this from *bona fide* scientific knowledge, or from respectable but discarded hypotheses in science? Has Laudan another word in store for us? It is telling that, in the course of his diatribe against the demarcation project, Laudan himself uses language that implicitly commits him to a quite definite and prescriptive view as to what is and is not science. For example, he talks about a "false stereotype of what science is and how it works" (1996b, p. 355) and elsewhere he states that "this requirement [about natural laws] is an

altogether inappropriate standard for ascertaining whether a claim is scientific" (1996b, p. 353).

While reading the different chapters on pseudoscience in this work, some readers may have felt that Laudan's critique looms large over this whole project. Nevertheless, we have decided to postpone an extensive discussion of Laudan's famous essay till the very end of this dissertation, because we think that the proof of the pudding is in the eating. Now that we have covered quite some ground in the hinterland of science, we are in a better position to counter some of Laudan's objections. Hopefully, the present work has shown that the study of pseudosciences is worth pursuing and enriching for philosophers. Is it fair then to say about demarcationism, with Mark Twain, that the rumors of its death have been greatly exaggerated?

List of abbreviations

MN	Methodological Naturalism
IMN	Intrinsic Methodological Naturalism
PMN	Provisory or Pragmatic Methodological Naturalism
PN	Philosophical Naturalism
IC	Irreducible Complexity
NFL	No Free Lunch (Theorem)
NS	Natural Selection
IDC	Intelligent Design Creationism
SC	Social Constructivism
CSR	Cognitive Science of Religion

Bibliography

Kitzmiller vs Dover Area School District. Transcript of proceedings. Afternoon session (U.S. District Court for the Middle District of Penssylvania 2005).

Aaronovitch, D. (2010). *Voodoo histories : the role of the conspiracy theory in shaping modern history*. New York: Riverhead Books.

Alberts, B. (1998). The Cell as a Collection of Protein Machines: Preparing the Next Generation of Molecular Biologists. *Cell, 92*(3), 291-294.

Allen, D. (1989). *Christian belief in a postmodern world: the full wealth of conviction.* Louisville: Westminster John Knox Press.

Amanzio, M., G. Geminiani, D. Leotta and S. Cappa (2008). Metaphor comprehension in Alzheimer's disease: Novelty matters. *Brain and Language, 107*(1), 1-10.

Anderson, C. A., M. R. Lepper and L. Ross (1980). Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology, 39*(6), 1037-1049.

ap Siôn, T. (2009). Ordinary prayer and the empirical study of prayer cards. *Rural Theology, 7*(1), 17-31.

Aristotle (350BE/1991). *De anima*. Buffalo: Prometheus Books.

Aronson, E. (1992). The return of the repressed: Dissonance theory makes a comeback. *Psychological Inquiry, 3*(4), 303-311.

Ashworth, W. B. (2003). Christianity and the Mechanistic Universe'. In D. C. Lindberg and R. L. Numbers (Eds.), *When Science and Christianity Meet.* (pp. 61-84). Chicago: University of Chicago Press.

Atran, S. (2002). *In gods we trust : The evolutionary landscape of religion*. Oxford: Oxford University Press.

Avise, J. C. (2010a). Footprints of nonsentient design inside the human genome. *Proceedings of the National Academy of Sciences, 107*, 8969-8976. doi:10.1073/pnas.0914609107

------ (2010b). *Inside the Human Genome: A Case for Non-Intelligent Design*. Oxford: Oxford University Press.

Balch, R. W., J. Domitrovich, B. L. Mahnke and V. Morrison (1997). Fifteen years of failed prophecy: Coping with cognitive dissonance in a Baha'i sect. In T. Robbins and S. J. Palmer (Eds.), *Millennium, messiahs, and mayhem* (pp. 73-90). New York: Routledge.

Bamford, G. (1993). Popper's Explications of Ad Hocness: Circularity, Empirical Content, and Scientific Practice. *The British Journal for the Philosophy of Science, 44*(2), 335-355.

------ (1999). What is the Problem of Ad Hoc Hypotheses? *Science & Education, 8*(4), 375-386.

Barnes, P. M., E. Powell-Griner, K. McFann and R. L. Nahin (2004). Complementary and alternative medicine use among adults: United States, 2002. *Seminars in Integrative Medicine, 2*(2), 54-71.

Barrett, J. L. (1999). Theological Correctness: Cognitive Constraint and the Study of Religion. *Method & Theory in the Study of Religion, 11*, 325-339.

------ (2000). Exploring the natural foundations of religion. *Trends in Cognitive Sciences, 4*(1), 29-34.

------ (2001). How ordinary cognition informs petitionary prayer. *Journal of Cognition and Culture, 1*(3), 259-269.

------ (2004). *Why would anyone believe in God?* Walnut Creek (Calif.): AltaMira Press.

------ (2007). Cognitive science of religion: What is it and why is it? *Religious Compass, 1*(6), 768-786.

Barrett, J. L. and F. C. Keil (1996). Conceptualizing a nonnatural entity: Anthropomorphism in God concepts. *Cognitive Psychology, 31*(3), 219-247.

Barrett, J. L. and J. A. Lanman (2008). The science of religious beliefs. *Religion, 38*(2), 109-124.

Batson, C. D. (1975). Rational processing or rationalization?: The effect of discontinuing information on stated religious belief. *Journal of Personality and Social Psychology, 32*(1), 176-184.

Behe, M. J. (1996, October 29). Darwin Under the Microscope. *The New York Times,* p. 25

------ (2001). Reply to my critics: A response to reviews of Darwin's Black Box: the biochemical challenge to evolution. *Biology & Philosophy, 16*(5), 683-707.

------ (2004). Irreducible complexity: obstacle to Darwinian evolution. In W. Dembski and M. Ruse (Eds.), *Debating design: from Darwin to DNA* (pp. 352-370). New York: Cambridge University Press.

------ (2006). *Darwin's black box : the biochemical challenge to evolution (10th Anniversary Edition)*. New York, NY: Simon and Schuster.

------ (2007). *The edge of evolution : the search for the limits of Darwinism*. New York (N.Y.): Free press.

Beloff, J. (1994). Lessons of history. *Journal of the American Society for Psychical Research, 88*(7), 7-22.

Benassi, V. A., B. Singer and C. B. Reynolds (1980). Occult belief: Seeing is believing. *Journal for the Scientific Study of Religion, 19*(4), 337-349.

Benson, H., J. A. Dusek, J. B. Sherwood, P. Lam, C. F. Bethea, W. Carpenter, S. Levitsky, P. C. Hill, D. W. Clem and M. K. Jain (2006). Study of the Therapeutic Effects of Intercessory Prayer (STEP) in cardiac bypass patients: a multicenter randomized trial of uncertainty and certainty of receiving intercessory prayer. *American Heart Journal, 151*(4), 934-942.

Benvenuto, B. and R. Kennedy (1986). *The works of Jacques Lacan : an introduction*. London: Free Association Books.

Bering, J. M. and D. D. P. Johnson (2005). "O Lord You Perceive my Thoughts from Afar": Recursiveness and the Evolution of Supernatural Agency. *Journal of Cognition and Culture, 5*, 118-142.

Blackmore, S. J. (2000). *The meme machine*. Oxford: Oxford University Press.

Blancke, S., M. Boudry and J. Braeckman (2010). Simulation of biological evolution under attack, but not really: a response to Meester. *Biology & Philosophy, 26*(1), 113-118.

Blancke, S., M. Boudry, J. Braeckman, J. De Smedt and H. De Cruz (in press). Dealing with creationist challenges. What European biology teachers might expect in the classroom. *Journal of Biological Education*.

Bledsoe, B. (2006). *Every Reason to Be a Christian*. Longwood, Fl: Xulon Press.

Bloor, D. (1991). *Knowledge and Social Imagery*. Chicago: University Of Chicago Press.

------ (2007). Epistemic grace: Antirelativism as theology in disguise. *Common Knowledge, 13*(2-3), 250.

Boghossian, P. A. (2006). *Fear of Knowledge: Against Relativism and Constructivism*. Oxford: Oxford University Press.

Bonewitz, I. (1989). Real Magic. York Beach (Maine): Samuel Weiser.

Borch-Jacobsen, M. (1991). *Lacan: the absolute master*. Stanford, CA: Stanford University Press.

------ (2005). Une théorie zéro. In C. Meyer (Ed.), *Le livre noir de la psychanalyse* (pp. 178-183). Paris: Les Arènes.

Borch-Jacobsen, M. and S. Shamdasani (2006). *Le dossier Freud: enquête sur l'histoire de la psychanalyse*. Paris: Empêcheurs de penser en rond.

------ (2008). Interprefactions: Freud's legendary science. *History of the human sciences, 21*(3), 1-25.

Boudry, M. (2009a). In de ban van Freud & Co - het epistemische drijfzand van de psychoanalyse. *Tijdschrift Klinische Psychologie, 39*(1), 28-35.

------ (2009b). *Methodological Naturalism as an Intrinsic Property of Science: Grist to the Mill of Intelligent Design Theory*. Paper presented at the 150 Years After Origin: Biological, Historical and Philosophical Perspectives.

------ (2010). Exploring the hinterland of science. Book review of Massimo Pigliucci: Nonsense on stilts: How to tell science from bunk. *Metascience, 20*(1), 173-176.

------ (in press). The hypothesis that saves the day. Ad hoc reasoning in pseudoscience. *Logique et Analyse*.

Boudry, M., S. Blancke and J. Braeckman Grist to the mill of anti-evolutionism: the failed strategy of ruling the supernatural out of science a priori (submitted).

------ (2010a). How not to attack Intelligent Design Creationism: Philosophical misconceptions about Methodological Naturalism. *Foundations of Science, 15*(3), 227–244.

------ (2010b). Irreducible incoherence and intelligent design. A look into the conceptual toolbox of a pseudoscience. *Quarterly Review of Biology, 85*(4), 473-482.

Boudry, M. and J. Braeckman (2010). Immunizing strategies & epistemic defense mechanisms. *Philosophia*. doi:10.1007/s11406-010-9254-9

------ (forthcoming). How Convenient! The Epistemic Rationale of Self-validating Belief Systems. *Philosophical Psychology*.

Boudry, M. and F. Buekens (2011). The Epistemic Predicament of a Pseudoscience: Social Constructivism Confronts Freudian Psychoanalysis. *Theoria, 77*. doi:10.1111/j.1755-2567.2011.01098.x

Bowler, P. J. (2007). *Monkey trials and gorilla sermons : evolution and Christianity from Darwin to intelligent design*. Cambridge (Mass.): Harvard university press.

Boyer, P. (1994). *The naturalness of religious ideas : A cognitive theory of religion*. Berkeley (Calif.): University of California Press.

------ (1998). Cognitive tracks of cultural inheritance: How evolved intuitive ontology governs cultural transmission. *American Anthropologist, 100*(4), 876-889.

------ (2001). *Religion explained: The evolutionary origins of religious thought*. New York (N.Y.): Basic Books.

Boyer, P. and B. Bergstrom (2008). Evolutionary perspectives on religion. *Annual review of anthropology, 37*, 111-130.

Brown, T. L. (2003). Making Truth: The Roles of Metaphor in Science. *Urbana: University of Illinois Press*.

Buekens, F. (2005). Pourquoi Lacan est-il si obscur? In C. Meyer (Ed.), *Le livre noir de la psychanalyse* (pp. 269-277). Paris: Les Arènes.

------ (2006a). *Freuds Vergissing. De illusies van de psychoanalyse.* Leuven: Van Halewyck.

------ (2006b). *Jacques Lacan: proefvlucht in het luchtledige*. Leuven: Acco.

Buekens, F. and M. Boudry (2010). Unintended Institutional Facts. On the Structure of Psychoanalytic Hermeneutics. In K. François, B. Löwe, T. Müller and B. van Kerkhove (Eds.), *Foundations of the Formal Sciences VII*. London: King's College Publications.

------ (in press). Freudian Hermeneutics and Unintended Institutional Facts. *Philosophy of the Social Sciences*.

Cairns-Smith, A. G. (1986). *Seven clues to the origin of life: a scientific detective story*. Cambridge, UK: Cambridge University Press.

Carretta, T. R. and R. L. Moreland (1982). Nixon and Watergate: A field demonstration of belief perseverance. *Personality and Social Psychology Bulletin, 8*(3), 446-453.

Carroll, L. (1993). *Alice's adventures in Wonderland & Through the looking glass*. Hertfordshire: Wordsworth.

Carroll, R. T. (2003). *The skeptic's dictionary: a collection of strange beliefs, amusing deceptions, and dangerous delusions*. Hoboken, NJ: John Wiley & Sons.

Caudill, D. S. (1997). *Lacan and the subject of law : toward a psychoanalytic critical legal theory*. Atlantic Highlands, N.J.: Humanities Press.

Cervantes, M. (2008). *The history of Don Quixote, Volume 1*. Charleston, SC: Bibliobazaar.

Ciliberti, S., O. C. Martin and A. Wagner (2007). Innovation and robustness in complex regulatory gene networks. *Proceedings of the National Academy of Sciences, 104*(34), 13591-13596.

Cioffi, F. (1985). Psychoanalysis, Pseudoscience and Testability. In G. Currie and A. Musgrave (Eds.), *Popper and the human sciences*. Dordrecht: Martins Nijhoff Publishers.

------ (1998). *Freud and the question of pseudoscience*. Chicago: Open Court.

Clarke, S. (2002). Conspiracy theories and conspiracy theorizing. *Philosophy of the Social Sciences, 32*(2), 131-150.

Cleland, C. E. (2002). Methodological and epistemic differences between historical science and experimental science. *Philosophy of Science, 69*(3), 474-496.

Collins, H. M. (1981a). Son of Seven Sexes: The Social Destruction of a Physical Phenomenon. *Social Studies of Science, 11*(1), 33-62.

------ (1981b). Stages in the empirical program of relativism – Introduction. *Social Studies of Science, 11*(1), 3-10.

Condit, C. M., B. R. Bates, R. Galloway, S. B. Givens, C. K. Haynie, J. W. Jordan, G. Stables and H. M. West (2002). Recipes or blueprints for our genes? How contexts selectively activate the multiple meanings of metaphors. *Quarterly Journal of Speech, 88*(3), 303-325.

Coyne, J. A. (2009a, February 04, 2009). Seeing and believing. The never-ending attempt to reconcile science and religion, and why it is doomed to fail. *New Republic*

------ (2009b). *Why evolution is true*. New York (N.Y.): Viking.

Crews, F. C. (1986). *Skeptical engagements*. Oxford: Oxford University Press.

------ (2006). *Follies of the Wise: Dissenting Essays*. Emeryville, CA: Shoemaker & Hoard.

Darwin, C. (1964). *On the origin of species* (Facsimile of the first edition ed.). Cambridge, Mass.: Harvard University Press.

------ (1965 (1872)). *The expression of the emotions in man and animals*. Chicago (Ill.): University of Chicago press.

------ (1998 (1859)). *The origin of species*. Oxford: Oxford university press.

------ (2000). *The Life and Letters of Charles Darwin (Volume 2)*. Boston, MA: Adamant Media Corporation.

------ (2006). *The origin of species: a variorum text*. Philadelphia: University of Pennsylvania Press.

Davis, P. W., D. H. Kenyon and C. B. Thaxton (1993). *Of pandas and people: the central question of biological origins*. Dallas, TX: Haughton Pub Co.

Dawes, G. W. (2007). What is wrong with intelligent design? *International Journal for Philosophy of Religion, 61*(2), 69-81.

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

------ (1986). *The blind watchmaker*. Harlow: Longman Scientific & Technical.

------ (1989, April 9, 1989). Review of blueprints: Solving the mystery of evolution. By Maitland A. Edey and Donald C. Johanson. *New York Times,* p. 34 from http://www.philvaz.com/apologetics/p88.htm.

------ (1991). *The blind watchmaker* (Repr. ed.). Harmondsworth: Penguin books.

------ (1993). Viruses of the mind. In B. Dahlbom (Ed.), *Dennett and his critics: Demystifying mind* (pp. 13-27). Oxford: Blackwell.

------ (1997). Obscurantism to the Rescue. *Quarterly Review of Biology, 72*, 397-399.

------ (2006). *The God delusion*. London: Bantam.

------ (2009). *The greatest show on Earth : the evidence for evolution*. London: Bantam.

Dawkins, R. and Y. Wong (2005). *The ancestor's tale: a pilgrimage to the dawn of life*. London: Phoenix.

Dawson, L. L. (1999). When prophecy fails and faith persists: A theoretical overview. *Nova Religio, 3*(1), 60-82.

De Cruz, H. and J. De Smedt (2010). Science as structured imagination. *The Journal of Creative Behavior, 44*(1), 29-44.

Dein, S. (2001). What really happens when prophecy fails: The case of Lubavitch. *Sociology of Religion, 62*(3), 383-401.

Dembski, W. A. (1999). *Intelligent design : the bridge between science and technology*. Downers Grove, IL: InterVarsity press.

------ (2002). *No Free Lunch: Why specified complexity cannot be purchased without intelligence*. Lanham, Ld.: Rowman & Littlefield Publishers.

------ (2004). *The design revolution: Answering the toughest questions about intelligent design*. Downers Grove, IL: InterVarsity Press.

Dennett, D. C. (1991). *Consciousness explained*. Boston: Little, Brown and Company.

------ (1996). *Darwin's dangerous idea : Evolution and the meanings of life*. New York: Simon and Schuster.

------ (2003). In Darwin's wake, where am I? In J. Hodge and G. Radick (Eds.), *The Cambridge companion to Darwin* (pp. 393-412). Cambridge: Cambridge university press.

------ (2006). *Breaking the spell: Religion as a natural phenomenon*. New York, N.Y.: Viking (Penguin).

------ (2007). Atheism and evolution. In M. Martin (Ed.), *The Cambridge companion to atheism* (pp. 135-148). Cambridge: Cambridge University Press.

Dennett, D. C. and L. LaScola (2010). Preachers who are not believers. *Evolutionary Psychology, 8*(1), 122-150.

Derksen, A. A. (1993). The seven sins of pseudo-science. *Journal for General Philosophy of Science, 24*(1), 17-42.

Descartes, R. (1972 [1648]). *Treatise of man*. Cambridge: Harvard University Press.

deVries, P. (1986). Naturalism in the Natural Sciences: A Christian Perspective. *Christian Scholar's Review, 15*(4), 388-396.

Dietz, R. S. (1983). Gish's Law. *Geotimes, 28*(8), 11-12.

Dilley, S. C. (2010). Philosophical naturalism and methodological naturalism: Strange bedfellows. *Philosophia Christi, 12*(1), 118-141.

Ditto, P. H. and D. F. Lopez (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology, 63*, 568-568.

Duhem, P. M. M. (1954). *The aim and structure of physical theory*. Princeton (N.J.): Princeton university press.

Dunkelberg, P. (2003). Irreducible complexity demystified. *Talk Reason*, from www.talkdesign.org/faqs/icdmyst/ICDmyst.html

Eagle, M. (1993). The Dynamics of Theory Change in Psychoanalysis. In J. Earman, A. Janis, N. Rescher and G. Massey (Eds.), *Philosophical Problems of the Internal and External Worlds: Essays on the Philosophy of Adolf Grünbaum* (pp. 373-408). Pittsburgh: University of Pittsburgh Press.

Eagle, M. N. (1988). Skeptical Engagements - Crews, F. (book review). *Contemporary Psychology, 33*(5), 404-405.

------ (2003). The Postmodern Turn in Psychoanalysis: A Critique. *Psychoanalytic Psychology, 20*(3), 411-424.

Earman, J. (2000). *Hume's abject failure : the argument against miracles*. Oxford: Oxford university press.

Edgerton, R. B. (1992). *Sick societies: Challenging the myth of primitive harmony*. New York & Toronto: Free Press.

Edis, T. (1998). Taking creationism seriously. *Skeptic, 6*(2), 56-65.

------ (2002). *The ghost in the universe : God in light of modern science*. Amherst (N.Y.): Prometheus books.

------ (2006). Grand Themes, Narrow Constituency. In M. Young and T. Edis (Eds.), *Why intelligent design fails : a scientific critique of the new creationism* (pp. 9-19). New Brunswick, N.J.: Rutgers University Press.

Erwin, E. (1996). *A Final Accounting: Philosophical and Empirical Issues in Freudian Psychology*. Cambridge, MA: MIT Press.

------ (2002). *The Freud Encyclopedia: Theory, Therapy, and Culture*. New York: Routledge.

Esterson, A. (1993). *Seductive Mirage: An Exploration of the Work of Sigmund Freud*. Chicago: Open Court.

Estling, R. (2005). Stupid things. *Skeptical Inquirer, 29*(1), 56-57.

Evans-Pritchard, E. E. (1965). *Witchcraft, oracles and magic among the Azande*. Oxford: Clarendon press.

Fales, E. (2009). Animadversions on Kitzmiller v. Dover: Correct Ruling, Flawed Reasoning. *Infidels*. Retrieved from http://www.infidels.org/library/modern/evan_fales/intelligent-design.html

------ (2010). *Divine Intervention: Metaphysical and Epistemological Puzzles*: Taylor & Francis.

Farrell, J. (1996). *Freud's paranoid quest: Psychoanalysis and modern suspicion*. London: New York University Press.

Festinger, L., S. Schachter and H. W. Riecken (1964). *When prophecy fails : a social and psychological study of a modern group that predicted the destruction of the world*. New York (N.Y.): Harper and Row.

Fink, B. (1995). *The Lacanian subject : between language and jouissance*. Princeton (N.J.): Princeton university press.

------ (1997). *A clinical introduction to Lacanian psychoanalysis : theory and technique*. Cambridge (Mass.): Harvard university press.

Fishman, Y. I. (2009). Can Science Test Supernatural Worldviews? *Science & Education, 18*(6-7), 813-837.

Forrest, B. (2000). Methodological Naturalism And Philosophical Naturalism: Clarifying the Connection. *Philo, 3*(2), 7–29.

Forrest, B. C. and P. R. Gross (2007). Biochemistry by design. *Trends in Biochemical Sciences, 32*(7), 301-310.

------ (2007). *Creationism's Trojan horse : the wedge of intelligent design (Updated)*. Oxford: Oxford university press.

Forster, M. and E. Sober (1994). How to Tell When Simpler, More Unified, or Less Ad-Hoc Theories Will Provide More Accurate Predictions. *British Journal for the Philosophy of Science, 45*(1), 1-35.

Foster, J. G. (1971). *Enquiry into the Practice and Effects of Scientology*. London: Her Majesty's Stationery Office.

Foster, K. R. and H. Kokko (2009). The evolution of superstitious and superstition-like behaviour. *Proceedings of the Royal Society of London, Series B, 276*, 31-37.

Fox Keller, E. (1995). *Refiguring life: Metaphors of twentieth-century biology*. New York: Columbia University Press.

France, A. (1894). *Le Jardin d'Épicure*: Pitbook.com. Retrieved from http://www.pitbook.com/textes/pdf/jardin_epicure.pdf

Francis, L. J. and T. E. Evans (1995). The psychology of christian prayer: a review ofempirical research. *Religion, 25*(4), 371-388.

Freud, S. (1924). *Collected papers. Vol. 2*. London: The Hogarth Press Ltd and The Institute of Psycho-Analysis.

------ (1953a). *Standard edition. Vol. 4, (1900) The interpretation of dreams (first part)*. London: Hogarth Press.

------ (1953b). *Standard edition. Vol. 5, (1900-1901) The interpretation of dreams (second part) and, On dreams*. London: Hogarth Press.

------ (1953c). *Standard edition. Vol. 7, (1901-1905) : A case of hysteria ; Three essays on sexuality; and other works*. London: Hogarth Press.

------ (1955a). *Standard edition. Vol. 10, Two case histories: 'Little Hans' and the 'Rat man'*. London: Hogarth Press.

------ (1955b). *Standard edition. Vol. 17, (1917-1919) : An infantile neurosis, and other works*. London: Hogarth Press.

------ (1957). *Standard edition. Vol. 11, (1910) : Five lectures on psycho-analysis, Leonardo da Vinci, and other works*. London: Hogarth Press.

------ (1961). *Standard edition. Vol. 19, (1923-1925) : The ego and the id, and other works*. London: Hogarth Press.

------ (1964). *Standard edition. Vol. 22, (1932-1936) : New introductory lectures on psycho-analysis and other works*. London: Hogarth Press.

Fuller, S. (2007). *Science vs religion? Intelligent design and the problem of evolution*. Cambridge: Polity press.

Fusfield, W. D. (1993). Some Pseudoscientific Features of Transcendental-Pragmatic Grounding Projects. In H. Albert and K. Salamun (Eds.). Amsterdam-Atlanta: Mensch und Gesellschaft aus der Sicht des kritischen Rationalismus.

Galton, S. F. (1872). Statistical inquiries into the efficacy of prayer. *Fortnightly Review, 12*, 125-135.

Gardner, S. (2000). Psychoanalysis and the personal/sub-personal distinction. *Philosophical Explorations: An International Journal for the Philosophy of Mind and Action, 3*(1), 96 - 119.

Gellner, E. (1985). *The Psychoanalytic Movement: The Cunning of Unreason*. London: Paladin.

Gillett, E. (1998). Relativism and the Social-constructivist Paradigm. *Philosophy, Psychiatry, and Psychology, 5*(1), 37-48.

Gilovich, T. (1983). Biased evaluation and persistence in gambling. *Journal of Personality and Social Psychology, 44*(6), 1110-1126.

------ (1991). *How we know what isn't so : the fallibility of human reason in everyday life*. New York (N.Y.): Free press.

Gish, D. T. (1973). *Evolution: the fossils say no!* San Diego: Creation-Life Publishers.

Gliboff, S. (2000). Paley's design argument as an inference to the best explanation, or, Dawkins' dilemma. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 31*(4), 579-597.

Godfrey, L. (1979). Science and Evolution in the Public Eye. *Skeptical Inquirer, 4*(1), 21-32.

Golinski, J. (2005). *Making natural knowledge: Constructivism and the history of science*. Chicago: University Of Chicago Press.

Gordon, J. S. (1996). *Manifesto for a new medicine*. Reading, MA: Addison-Wesley.

Gosse, P. H. (1857). *Omphalos: an attempt to untie the geological knot*. London: J. Van Voorst.

Gould, S. J. (1980). *The panda's thumb : more reflections in natural history* (1st ed.). New York: Norton.

------ (1997). Foreward to "Why people believe weird things : pseudoscience, superstition, and other confusions of our time". In M. Shermer (Ed.). New York: Freeman.

------ (1999). Non-overlapping magisteria. *Skeptical Inquirer, 23*, 55-61.

Gregory, R. L. (1997). Knowledge in perception and illusion. *Philosophical Transactions of the Royal Society B: Biological Sciences, 352*(1358), 1121.

Gross, B. R. (1983). Philosophers at the Bar - Some Reasons for Restraint. *Science, Technology, and Human Values, 8*(4), 30-38.

Grünbaum, A. (1976). Ad hoc auxiliary hypotheses and falsificationism. *The British Journal for the Philosophy of Science, 27*(4), 329-362.

------ (1979). Is Freudian psychoanalytic theory pseudo-scientific by Karl Popper's criterion of demarcation? *American Philosophical Quarterly, 16*(2), 131-141.

------ (1984). *The foundations of psychoanalysis: a philosophical critique*. Berkeley: University of California Press.

------ (2008). Popper's Fundamental Misdiagnosis of the Scientific Defects of Freudian Psychoanalysis and of their Bearing on the Theory of Demarcation. *Psychoanalytic Psychology, 25*(4), 574-589.

Guthrie, S. (1993). *Faces in the clouds : a new theory of religion*. New York (N.Y.): Oxford university press.

Guttentag, O. E. (1940). Trends toward homeopathy: Present and past. *Bulletin of the History of Medicine*, 1172–93.

Habermas, J. and J. J. Shapiro (1981). *Knowledge and human interests*. London: Heinemann.

Hacking, I. (1999). *The social construction of what?* Harvard: Harvard University Press.

Häggström, O. (2007). Intelligent design and the NFL theorems. *Biology and Philosophy, 22*(2), 217-230.

Haldane, J. B. S. (1934). *Fact and faith*. London: Watts.

Harker, D. (2008). On the predilections for predictions. *The British Journal for the Philosophy of Science, 59*(3), 429.

Hartmann, M., P. C. Haddow and P. K. Lehre (2007). The genotypic complexity of evolved fault-tolerant and noise-robust circuits. *BioSystems, 87*(2-3), 224-232.

Haslanger, S. (2003). Social Construction: The 'Debunking' Project. In F. Schmitt (Ed.), *Socializing Metaphysics* (pp. 301-325). Lanham, MD: Rowman & Littlefield.

Hassoun, R., P. Schwartz, K. Feistel, M. Blum and C. Viebahn (2009). Axial differentiation and early gastrulation stages of the pig embryo. *Differentiation, 78*(5), 301-311.

Haught, J. F. (2000). *God after Darwin : a theology of evolution*. Boulder (Colo.): Westview press.

------ (2003). Is nature enough? No. *Zygon: Journal of Religion and Science, 38*(4), 769-782.

------ (2004). Darwin, design, and divine providence. In M. Ruse and W. A. Dembski (Eds.), *Debating Design: From Darwin to DNA* (pp. 229-245). Cambridge: Cambridge University Press.

Heller, J. (1994 [1962]). *Catch 22*. London: Vintage Books.

Hendrikse, J. L., T. E. Parsons and B. Hallgrimsson (2007). Evolvability as the proper focus of evolutionary developmental biology. *Evolution & Development, 9*(4), 393-401.

Hines, T. (2003). *Pseudoscience and the paranormal (2nd ed.)*. Amherst, NY: Prometheus Books

Hitchcock, C. and E. Sober (2004). Prediction versus accommodation and the risk of overfitting. *The British Journal for the Philosophy of Science, 55*(1), 1.

Hook, D. (2008). Absolute other: Lacan's 'big Other'as adjunct to critical social psychological analysis? *Social and Personality Psychology Compass, 2*(1), 51–73.

Howard-Snyder, D. and P. K. Moser (2002). *Divine hiddenness: new essays*. Cambridge Cambridge University Press.

Hume, D. (1998 [1779]). *Dialogues Concerning Natural Religion (2nd ed)*. Indianapolis: Hackett.

------ (2000 [1748]). *An enquiry concerning human understanding*. Oxford: Clarendon Press.

------ (2007 [1779]). *Dialogues concerning natural religion and other writings*. Cambridge: Cambridge University Press.

Humphrey, N. (1995). *Soul searching : Human nature and supernatural belief*. London: Chatto & Windus.

------ (1996). *Soul searching : Human nature and supernatural belief*. London: Vintage.

Iimura, T., N. Denans and O. Pourquié (2009). Establishment of Hox Vertebral Identities in the Embryonic Spine Precursors. *Current Topics in Developmental Biology, 88*, 201-234.

Irwin, H. J. (2009). *The psychology of paranormal belief*. Hatfield, Hertfordshire: University of Hertfordshire Press.

Jablonka, E. (2007). The developmental construction of heredity. *Developmental Psychobiology, 49*(8), 808-817.

Jacobs, D. M. (1998). *The threat: The secret agenda: What the aliens really want ... and how they plan to get it*. New York: Simon & Schuster.

Janssen, J., J. De Hart and C. Den Draak (1990). A content analysis of the praying practices of Dutch youth. *Journal for the Scientific Study of Religion, 29*(1), 99-107.

Jaroff, L. (1989, March 20). The gene hunt. *Time,* pp. 62-67.

Johannes, F., V. Colot and R. C. Jansen (2008). OPINION Epigenome dynamics: a quantitative genetics perspective. *Nature Reviews Genetics, 9*(11), 883-890.

Johnson, D. D. P. (2005). God's punishment and public goods. *Human Nature, 16*(4), 410-446.

Johnson, P. E. (1991). *Darwin on trial*. Washington, D.C. Lanham, MD: Regnery Gateway.

------ (1993). *Darwin on trial (2nd edition)*. Downers Grove, IL: InterVarsity Press.

------ (1995). *Reason in the balance : the case against naturalism in science, law and education*. Downers Grove (Ill.): InterVarsity press.

------ (1997). The Unraveling of Scientific Materialism. *First Things, 77*, 22-25.

------ (2001). Evolution as Dogma: The Establishment of Naturalism. In R. T. Pennock (Ed.), *Intelligent design creationism and its critics : philosophical, theological, and scientific perspectives* (pp. 59-76). Cambridge (Mass.): MIT press.

Kitzmiller v. Dover (400 F. Supp. 2nd 707 (M.D. Pa.) 2005).

Jopling, D. A. (2008). *Talking Cures and Placebo Effects*. Oxford: Oxford University Press.

Kant, I. (2007 [1790]). *Critique of judgment*. New York: Cosimo Books.

Kaplan, J. M. and M. Pigliucci (2001). Genes `for' Phenotypes: A Modern History View. *Biology and Philosophy, 16*(2), 189-213.

Keeley, B. L. (1999). Of conspiracy theories. *Journal of Philosophy, 96*(3), 109-126.

Kelemen, D. (2004). Are children "intuitive theists"? Reasoning about purpose and design in nature. *Psychological Science, 15*(5), 295-301.

Kelemen, D. and C. Di Yanni (2005). Intuitions about origins: purpose and intelligent design in children's reasoning about nature. *Journal of Cognition and Development, 6*(1), 3-31.

Kelemen, D. and E. Rosset (2009). The Human Function Compunction: Teleological explanation in adults. *Cognition, 111*(1), 138-143.

Kennedy, J. E. (2001). Why is psi so elusive? A review and proposed model. *Journal of Parapsychology, 65*(3), 219-246.

------ (2003). The capricious, actively evasive, unsustainable nature of psi: A summary and hypotheses. *The Journal of Parapsychology, 67*(1), 53-75.

Kitcher, P. (1981). Explanatory unification. *Philosophy of Science, 48*(4), 507-531.

------ (1982). *Abusing science: The case against creationism*. Cambridge (Mass.): MIT Press.

------ (1985). Darwin's achievement. In N. Rescher (Ed.), *Reason and Rationality in Natural Science* (pp. 127-189). Lanham: University Press of America.

------ (1993). *The advancement of science : science without legend, objectivity without illusions*. New York: Oxford University Press.

------ (1998). A Plea for Science Studies. In N. Koertge (Ed.), *A House Built on Sand: exposing postmodern myths about science* (pp. 32-56). Oxford/New York: Oxford University Press.

------ (2007). *Living with Darwin : evolution, design, and the future of faith*. Oxford ; New York: Oxford University Press.

Koertge, N. (1996). Wrestling with the Social Constructor. In P. Gross, N. Levitt and M. Lewis (Eds.), *The flight from science and reason* (pp. 266–273). Baltimore Johns Hopkins University Press.

------ (2000). 'New age' philosophies of science: constructivism, feminism and postmodernism. *The British Journal for the Philosophy of Science, 51*(4), 667-683.

Koperski, J. (2008). Two bad ways to attack Intelligent Design and two good ones. *Zygon, 43*(2), 433-449.

Kripke, S. (1980). *Naming and Necessity*. Oxford: Blackwell.

Kukla, A. (2000). *Social constructivism and the philosophy of science*. New York: Routledge.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*(3), 480-480.

Kurtz, P. (1986). *The transcendental temptation : a critique of religion and the paranormal*. Buffalo, N.Y.: Prometheus Books.

------ (1999). Should Skeptical Inquiry Be Applied to Religion? *Skeptical Inquirer, 23*, 24-28.

Kushner, H. S. (1981). *When bad things happen to good people*. New York: Avon Books.

Lacan, J. (1966). *Écrits,* (Vol. 912). Paris: Le Seuil.

------ (2004). *Le Séminaire de Jacques Lacan. 10: L 'angoisse*. Paris: Seuil.

------ (2005). *Le Séminaire de Jacques Lacan. 23: Le sinthome*. Paris: Seuil.

Lagerspetz, E. (2006). Institutional facts, performativity and false beliefs. *Cognitive Systems Research, 7*(2-3), 298-306.

Lakatos, I. (1968). Criticism and the methodology of scientific research programmes. *Proceedings of the Aristotelian Society, 69*, 149-186.

Lakatos, I. and A. Musgrave (1970). *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press.

Lakoff, G. and M. Johnson (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Laudan, L. (1977). *Progress and its problems: towards a theory of scientific growth*. London: Routledge & Kegan Paul.

------ (1982). Commentary: Science at the Bar-Causes for Concern. *Science, Technology, & Human Values, 7*(41), 16-19.

------ (1983). The demise of the demarcation problem. In R. S. Cohen and L. Laudan (Eds.), *Physics, Philosophy, and Psychoanalysis: Essays in Honor of Adolf Grünbaum.* (pp. 111–128). Dordrecht: D. Reidel.

------ (1996a). The demise of the demarcation problem. In M. Ruse (Ed.), *But is it science? The philosophical question in the creation/evolution controversy* (pp. 337-350). Amherst (N.Y.): Prometheus books.

------ (1996b). Science at the bar - Causes for concern. In M. Ruse (Ed.), *But is it science? The philosophical question in the creation/evolution controversy* (pp. 351-355). Amherst (N.Y.): Prometheus books.

Leavy, S. A. (1983). The Image and the Word. Further Reflections on Jacques Lacan. In J. H. Smith and W. Kerrigan (Eds.), *Interpreting Lacan* (pp. 3-20). New Haven: Yale University Press.

Leguil-Badal, C. (2006). Sur le cognitivisme. In J.-A. Miller (Ed.), *l'Anti-livre noir de la psychanalyse* (pp. 257-278). Paris: Seuil.

Lem, S. (1999). *A perfect vacuum*. Evanston, IL: Northwestern University Press.

Lenski, R. E., C. Ofria, R. T. Pennock and C. Adami (2003). The evolutionary origin of complex features. *Nature, 423*(6936), 139-144.

Leplin, J. (1975). The concept of an ad hoc hypothesis. *Studies in History and Philosophy of Science, 5*(4), 309-345.

Levine, J. S. and K. R. Miller (1994). *Biology: Discovering life*. Lexington: Heath Press.

Lewontin, R. (1983). Introduction. In L. Godfrey (Ed.), *Scientists Confront Creationism* (pp. xxiii-xxvi). New York: Norton.

------ (1997). Billions and billions of demons. Book review of 'The Demon-Haunted World: Science as a Candle in the Dark' by Carl Sagan. *The New York Review of Books, 44*(1), 31.

Lienard, P. and P. Boyer (2006). Whence collective rituals? A cultural selection model of ritualized behavior. *American Anthropologist, 108*(4), 814.

Lindeman, M. and K. Aarnio (2007). Superstitious, magical, and paranormal beliefs: An integrative model. *Journal of Research in Personality, 41*(4), 731-744.

Lord, C. G., L. Ross and M. R. Lepper (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37*(11), 2098-2109.

Lukes, S. (2007). The problem of apparently irrational beliefs. In S. P. Turner and M. W. Risjord (Eds.), *Philosophy of anthropology and sociology* (pp. 591-606). Amsterdam: Elsevier/North-Holland.

Lupfer, M. B., D. Tolliver and M. Jackson (1996). Explaining Life-Altering Occurrences: A Test of the 'God-of-the-Gaps' Hypothesis. *Journal for the Scientific Study of Religion, 35*(4), 379-391.

Macbeth, N. (1974). *Darwin retried: an appeal to reason*. London: Garnstone Press.

Mack, J. E. (1995). *Abduction : human encounters with aliens*. London: Simon and Schuster.

Mackay, C. (1974 [1841]). *Extraordinary popular delusions and the madness of crowds*. New York: Barnes & Noble Publishing.

Mackie, J. L. (1955). Evil and omnipotence. *Mind, 64*(254), 200-212.

Macmillan, M. (1997). *Freud Evaluated: The Completed Arc*. Cambridge, MA: MIT Press.

Maienschein, J. (2007). Untangling debates about science and religion. In N. C. Comfort (Ed.), *The panda's black box: opening up the intelligent design controversy* (pp. 83-108). Baltimore: Johns Hopkins University Press.

Marks, D. (2000). *The psychology of the psychic* (2nd ed.). Amherst (N.Y.): Prometheus books.

Marmor, J. (1962). Psychoanalytic Therapy as an Educational Process. In J. Messerman (Ed.), *Psychoanalytic Education Science Series* (Vol. 5, pp. 286-299). New York: Orune & Stratton.

Martin, E. (1994). *Flexible bodies: Tracking immunity in American culture from the days of polio to the age of AIDS*. Boston: Beacon Press.

Matthews, W. J., J. Conti and T. Christ (2000). God's HMO: Prayer, Faith, Belief and Physical Well-Being'. *Skeptic, 8*(2), 64-66.

Mayr, E. and W. B. Provine (1998). *The evolutionary synthesis: perspectives on the unification of biology*. Cambridge: Harvard Univ Press.

McCauley, R. N. and E. T. Lawson (2002). *Bringing ritual to mind: Psychological foundations of cultural forms*. Cambridge Cambridge University Press.

McGinn, C. (1994). The problem of philosophy. *Philosophical Studies, 76*(2), 133-156.

McMullin, E. (2001). Plantinga's defense of special creation. In R. T. Pennock (Ed.), *Intelligent design creationism and its critics : philosophical, theological, and scientific perspectives* (pp. 168-169). Cambridge (Mass.): MIT press.

Meester, R. (2009). Simulation of biological evolution and the NFL theorems. *Biology and Philosophy, 24*(4), 461-472.

Mele, A. R. (1994). Self-control and belief. *Philosophical Psychology, 7*(4), 419-435.

Melton, J. G. (1985). Spiritualization and reaffirmation: What really happens when prophecy fails. *American Studies, 26*(2), 17.

Mercier, H. (in press). Our pigheaded core: How we became smarter to be influenced by other people. In B. Calcott, R. Joyce and K. Sterelny (Eds.), *Signaling, Commitment and Emotion*. Cambridge (Mass.): MIT Press.

Miller, K. B. (2009). The misguided attack on methodological naturalism. In J. S. Schneiderman and W. D. Allmon (Eds.), *For the Rock Record : Geologists on Intelligent Design* (pp. 117-140). Berkeley and Los Angeles: University of California Press.

Miller, K. R. (2000). *Finding Darwin's God : a scientist's search for common ground between God and evolution*. New York (N.Y.): HarperCollins.

------ (2004). The flagellum unspun. The Collapse of "Irreducible Complexity". In W. Dembski and M. Ruse (Eds.), *Debating design: from Darwin to DNA* (pp. 81-98). Cambridge: Cambridge university press.

------ (2008). *Only a theory : evolution and the battle for America's soul*. New York (N.Y.): Viking.

Minelli, A. and G. Fusco (2005). Conserved versus innovative features in animal body organization. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution, 304B*(6), 520-525.

Moore, R. (1999). *The Creation of Reality in Psychoanalysis: A View of the Contributions of Donald Spence, Roy Schafer, Robert Stolorow, Irwin Z. Hoffman, and Beyond*. Hillsdale, NJ: Analytic Press.

Morris, H. (1974). *Scientific Creationism*. San Diego: Creation-Life Publishers.

Morris, H. M. (1963). *Twilight of evolution*. Grand Rapids, Mich: Baker Pub Group.

Muller, G. B. (2007). Evo-devo: extending the evolutionary synthesis. *Nat Rev Genet, 8*(12), 943-949.

Muller, H. J. (1918). Genetic Variability, Twin Hybrids and Constant Hybrids, in a Case of Balanced Lethal Factors. *Genetics, 3*, 422-499.

Muller, J. P. and W. J. Richardson (1982). *Lacan and language: A Reader's Guide to Ecrits* New York: International Universities Press.

Mullin, R. B. (2008). Science, miracles and the prayer-gauge debate. In D. C. Lindberg and R. L. Numbers (Eds.), *When Science and Christianity Meet* (pp. 203-224). Chicago and London: University of Chicago Press.

Murdock, G. P. (1980). *Theories of illness: A world survey*. Pittsburgh University of Pittsburgh Press.

Murray, M. J. (1993). Coercion and the Hiddenness of God. *American Philosophical Quarterly, 30*(1), 27-38.

Nagel, T. (2008). Public education and intelligent design. *Philosophy & Public Affairs, 36*(2), 187-205.

National Academy of Sciences (1998). *Teaching about evolution and the nature of science*. Washington, D.C.: National Academy Press.

Nelkin, D. (2001). Molecular metaphors: the gene in popular discourse. *Nat Rev Genet, 2*(7), 555-559.

Nelson, P. A. (1996). The role of theology in current evolutionary reasoning. *Biology & Philosophy, 11*(4), 493-517.

------ (1998). Jettison the Arguments, or the Rule? The Place of Darwinian Theological Themata in Evolutionary Reasoning. Retrieved 23/03/2009, from http://www.arn.org/docs/nelson/pn_jettison.htm

Nesse, R. M. (2001). The smoke detector principle. *Annals of the New York Academy of Sciences, 935*, 75-85.

Nickell, J. (1993). *Looking for a miracle: weeping icons, relics, stigmata, visions & healing cures*. New York: Prometheus Books.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*, 175-220.

Nieuwentyt, B. (1715). *Het regt gebruik der werelt beschouwingen, ter overtuiginge van ongodisten en ongelovigen aangetoont*. Amsterdam: Wolters and Pauli.

Nisbett, R. and L. Ross (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs (N.J.): Prentice-Hall.

Nobus, D. (1998). *Key concepts of Lacanian psychoanalysis*. London: Rebus press.

------ (2004a). Knowledge in failure: On the crises of legitimacy within Lacanian psychoanalysis. In A. Casement (Ed.), *Who owns psychoanalysis?* (pp. 203-225). London: Karnac.

------ (2004b). The punning of reason. *Angelaki, 9*(1), 189-201.

Numbers, R. (2003). Science without God: Natural laws and Christian beliefs. In D. C. Lindberg and R. L. Numbers (Eds.), *When Science and Christianity Meet* (pp. 265–85). Chicago / London: University of Chicago Press.

Oppy, G. (1996). Hume and the argument for biological design. *Biology & Philosophy, 11*(4), 519-534.

Orr, A. H. (1997). Darwin vs. Intelligent Design (again): The latest attack on evolution Is cleverly argued, biologically informed—and wrong. *Boston Review, 22*(6).

Orwell, G. (1945). Notes on nationalism. *Polemic*. Retrieved from http://orwell.ru/library/essays/nationalism/english/e_nat

Paley, W. (1802). *Natural theology*. London: Gould and Lincoln.

Pallen, M. J. and N. J. Matzke (2006). From The Origin of Species to the origin of bacterial flagella. *Nature Reviews Microbiology, 4*(10), 784-790.

Park, R. L. (2002). *Voodoo science: The road from foolishness to fraud*. New York: Oxford University Press.

------ (2008). *Superstition : belief in the age of science* (1st ed.). Princeton, N.J.: Princeton University Press.

Pennock, R. T. (1996). Naturalism, evidence and creationism: The case of Phillip Johnson. *Biology & Philosophy, 11*(4), 543-559.

------ (1999). *Tower of Babel: The evidence against the new creationism*. Cambridge (Mass.): MIT Press.

------ (2000). *Lions, Tigers and APES, Oh My! Creationism vs. Evolution in Kansas*. Paper presented at the AAAS Dialogue on Science, Ethics and Religion. from http://www.aaas.org/spp/dser/02_Events/Conferences/CF_2000_04_1415_Teach/Abstract_Pennock.shtml

------ (2007). God of the Gaps: The argument from ignorance and the limits of methodological naturalism. In A. J. Petto and L. R. Godfrey (Eds.), *Scientists Confront Intelligent Design and Creationism* (pp. 309-338). New York & London: WW Norton & Company.

------ (2009). Can't philosophers tell the difference between science and religion?: Demarcation revisited. *Synthese*, 1-30.

Pennock, R. T. and M. Ruse (2009). *But is it science? : the philosophical question in the creation/evolution controversy* (Updated ed.). Amherst (N.Y.): Prometheus books.

Perakh, M. (2002). A free lunch in a mousetrap. Retrieved January 7, 2010, from http://www.talkreason.org/articles/dem_nfl.cfm#salvage

------ (2004). There is a free lunch after all. William Dembski's wrong answers to irrelevant questions. In M. Young and T. Edis (Eds.), *Why intelligent design fails* (pp. 153–171). New Brunswick, N.J.: Rutgers University Press.

------ (2004). *Unintelligent design*. Amherst (N.Y.): Prometheus books.

------ (2008). Flagella Myths. How Intelligent Design Proponents Created the Myth That Bacteria Flagella Look Like Man-made Machines. *Skeptic, 14*(3).

Peters, S. F. (2008). *When prayer fails: Faith healing, children, and the law*. Oxford: Oxford University Press.

Pigliucci, M. (2001). *Phenotypic plasticity: beyond nature and nurture*: Johns Hopkins University Press.

------ (2002). *Denying evolution: Creationism, scientism, and the nature of science*. Sunderland, MA: Sinauer Associates.

------ (2009). An extended synthesis for evolutionary biology. *Annals of the New York Academy of Sciences, 1168*(1), 218-228.

------ (2010). *Nonsense on stilts: How to tell science from bunk*. Chicago: University Of Chicago Press.

Pigliucci, M. and M. Boudry (in press). Why Machine-Information Metaphors are Bad for Science and Science Education. *Science & Education*. doi:10.1007/s11191-010-9267-6

Plantinga, A. (1996). Methodological Naturalism? In J. van der Meer (Ed.), *Facets of Faith and Science, Volume 1: Historiography and Modes of Interaction* (pp. 177-221).

------ (2001a). Evolution, neutrality, and antecedent probability: a reply to McMullin and Van Till. In R. T. Pennock (Ed.), *Intelligent design creationism and its critics : philosophical, theological, and scientific perspectives* (pp. 197-236). Cambridge (Mass.): MIT press.

------ (2001b). Methodological Naturalism? In R. T. Pennock (Ed.), *Intelligent design creationism and its critics : philosophical, theological, and scientific perspectives* (pp. 339-361). Cambridge (Mass.): MIT press.

------ (2001c). When faith and reason clash: evolution and the Bible. In R. T. Pennock (Ed.), *Intelligent design creationism and its critics : philosophical, theological, and scientific perspectives* (pp. 113-145). Cambridge (Mass.): MIT press.

Popper, K. R. (2002). *Conjectures and refutations: The growth of scientific knowledge*. London: Routledge.

Pyszczynski, T. and J. Greenberg (1987). Toward an integration of cognitive and motivational perspectives on social inference: A biased hypothesis-testing model. In L. Berkowitz (Ed.), *Advances in experimental social psychology (Vol. 20)* (pp. 297–340). New York: Academic Press.

Quine, W. V. O. (1953). *From a logical point of view*. Cambridge, Mass.: Harvard University Press.

Quinn, P. L. (1996). The philosopher of science as expert witness. In M. Ruse (Ed.), *But is it science? The philosophical question in the creation/evolution controversy* (pp. 367-385). Amherst (N.Y.): Prometheus books.

Randi, J. (1981). Selective test selection. *Skeptical Inquirer, 5*, 12-13.

Rao, K. R. (1968). Spontaneous ESP in laboratory tests: the error phenomenon. *Journal of the American Society for Psychical Research, 62*, 63-72.

Ratzsch, D. (2005). How Not to Critique Intelligent Design Theory. *Ars Disputandi: The Online Journal for Philosophy of Religion, 5*.

Recanati, F. (2000). *Oratio Obliqua, Oratio Recta. An Essay on Meta-Representation*. Cambridge (Mass.): MIT Press.

Richerson, P. J. and R. Boyd (2005). *Not by genes alone: How culture transformed human evolution*. Chicago (Ill.): University of Chicago press.

Richter, R. (2002). What Science Can and Cannot Say: The Problems with Methodological Naturalism. *Reports of the National Center for Science Education, 22*(1-2), 18-22.

Ricœur, P. (1970). *Freud and philosophy : an essay on interpretation*. New Haven: Yale University Press.

Roggen, D., D. Federici and D. Floreano (2007). Evolutionary morphogenesis for multi-cellular systems. *Genetic Programming and Evolvable Machines, 8*(1), 61-96.

Rosenhouse, J. (2002). Probability, optimization theory and evolution. *Evolution, 56*(8), 1721-1722.

Ross, L., M. R. Lepper, F. Strack and J. Steinmetz (1977). Social explanation and social expectation: Effects of real and hypothetical explanations on subjective likelihood. *Journal of Personality, 35*(11), 817-829.

Roth, A. A. (1980). Implications of the various interpretations of the fossil record. *Origins, 7*(2), 71-86.

Ruse, M. (1982). *Darwinism defended : a guide to the evolution controversies*. Reading (Mass.): Addison-Wesley.

------ (1996a). *But is it science? The philosophical question in the creation/evolution controversy*. Amherst (N.Y.): Prometheus books.

------ (1996b). Pro judice. In M. Ruse (Ed.), *But is it science? The philosophical question in the creation/evolution controversy* (pp. 356-362). Amherst (N.Y.): Prometheus books.

------ (2001). *Can a Darwinian be a Christian? The relationship between science and religion*. Cambridge: Cambridge university press.

------ (2004). The Argument from Design: A Brief History. In W. A. Dembski and M. Ruse (Eds.), *Debating design : from Darwin to DNA* (pp. 13-31). Cambridge: Cambridge university press.

------ (2005). Methodological Naturalism Under Attack. *South African Journal of Philosophy, 24*(1), 44-61.

------ (2006). *Darwinism and its discontents*. Cambridge: Cambridge university press.

Rutka, J., P. Kongkham, P. Northcott, C. Carlotti, M. Guduk, H. Osawa, O. Moreno, H. Seol, A. Restrepo, A. Weeks, S. Nagai and C. Smith (2009). The evolution and application of techniques in

molecular biology to human brain tumors: a 25 year perspective. *Journal of Neuro-Oncology, 92*(3), 261-273.

Sagan, C. (1995). Wonder and skepticism. *Skeptical Enquirer, 19*(1), 24–30.

Sagan, E. (1979). Religion and magic: A developmental view. *Sociological Inquiry, 49*(2-3), 87-116.

Saminathan, R., A. Pachiappan, L. Feng, E. Rowan and P. Gopalakrishnakone (2009). Transcriptome Profiling of Neuronal Model Cell PC12 from Rat Pheochromocytoma. *Cellular and Molecular Neurobiology, 29*(4), 533-548.

Sarkar, S. (2007). *Doubting Darwin? Creationist designs on evolution*. Malden (Mass.): Blackwell.

------ (2009). The science question in intelligent design. *Synthese*. doi:10.1007/s11229-009-9540-x

Sarup, M. (1992). *Jacques Lacan*. New York (N.Y.): Harvester Wheatsheaf.

Sayers, J. (1987). Melanie Klein, psychoanalysis, and feminism. *Feminist Review, 25*, 23-37.

Schafer, R. (1992). *Retelling a Life: Narration and Dialogue in Psychoanalysis*. New York: Basic Books.

Schafersman, S. D. (1997). Naturalism is an Essential Part of Science and Critical Inquiry. Retrieved 30/03/2009, from http://www.freeinquiry.com/naturalism.html

Schellenberg, J. L. (2006). *Divine hiddenness and human reason*: Cornell University Press.

Schmied, G. (2007). God images in prayer intention books. *Implicit Religion, 5*(2), 121.

Schrödinger, E. (1992 [1944]). *What is life? The physical aspect of the living cell*: Cambridge University Press.

Scott, E. C. (1998). Two Kinds of Materialism: keeping them separate makes faith and science compatible. *Free Inquiry, 18*(Spring 1998), 20-20.

------ (1999). The 'Science and Religion' Movement. An Opportunity for Improved Public Understanding of Science? *Skeptical Inquirer, 23*(4), 29-31.

------ (2004). *Evolution vs. creationism : an introduction*. Berkeley (Calif.): University of California press.

------ (2009). *Evolution vs. creationism : an introduction* (2nd ed.). Westport (Conn.): Greenwood press.

Searle, J. R. (1969). *Speech acts: an essay in the philosophy of language*. London: Cambridge University Press.

------ (1995). *The Construction of Social Reality*. New York: Free Press.

------ (2010). *Making the social world : the structure of human civilization*. Oxford: Oxford University Press.

Shackel, N. (2005). The vacuity of postmodern methodology. *Metaphilosophy, 36*(3), 295-320.

Shallit, J. (2002). Book review of "No Free Lunch: Why Specified Complexity Cannot be Purchased Without Intelligence" by William Dembski. *BioSystems, 66*(1-2), 93-99.

Shanks, N. (2004). *God, the devil, and Darwin : a critique of intelligent design theory*. Oxford: Oxford university press.

Shanks, N. and K. Joplin (1999). Redundant Complexity: A Critical Analysis of Intelligent Design in Biochemistry. *Philosophy of Science, 66*(2), 268-282.

Sheldrake, R. (1995). *Seven experiments that could change the world: a do-it-yourself guide to revolutionary science*. New York, NY: Putnam Publishing Group.

Shermer, M. (2002). *Why people believe weird things : pseudoscience, superstition, and other confusions of our time* (Rev. and expanded. ed.). New York: A.W.H. Freeman/Owl Book.

------ (2003). Why smart people believe weird things. *Skeptic, 10*(2), 62-73.

Slezak, P. (1994). Sociology of scientific knowledge and scientific education: Part I. *Science & Education, 3*(3), 265-294.

Slone, D. J. (2004). *Theological incorrectness: why religious people believe what they shouldn't*. New York: Oxford University Press, USA.

Smith, K. C. (2001). Appealing to ignorance behind the cloak of ambiguity. In R. T. Pennock (Ed.), *Intelligent design creationism and its critics : philosophical, theological, and scientific perspectives* (pp. 705-735). Cambridge (Mass.): MIT press.

Snow, C. P. (1993 [1959]). *The two cultures*. Cambridge: Cambridge University Press.

Snow, D. A. and R. Machalek (1982). On the presumed fragility of unconventional beliefs. *Journal for the Scientific Study of Religion, 21*(1), 15-26.

Sober, E. (2000). *Philosophy of biology (2nd edition)*. Boulder, Colorado: Westview Press.

------ (2002). Intelligent design and probability reasoning (Developing an philosophical epistemology of irreducible complexity). *International Journal for Philosophy of Religion, 52*(2), 65-80.

------ (2004). The design argument. In W. Mann (Ed.), *The Blackwell guide to philosophy of religion* (pp. 117-147). Oxford: Blackwell.

------ (2007). What is wrong with intelligent design? *Quarterly Review of Biology, 82*(1), 3-8.

------ (2008). *Evidence and Evolution: the logic behind the science*. Cambridge: Cambridge University Press.

------ (2010). Evolution without naturalism. In J. Kvanvig (Ed.), *Oxford Studies in Philosophy of Religion* (Vol. 3). Oxford: Oxford University Press.

Sokal, A. and J. Bricmont (1997). *Impostures intellectuelles*. Paris: Jacob.

Spelke, E. (1994). Initial knowledge: six suggestions. *Cognition, 50*, 431-445.

Spelke, E. S. and K. D. Kinzler (2007). Core knowledge. *Developmental Science, 10*(1), 89-96.

Spelke, E. S., A. Phillips and A. L. Woodward (1995). Infants' knowledge of object motion and human action. In D. Sperber, D. Premack and A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 44–78).

Spence, D. P. (1982). *Narrative truth and historical truth : meaning and interpretation in psychoanalysis*. New York / London: Norton.

Sperber, D. (1985). Anthropology and psychology: Towards an epidemiology of representations. *Man, 20*(1), 73-89.

------ (1990). The epidemiology of beliefs. In G. Gaskell and C. Fraser (Eds.), *The social psychological study of widespread beliefs* (pp. 25-44). Oxford: Clarendon Press.

------ (1996). *Explaining culture: A naturalistic approach*. Cambridge, Mass.: Blackwell.

------ (1997). Intuitive and reflective beliefs. *Mind & Language, 12*(1), 67-83.

------ (2000). An objection to the memetic approach to culture. In R. Aunger (Ed.), *Darwinizing culture: The status of memetics as a science* (pp. 163-173). Oxford: Oxford University Press.

Sperber, D., F. Clément, C. Heintz, O. Mascaro, H. Mercier, G. Origgi and D. Wilson (2010). Epistemic vigilance. *Mind & Language, 25*(4), 359–393.

Sperber, D. and D. Wilson (1995). *Relevance : communication and cognition* (2nd ed.). Oxford: Blackwell.

Stanley, A. (2000, May 14). Vatican Discloses the 'Third Secret' of Fatima. *The New York Times*

Stanley, K. (2007). Compositional pattern producing networks: A novel abstraction of development. *Genetic Programming and Evolvable Machines, 8*(2), 131-162.

Stenger, V. J. (2008). *God : the failed hypothesis : how science shows that God does not exist*. Amherst (N.Y.): Prometheus books.

Stern, D. (1992). Commentary on constructivism in clinical psychoanalysis. *Psychoanalytic Dialogues, 2*, 331-363.

Strahler, A. N. (1992). *Understanding science: An Introduction to Concepts and Issues*. Buffalo: Prometheus Books.

Streatfeild, D. (2007). *Brainwash: The secret history of mind control*. New York: Thomas Dunne Books.

Stump, E. (1979). Petitionary prayer. *American Philosophical Quarterly, 16*(2), 81-91.

Sturm, R. A. and M. Larsson (2009). Genetics of human iris colour and patterns. *Pigment Cell & Melanoma Research, 22*(5), 544-562.

Swinburne, R. (2004). *The existence of God*. Oxford: Clarendon Press.

Talmont-Kaminski, K. (2007). Reason, Red in Tooth and Claw: Naturalising Enlightenment Thinking. In G. Gasser (Ed.), *How successful is naturalism?* (pp. 183-199). Frankfurt: Ontos Verlag.

------ (2008). In a mirror, darkly: Does superstition reflect rationality? *Skeptical Enquirer, 32*(4), 48-51.

------ (2009). The fixation of superstitious beliefs. *Teorema, 28*(3), 81-95.

Tavris, C. and E. Aronson (2008). *Mistakes were made (but not by me) : Why we justify foolish beliefs, bad decisions, and hurtful acts*. London: Pinter and Martin.

Thurston, L. (1998). Ineluctable Nodalities: On the Borromean Knot. In D. Nobus (Ed.), *Key Concepts of Lacanian Psychoanalysis* (pp. 139-163). London: Rebus press.

Tian, C., P. K. Gregersen and M. F. Seldin (2008). Accounting for ancestry: population substructure and genome-wide association studies. *Human Molecular Genetics, 17*(R2), R143-R150.

Timpanaro, S. (1976). *The Freudian slip : psychoanalysis and textual criticism*. London: Verso.

Tumminia, D. (1998). How prophecy never fails: Interpretive reason in a flying-saucer group. *Sociology of Religion, 59*(2), 157-170.

Tyson, N. D. (2005, November). The Perimeter of Ignorance. *Natural History, 114*.

Uttamchandani, M., C. H. S. Lu and S. Q. Yao (2009). Next Generation Chemical Proteomic Tools for Rapid Enzyme Profiling. *Accounts of Chemical Research, 42*(8), 1183-1192.

Van Rillaer, J. (1980). *Les illusions de la psychanalyse*. Bruxelles: Mardaga.

Van Till, H. J. (1996). Basil, Augustine, and the doctrine of creation's functional integrity. *Science and Christian Belief, 8*, 21-38.

------ (2001). When faith and reason cooperate. In R. T. Pennock (Ed.), *Intelligent design creationism and its critics : philosophical, theological, and scientific perspectives* (pp. 147-163). Cambridge (Mass.): MIT press.

Van Till, H. J. and P. E. Johnson (1993). God and Evolution: An Exchange. *First Things, 34*, 32-41.

Verhaeghe, P. (2004). *On Being Normal and Other Disorders. A Manual for Clinical Psychodiagnostics*. New York: Other Press.

Vermeir, K. (2009). A Family of Sciences. From a Wittgensteinian to a Procedural Analysis of the Demarcation Problem. *Tijdschrift voor Filosofie, 71*(1), 119-145.

von Hippel and R. Trivers (in press). The evolution and psychology of self-deception. *Behavioral and Brain Sciences*.

Vyse, S. A. (1997). *Believing in magic : the psychology of superstition*. New York (N.Y.): Oxford university press.

Watson, J. D. and F. H. C. Crick (1953). Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature, 171*(4356), 737-738.

Weisberg, J. (2005). Firing squads and fine-tuning: Sober on the design argument. *British Journal for the Philosophy of Science, 56*(4), 809-821.

Whewell, W. (1840). *The philosophy of the inductive sciences: founded upon their history*. London: J. W. Parker.

Williams, R. J. (1980). *Biochemical individuality*. Austin, TX: University of Texas Press.

Wiseman, R. (1997). *Deception & self-deception: Investigating psychics*. Amherst, N.Y.: Prometheus Books.

------ (2010). 'Heads I win, tails you lose'. How parapsychologists nullify null results. *Skeptical Inquirer, 34*(1), 36-39.

Wittgenstein, L. (1951 [2009]). *Philosophical investigations*. New York: Wiley.

------ (1972). *The blue and brown books. Preliminary studies for the 'philosophical investigations'*. Oxford: Basil Blackwell.

Woetzel, D. (2005). Can science consider supernatural explanations? *Creation Matters, 10*(4), 1-5.

Wolpert, D. H. (2002). *William Dembski's treatment of the no free lunch theorems is written in jello*. Retrieved from http://www.talkreason.org/articles/jello.cfm

Wolpert, D. H. and W. G. Macready (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation, 1*(1), 67–82.

Wolpert, L. and D. Skinner (1993). *The triumph of the embryo*. Oxford: Oxford university press.

Woodward, A. L., A. T. Phillips and E. S. Spelke (1993). Infants' expectations about the motion of animate versus inanimate objects. *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 1087-1091). Boulder, CO.

Worrall, J. (2004). Science discredits religion. In M. L. Peterson and R. J. VanArragon (Eds.), *Contemporary debates in philosophy of religion* (pp. 59-71). Malden (Mass.): Blackwell.

Wu, Z. and H. Zhao (2009). Statistical Power of Model Selection Strategies for Genome-Wide Association Studies. *PLoS Genet, 5*(7), e1000582.

Young, M. (2004). Grand designs and facile analogies. Exposing Behe's mousetrap and Dembski's arrow. In M. Young and T. Edis (Eds.), *Why intelligent design fails: A scientific critique of the new creationism* (pp. 20-31). New Brunswick, New Jersey and London: Rutgers University Press.

Young, M. and T. Edis (2006). *Why intelligent design fails : a scientific critique of the new creationism*. New Brunswick, N.J.: Rutgers University Press.

Zaleski, P. and C. Zaleski (2005). *Prayer : a history*. Boston, New York: Houghton Mifflin Co.

Zuefle, D. M. (1999). Tracking Bigfoot on the Internet. *Skeptical Inquirer, 23*, 26-29.

Zygmunt, J. F. (1970). Prophetic failure and chiliastic identity: The case of Jehovah's Witnesses. *The American Journal of Sociology, 75*(6), 926-948.