

Introduction

In his 1814 *Philosophical Essay on Probability*, Pierre-Simon Laplace wrote:

An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes.

Laplace suggests that given the right basic information, and sufficiently powerful reasoning, all truths about the universe can be determined. For Laplace, this basic information included truths about the fundamental laws of physics and truths about the location of all fundamental entities at a time. Let us call these the *Laplacean truths*. The reasoning requires an idealized ‘vast enough intellect’, which we might call a *Laplacean intellect*. A Laplacean intellect who knows all the Laplacean truths is a *Laplacean demon*.

The key claim in Laplace’s text is that a Laplacean demon would be uncertain of nothing. In effect, Laplace is saying that for any proposition that the demon can entertain, the demon will not be uncertain about that proposition. Or in a small variation: for any proposition the demon can entertain, the demon will be in a position to know whether it is true.

Suppose that there will be an election tomorrow. I can entertain the proposition that the left-wing candidate will win, the proposition that the right-wing candidate will win, and the proposition that the third-party candidate will win. If Laplace’s thesis is right, a Laplacean demon in my shoes will be able to know which if any of these propositions is true. If the left-wing candidate will win, the demon will be in a position to know it; if the right-wing candidate will win, the demon will be in a position to know it; if the third-party candidate will win, the demon will be in a position to know it.

Laplace’s thesis is an instance of what I call a *scrutability* thesis. It says that the world is in a certain sense comprehensible, at least given a certain class of basic truths about the world. In particular, it says that all truths about the world are *scrutable* from some basic truths. This means roughly that there is a connection in the realm of knowledge between the basic truths and all the rest: given the basic truths, the rest of the truths can be determined.

We might then put a version of Laplace’s thesis as follows:

Laplacean Scrutability: For all true propositions p , a Laplacean intellect who knew all the Laplacean truths would be in a position to know that p .

In the years since Laplace wrote, Laplace's demon has come in for something of a battering. But I think that there remains much value in Laplace's pregnant idea. One can extract some of the value by examining the problems that arise for Laplace's thesis, and by reformulating the scrutability thesis in a way that avoids them.

One sort of problem arises from the information that Laplace allows in the base. Most famously, the apparent failure of determinism in quantum mechanics suggests that the demon could not predict the future just from facts about physical laws and about the present. It may be, for example, that futures in which the left-wing candidate, the right-wing candidate, and the third-party candidate win are all left open by these facts. All three futures could evolve from the present state of the world given the right sort of quantum-mechanical evolution.

There are other limitations. Many have argued that complete physical information is not enough to know all truths about the mind: if Laplace's demon has never experienced colors, for example, it will not know what it is like to see red. Others have argued that complete objective information is not enough to determine perspectival truths about the current time, or one's own identity: even given complete physical information, Laplace's demon might not know that today is Tuesday. Others find gaps for mathematics, morality, and other areas.

To avoid these problems, however, we need only give Laplace's demon more information than Laplace allows. To accommodate nondeterminism, we might give the demon full information about the distribution of fundamental physical entities throughout space and time. To handle problems involving the mind and the self, we might give the demon information about consciousness or the principles governing it, along with information about its own location in spacetime. If there are gaps for mathematics or morality, we can give Laplace's demon mathematical and moral principles as well. It is not clear precisely what information is required, but here the key claim is that there is *some* limited class of base truths that will allow Laplace's demon to do its work.

We might say that a *compact* class of truths is a set of truths that involves only a limited class of concepts and that avoids trivializing mechanisms such as coding the entire state of the world into a single number. I will elaborate on this rough characterization in the first chapter. For now, we can say that the class of physical truths will be a compact class, as will the expanded class of truths suggested above. We can then put a generalized Laplacean thesis as follows:

Inferential Scrutability: There is a compact class of truths such that for all true propositions p , a Laplacean intellect who knew all the truths in that class would be in a position to know that p .

Inferential Scrutability allows a broader class of base truths than Laplacean Scrutability, but otherwise it shares a similar form. In both theses, the idea is that if the demon knew all the basic truths, it could come to know all the truths, perhaps by inference from those basic truths. For the demon to know all the basic truths, they must be true in the demon's own world. So Inferential Scrutability in effect requires that the demon inhabits our world or one very much like it, knows all the basic truths about it, and comes to know all truths from there.

This requirement gives rise to a second sort of problem for Laplace's demon. In the actual world, we may suppose, one truth is that there are no Laplacean demons. But no Laplacean demon could know that there are no Laplacean demons. To avoid this problem, we could require the demon to know all truths about its own modified world rather than about the actual world. But now the demon has to know about itself, and a number of paradoxes threaten. There are paradoxes of complexity: to know the whole universe, the demon's mind needs to be as complex as the whole universe, even though it is just one part of the universe. There are paradoxes of prediction: the demon will be able to predict its own actions and then try to act contrary to the prediction. And there are paradoxes of knowability: if there is any truth q that the demon never comes to know, perhaps because it never entertains q , then it seems that the demon could never know the true proposition that q is a truth that it does not know.

To avoid these paradoxes, we can think of the demon as lying outside the world it is trying to know. Or better, we can think of the demon as contemplating the universe conditionally: *if* the Laplacean truths obtain, *then* this is what follows. Even if our own world does not contain a demon, we can still ask what a demon in some other world could come to know about our world, if it were given the relevant information in conditional form. Such a demon need not contemplate its own existence. What results is a conditional version of the scrutability thesis.

Conditional Scrutability: There is a compact class of truths such that for all true propositions p , a Laplacean intellect would be in a position to know that *if* the truths in that class obtain, then p .

We can make one final change. A key element of Laplace's idea is that the Laplacean truths are *all* the truths that the demon needs. No other empirical information is needed for the demon to do its job. Here, the idea is that to know the conditional above—*if* the basic truths hold, then p holds—the demon does not need any further empirical information in the background. That is, to know the conditional, the demon need not rely on a posteriori sources such as perception, introspection, or testimony. The demon can know the conditional a priori, or with justification independent of experience. We might put this as follows.

A Priori Scrutability: There is a compact class of truths such that for all true propositions p , a Laplacean intellect would be in a position to know a priori that if the truths in that class obtain, then p .

The three preceding theses are all scrutability theses. They say that there is a compact class of basic truths from which all truths can be determined, given sufficiently powerful reasoning. The A Priori Scrutability thesis is the most important for my purposes, so I will sometimes refer to it as simply ‘the scrutability thesis’. But the other theses above will also play a role.

All sorts of questions immediately arise. How can these scrutability theses be made precise? Why should we believe them, and how can one argue for them? Just which truths are among the basic truths, and how small can the basis be? What about hard cases, such as knowledge of social truths, or moral truths, or mathematical truths? What does the scrutability framework tell us about language, thought, knowledge, and reality? All of these questions are pursued in this book.

I suspect that to many readers, the scrutability theses just discussed will seem obvious. I hope that they at least seem plausible to many more. But to paraphrase Russell, philosophy is the art of moving from obvious premises to interesting conclusions. Even if scrutability theses are obvious, there are many interesting conclusions downstream from them. Of course theses requiring only a compact base do not do everything that Laplace’s stronger thesis could do. If a demon is given a full specification of how basic physical entities are distributed throughout space and time, for example, then its ability to predict the future is less impressive than it would have been for Laplace’s original demon. I think that nevertheless, the weaker thesis is powerful, because of its many applications.

Applications to epistemology, the study of knowledge, are perhaps the most obvious. For example, the scrutability thesis is at least a cousin of the knowability thesis, the thesis that all truths can be known. In addition, I will argue later that a version of the scrutability thesis can help with the problem of skepticism about the external world.

There are also applications in many other areas. In metaphysics, specific versions of the scrutability thesis can be used to help determine what is true and what is fundamental. In the philosophy of science, the scrutability thesis can be used to shed light on reductive explanation and the unity of science. In the philosophy of mind, the scrutability thesis can be used to help understand primitive concepts and the content of thought. And perhaps most importantly, the thesis has powerful applications in the philosophy of language, helping us to analyze notions of meaning and content that are tied to thought and knowledge.

In fact, the scrutability framework bears directly on many of the central debates in philosophy. One version of the thesis can be used to defend a Fregean approach to meaning (an analysis of meaning grounded in rationality and the a

priori) over a purely Russellian approach (an analysis grounded in reference and the external world). Another can be used to defend internalism about mental content, defining a sort of content that is largely intrinsic to the subject, against a strong externalism on which all content depends on the environment. Another can be used as a key premise in an argument against materialism about consciousness. Another can be used to deflate many traditional skeptical arguments about knowledge. Another can be used to support a version of structural realism about science.

Different versions of the scrutability thesis are relevant to different applications, so the issues do not all stand and fall together.¹ But in different ways, scrutability provides a powerful fulcrum through which we can gain leverage on these issues. In some cases, one can make related arguments without a direct appeal to scrutability, so the conclusions are not wholly beholden to the scrutability framework and can be cast in different terms. But in every case, thinking in terms of scrutability reframes the issues in a way that can make old, murky problems a little clearer.

The scrutability framework tends in a direction contrary to a number of trends in post-1950 philosophy: trends including direct reference theories of meaning, externalism about mental content, and rejection of the analytic/synthetic distinction. In various respects, it helps to support ideas from an earlier era in philosophy. It supports Gottlob Frege's distinction between sense and reference, and helps provide a concrete account of what Fregean senses are. It coheres well with Bertrand Russell's ideas about constructions of the external world and about the role of acquaintance in thought and knowledge. And above all, it provides support for many key ideas of the great logical empiricist, Rudolf Carnap.

In many ways, Carnap is the hero of this book. Like the other twentieth-century logical empiricists, he is often dismissed as a proponent of a failed research program. But I am inclined to think that Carnap was fundamentally right more often than he was fundamentally wrong. I do not think that he was right about everything, but I think that many of his ideas have been underappreciated. So one might see this project, in part, as aiming for a sort of vindication.

The title of this book is a homage to Carnap's 1928 book *Der Logische Aufbau der Welt*, usually translated as either *The Logical Construction of the World* or *The Logical Structure of the World*. The title (like Carnap's?) should be heard as self-consciously absurd. I am not really constructing the world. But one can see the

¹ The application to Fregeanism requires a generalization of the A Priori Scrutability thesis already stated, while the other four applications respectively involve what I later call Narrow, Fundamental, Nomic, and Structural Scrutability respectively. The first application is outlined in the eleventh excursus, and the other four applications are discussed in and around chapter 8.

current book as trying to carry off a version of Carnap's project in the *Aufbau*: roughly, constructing a blueprint of the world, or at least constructing a blueprint for a blueprint, by providing a vocabulary in which such a blueprint can be given, and making a case that the blueprint would truly be a blueprint for the world. More specifically, the aim is to specify the structure of the world in the form of certain basic truths from which all truths can be derived. To do this, I think one has to expand Carnap's class of basic truths and change the derivation relation, just as we had to for Laplace. But with these changes made, I think that the project is viable and that some of the spirit of the *Aufbau* remains intact.

I did not set out to write a Carnapian book. Instead, the connections between my project and Carnap's crept up on me to the point where they could not be ignored. The connections to Carnap go beyond the *Aufbau*. The approach to Fregean sense in terms of intensions is very much a descendant of Carnap's approach in *Meaning and Necessity*. The reply to Quine in chapter 5 can be seen as an adaptation of Carnap's analysis in 'Meaning and Synonymy in Natural Language'. My approaches to the unity of science, to ontology, to skepticism, to inferentialism, and to verbal disputes all have something in common with different elements of Carnap's work. In some cases I was not conscious of the connection to Carnap until well into the process, but his presence here is clear all the same.

I should not overstate the extent to which my views and my motivations are Carnap's. I am not a logical empiricist or a logical positivist. I do not share Carnap's sometime inclination toward verificationism and phenomenalism. Where Carnap invokes a semantic notion of analyticity, I invoke an idealized epistemological notion of apriority. Logic plays a less central role for me than for Carnap, and unlike him, I eschew explicit definitional constructions. Carnap would not have approved of my views on the mind–body problem. Where Carnap saw the *Aufbau* as an attempt to make the content of science wholly objective and communicable, vindicating science serves less as a motivation for me, and my version of the project has subjective and nonstructural elements right in the base.

So this book picks up only on certain strands in Carnap, and not on his project as a whole. To oversimplify, one might say that where Carnap leans toward empiricism, I lean toward rationalism. The project as a whole might be seen as a sort of Carnapian rationalism. To some, that label might seem oxymoronic, but this just brings out that there is more to Carnap than traditional caricatures may suggest.

That said, I would like to think that those who share more of Carnap's empiricism than I do will find that there are still many elements of the current picture that they can accept. Later in the book, I discuss ways in which a version of this project might be used to vindicate something quite close to the Carnapian picture, coming as close as possible to the structural and definitional picture in the *Aufbau*.

Here is roughly what happens in this book. Chapter 1 introduces the project using the *Aufbau* as a guide. I go over various objections to the *Aufbau*, and sketch a version of the project that has the potential to overcome all these objections. This chapter in effect motivates and gives an overview of the project of the book as a whole. Chapter 2 goes over preliminaries, formulating scrutability theses in detail and addressing a number of other preliminary issues.

Chapters 3 and 4 mount the core arguments for scrutability. I argue for a limited scrutability thesis concerning the scrutability of all ‘ordinary truths’ from a certain base. Chapter 3 focuses on Inferential and Conditional Scrutability, using a hypothetical device, the Cosmoscope, to make things vivid. Chapter 4 extends these arguments to A Priori Scrutability. Many epistemological issues come up along the way in these chapters, and numerous objections are addressed.

Chapter 5 uses the framework to respond to Quine’s arguments against analyticity and apriority, by providing an analysis of conceptual change. Along the way, it develops a notion of meaning inspired by Carnap and grounded in the scrutability framework. Chapter 6 extends the arguments of chapters 3 and 4 to the scrutability of all truths, by considering various ‘hard cases’ such as mathematical truths, normative truths, intentional truths, ontological truths, and many others.

Chapters 7 and 8 investigate the character of a scrutability base. Chapter 7 tries to whittle down the base to the smallest possible class, proceeding through various domains to see whether they involve primitive concepts and need to be in the base or whether they can be eliminated. Chapter 8 builds on this to investigate the prospects for certain principled scrutability theses, in part to see to what extent the projects of Carnap and Russell can be vindicated, and in part to develop various applications. I see these two chapters as perhaps the central chapters of the book. Chapter 7 goes over many important issues concerning what should be in the base, while chapter 8 gives a sense of the upshot and rewards of the project. A summation after chapter 8 reviews the prospects for *Aufbau*-like projects, arguing that projects in the spirit of Carnap and Russell look surprisingly good.

Along the way, a series of excursuses after each chapter explore all sorts of connected issues. ‘Excursus’ is usefully ambiguous between ‘a detailed discussion of a particular point in a book, usually in an appendix’ and ‘a digression in a written text.’ Some of my excursuses are detailed discussions of points within the framework. Some of the more important excursuses in this group are the third (on sentential and propositional scrutability), the fifth (on insulated idealization), and the fourteenth (on epistemic rigidity and super-rigidity). Others are digressions that draw connections to the philosophical literature or develop applications. The more important excursuses in this group include the eleventh (on meaning), the fifteenth (on skepticism), and the sixteenth (on metaphysics).

The excursuses can be read in order along with the rest of the book, but they can be read in many different ways, and they can also be skipped as the reader pleases.

I originally intended that this book would contain a ninth chapter on verbal disputes, bringing out a way to use the scrutability framework to help resolve philosophical debates and shedding more light on primitive concepts and the analytic/synthetic distinction. The approach taken there also conveys a more flexible and dynamic version of the framework, without as much philosophical baggage as the earlier chapters. In the end I have left that chapter out: the book is long enough already, and that chapter (which has been published as a separate article) is not quite essential to the overall narrative. Still, I think of that chapter as part of this book in spirit, and it can be found online as part of an extended edition of this book. The same goes for four additional excursuses that I have ended up omitting: on inferentialism and analyticity, Twin-Earthability and narrow content, reference magnets and the grounds of intentionality, and conceptual analysis and ordinary language philosophy.

I have been asked a few times what area of philosophy this book falls into. The answer is not obvious, even to me. The book is an unholy stew of epistemology, philosophy of language, metaphysics, and philosophy of mind, with some philosophy of science and metaphilosophy thrown in along the way. But it approaches each of these areas in a distinctive way and with the other areas in mind.

Scrutability theses concern knowledge, so epistemology is at the heart of the project. But the analysis of knowledge, justification, and related notions, which form the core of contemporary epistemology, are only occasionally in focus here. Rather, I am doing a sort of *metaphysical epistemology* (or should that be epistemological metaphysics?): roughly, epistemology in service of a global picture of the world and of our conception thereof.

The metaphysical epistemology in this book breaks down into a number of components. To a first approximation, the early chapters (especially 3 and 4) focus on *global epistemology*: articulating and supporting general theses about what can be known and about the epistemological relations between truths about the world. The middle part of the book (especially chapter 5 and thereabouts) focuses on *epistemological semantics*: understanding notions of meaning and content that are tied to epistemological notions such as rationality and the a priori. The latter part of the book (chapter 7 onward) focuses on *conceptual metaphysics*: roughly, investigating the structure of our conception of reality, with one eye on how well this structure corresponds to reality itself.

The global epistemology in the early chapters serves as the motor that drives the arguments for scrutability for those who are skeptical. Scrutability theses can be seen as global epistemological theses akin to knowability theses and the like. I start by articulating these theses, and then try to argue for them in detail.

Along the way, a lot of epistemology takes place: epistemological issues about warrant, self-doubt, idealization, skepticism, conditionalization, evidence, recognitional capacities, inference, and the a priori take center stage.

The conceptual metaphysics of chapters 7 and 8 serves as the culmination of the book, giving a sense of the full picture that emerges for those who are sympathetic. Here the aim is to boil down our conception of reality to its most basic elements, isolating primitive elements in which our concepts are grounded, and to draw out consequences for mind, language, and reality. The sixteenth excursus draws out the application to issues in metaphysics, fleshing out the projects of conceptual metaphysics and connecting the epistemological notion of scrutability to the related metaphysical notions of supervenience and grounding.

The epistemological semantics of chapter 5 and the excursuses that follow gives a sense of one important application of the framework. Chapter 5 serves to motivate the framework of epistemically possible scenarios and intensions defined over them. The tenth excursus develops the modal framework in more detail. The eleventh excursus develops the semantic framework a little further and argues that the intensions so defined can play many of the key roles of Fregean senses. In chapter 8, I argue that these intensions can serve as a sort of narrow content of thought.

I sketch semantic applications only briefly in this book, but I develop the semantic applications much further in a forthcoming companion volume, *The Multiplicity of Meaning* (and also in various existing articles on which that book is partly based). Where this book starts with Carnap, that book starts with Frege, developing a Fregean approach to language and an internalist approach to thought. There the framework of epistemic two-dimensional semantics, which is itself grounded in the framework of scrutability, plays a central role. The books are written so that either can be read independently of the other, but I think that they work especially well together. They can be read in either order, proceeding either from epistemological foundations to semantic applications or vice versa.

I expect that there will also be a third book at some point, exploring related issues about modality and metaphysics. That book will develop the framework of epistemically possible scenarios, explore its relationship to the space of metaphysically possible worlds, and explore connections to related metaphysical issues. Between them, these three books can be seen as forming a sort of trilogy on the three vertices of the ‘golden triangle’ of reason, meaning, and modality.

The ideas in these books have grown indirectly out of some ideas in my 1996 book *The Conscious Mind*. An early version of the scrutability thesis is explored in chapter 2 of that book, as is a version of the two-dimensional semantic framework that plays a central role in *The Multiplicity of Meaning*. Some of the central themes in the early chapters got their initial airing in the 2001 article ‘Conceptual Analysis and Reductive Explanation’, co-authored with Frank

Jackson, an article that was at least putatively driven by issues about the mind–body problem.

Despite this connection, it would be a mistake to think of this book as intended mainly to provide a foundation for arguments about the metaphysics of consciousness. If I had been trying to bolster those arguments, I would have written a very different book. In this book, purely metaphysical issues (conceptual metaphysics aside) are most often in the background, while epistemological and semantic issues are in the foreground. In a few places I have articulated theses that might connect the epistemology to the metaphysics (notably the Fundamental Scrutability, Apriority/Necessity, and Conceptual/Metaphysical theses), but I have not tried to argue for them at any length. I have devoted much more energy to arguing for weaker scrutability theses that even thoroughgoing physicalists can accept. The stronger theses and associated metaphysical issues come into focus briefly in chapter 8 and the sixteenth excursus, but they will be more central in the book on modality mentioned above.

It would be somewhat closer to the mark to think of this book as intended to provide a foundation for the ideas about two-dimensional semantics that I have developed in other work. It has gradually become clear to me that the key issue here is scrutability: once an appropriate scrutability thesis is accepted (as I argue in the eleventh excursus), a version of the epistemic two-dimensional framework follows. In fact, this book started its life as a chapter or two in *The Multiplicity of Meaning*, before taking on a life of its own. Still, by now I think that the scrutability thesis has interest for all sorts of purposes, and that while the applications to the theory of meaning and content are important, there are certainly many others as well.

I have tried not to assume too much in the way of theoretical principles from the start. Instead, I have tried to proceed by working through cases and mounting arguments to see what sort of theses emerge at the other side. In this way my approach differs from that of Carnap in the *Aufbau*, who starts with a strong structuralist thesis. I was tempted to write another version of this book, one that first articulates one of the principled scrutability theses in chapter 8 and then uses it to drive a construction from the ground up while also defending it from objections. That principled approach would have been more theoretically elegant and cohesive. But the relatively unprincipled approach of the current book has the advantage of letting the chips fall where they may. This way, by the end of the book we are in a position to judge the prospects for numerous different principled approaches.

Of course I do not proceed with complete philosophical neutrality. There is no such thing, and the discussion here is inevitably filtered through my own philosophical sensibilities. Still, I have tried to acknowledge alternative viewpoints where I can, to find a way for opponents to come at least part of the way with

me if possible, to argue against them where I can, and to see where they will get off the bus if they must.

My philosophical sensibilities play a role when I consider some of the most famous arguments in recent philosophy: Quine's arguments against analyticity and the a priori, Kripke's arguments against Fregean views, Putnam's and Burge's arguments against internalism. I use the scrutability framework to rebut some of these famous arguments and to limit the consequences of others, thereby defending key elements of the traditional views (internalism, Fregeanism, belief in the a priori, and so on) against which these arguments are directed. I generally take it that the traditional views here have a sort of default status, so that if they are to be rejected it must be on the basis of argument. Because of this way of proceeding, I do not know how much I will do to bring around someone who is entirely unsympathetic with the traditional views, not as a matter of argument but as a matter of starting point. But I am happy enough for now with the conclusion that if these views (or the versions of them that I accept) are wrong, it is for reasons that are interestingly different from the familiar reasons that I argue against.

More generally, it will not surprise me if some of the key conclusions in this book are wrong. Even if ideal reasoners can be certain of the philosophical truth, I am not an ideal reasoner. But I hope that if I am wrong, it is not for old reasons, or not only for old reasons, but also for new and interesting reasons that lead to new and interesting philosophy.

That said, I think of scrutability as supporting a sort of philosophical optimism. Conditional on knowledge of certain fundamental truths and ideal reasoning, everything can be known. In particular, this means that any failures of philosophical knowledge can be ascribed either to the non-ideality of our reasoning or to our ignorance of fundamental truths. Now, it is far from clear to what extent the fundamental truths are knowable, and it is far from clear to what extent we approach the ideal in relevant respects. Still, it is also far from clear that fundamental truths are beyond our grasp, and it is far from clear that reasoning that is needed to determine philosophical truths is beyond our grasp. Philosophy is still young, and the human capacity for reasoning is strong. In a scrutable world, truth may be within reach.