

Chapter 6: Hard Cases

David J. Chalmers

1 Introduction

In this chapter, I consider hard cases: truths about mathematics, morality, ontology, intentionality, modality, and a number of other domains. These are all domains outside the scope of chapters 3 and 4, such that there is some question about whether all truths in these domains are scrutable from more basic truths or from *PQTI*.

It is easy for the treatment of hard cases to be comically brief. In the *Aufbau*, Carnap's treatment of value and culture get a few pages each. The following passage gives the essence of his treatment of culture:

‘The custom of greeting through the lifting of one’s hat would perhaps have to be constructed in the following form: “The custom of “greeting through the lifting of one’s hat” is present in a society (or in some other sociological grouping) at a certain time, if, among the members of this society at that time, there is present a psychological disposition of such a kind that, in situations of such and such a sort, a voluntary act of such and such a sort takes place.”’

My own treatment of hard cases will share something of this comical character. So it is worth saying up front that I am not trying execute a construction or a reduction of truths in these domains, or to develop a positive account of the domains in question. All I will do here is consider, for each domain, whether there is a prima facie challenge to scrutability, and when there is such a challenge, consider the options for answering it. I will outline the views that I think are most plausible in these cases, but I will also indicate how things go under alternative views.

In each case, I will offer reasons to believe in the scrutability of the relevant truths. Often this will be a case for inferential or conditional scrutability in the first instance. As in chapters 3 and 4, this provides the basis of a case for a priori scrutability, and I will consider any specific

worries about a priori scrutability where they exist. In some cases, I will also consider how things look if the relevant truths are not scrutable, so that the scrutability base has to be expanded. This discussion will not settle the question of a scrutability base for once and for all. But I hope that it gives a sense of the lay of the land, and makes a case that some limited scrutability base will suffice. Readers with different views about specific domains might come to different verdicts about them, and can adjust their results accordingly.

When faced with a problematic sentence M —a putative mathematical or moral truth, say—we can first ask whether M (or $\neg M$) is scrutable from $PQTI$. There are at least four options available here. The first is *rationalism*, holding that M is a priori. The second is *empiricism*, holding that M is not a priori, but nevertheless is scrutable from more basic truths such as $PQTI$. The third is *indeterminism*, holding that neither M nor $\neg M$ is determinately true at all, so that M is indeterminate. The fourth is *expansionism*, holding that M is true and is not scrutable from $PQTI$, so that the scrutability base needs to be expanded.

I think that each of the first three options is right in some cases. These options are clearly no threat to scrutability theses. The fourth option does not threaten the compact scrutability thesis as long as the expansion is sufficiently limited, and as long as it is not taken too often. If it turns out that to accommodate three or four problem cases then we need to conjoin three or four new families with $PQTI$ in the scrutability base, then as long as there are no trivializing mechanisms introduced, the thesis will remain intact. Many of the applications of scrutability theses, especially to issues about meaning and content, will be unthreatened by such an expansion.

Still, I am interested in keeping the scrutability base as small as possible. Some applications of the scrutability thesis depend on a very small base: for example, some applications in metaphysics depend on Fundamental Scrutability, the claim that all truths are scrutable from metaphysically fundamental truths. My own view is that it is not obvious that $PQTI$ needs to be expanded at all, and that in any case expansion should be very rare, probably limited to metaphysically fundamental truths and primitive indexicals. I will not now attempt to defend this claim, but I will keep an eye on it in what follows.

Toward the end of this chapter, I will examine the question of whether all truths are scrutable not just from $PQTI$ but from a more stripped-down cousin $PQTI-$ that dispenses with macrophysical truths and truths about counterfactuals: it contains merely microphysical truths, phenomenal truths, laws involving these, a that's-all clause, and indexical truths. This in effect requires considering the “hard cases” of macrophysical and counterfactual truths. The project of seeing whether $PQTI-$ can be stripped down further still is the subject of the next chapter.

2 Mathematical Truths

Perhaps the hardest case for scrutability from *PQTI* is mathematics. It is plausible that all ordinary mathematical truths are a priori: at least insofar as they are knowable, they are knowable without justification from experience. But it might be argued that some mathematical truths are not knowable at all, even on ideal rational reflection. If so, then these truths are not knowable a priori, and will plausibly not be a priori at all. Furthermore, empirical information does not seem to help, so if these truths are not a priori, they are also not a priori scrutable from *PQTI*.

Most obviously, certain Gödelian statements in arithmetic are not provable from standard axioms, and one might argue that they are not knowable a priori, even under an idealization. Something similar goes for certain statements of higher set theory, perhaps including the Continuum hypothesis or its negation, which do not seem to be provable from obvious axioms. Likewise, in second-order logic there are sentences that are intuitively true, but that do not seem to be provable. If there are sentences in these domain that are true but not a priori, then they are not a priori scrutable from *PQTI*.

Still, it is far from clear that there are any mathematical truths that are not a priori. Given a putative example of such a truth, one can argue that either the truth in question is knowable under some idealization of rational reasoning, or that the sentence in question is not determinately true at all.

Take the Gödelian case first. Certainly, the fact that a statement is not provable from the Peano axioms does not entail that it is not knowable a priori. The Gödel sentence of the Peano axioms themselves is not provable from those axioms, but it is plausibly knowable a priori: it is equivalent to the claim that the axioms are consistent, and we plausibly know that the axioms are consistent in virtue of knowing that they are true. All that follows is that the Peano axioms do not encapsulate an axiomatic basis of our mathematical knowledge.

It might be suggested that there are more complex Gödel sentences that we cannot know. For example, if our brains are finite and computational, it is arguable that there is an axiomatic system that either encapsulates or exceeds our arithmetical competence. If that system is consistent, then its Gödel sentence will be a truth that we cannot know. But the construction of this example clearly stems from our contingent cognitive limitations. Just as we can know the truth of Gödel sentences for more limited systems, it is reasonable to hold that reasoners less limited than us could know the truth of this Gödel sentence.

Still, one might wonder whether this will always be so. Might there not be truths of arithmetic

that no reasoner could know a priori? I think that there are at least two good reasons to deny this.

First, Feferman's completeness theorem (1962) shows that the truth of any statement of arithmetic can be settled by repeatedly adjoining Gödel sentences to the Peano axioms a relevantly infinite number of times. This construction requires an ability to count through sufficiently large infinite ordinals, and for any finite creature this ability will give out at some point. But the ability gives out at different points for different creatures. It is not unreasonable to hold that for any such ordinal, some idealized creature could count that high, and so could know the relevant truth a priori. There is some question about whether the character of the repeated adjoining has to presuppose the arithmetical truths in question, however; if so, it is not obvious that this adjoining should be considered a priori.¹ So I will not rely on the Feferman result here.

More straightforwardly, if more distantly, one can idealize away from our capacity to consider only a finite number of cases at once. Russell famously said that our inability to count through all of the integers is a "mere medical impossibility". Likewise, we can at least conceive of a creature with infinite capacity for parallel reasoning. Whenever this creature has to evaluate the truth of $\forall n\phi(n)$ where n ranges over integers, it simultaneously evaluates $\phi(n)$ for each n (perhaps thinking faster for larger n , to ensure a bounded thinking time), and responds with the verdict "true" if and only if $\phi(n)$ is true for each n . No paradoxes arise here. In fact, there exist programming languages that can specify an algorithm to handle these cases using infinite branching and infinite conjunctions, although of course the algorithm could not be implemented effectively on a finite computer.² With this sort of process in place, one can demonstrate that the truth of any statement of arithmetic can be settled.

Of course this is a relatively taxing idealization, but it is not out of bounds. It is grounded in familiar cognitive capacities, such as the capacity to evaluate $\phi(n)$ for any given n , and is extended merely by dropping a limitation on parallel processing. One might worry that if we allow idealizations like this, then anything goes. But this seems wrong: even under this sort of idealization, there is no reason to believe that a statement such as 'There is a cup in this room' is knowable a priori. In any case, if the relevant notion of ideal reasoning is stipulated such that idealizations of this sort are allowed, then arithmetical truths will clearly be scrutable from any base.

The case of unprovable statements of higher set theory is murkier. Here, it is not at all clear

¹Thanks to Hartry Field for discussion here. See also Torkel Franzen's book *Inexhaustibility*.

²See, for example, work on infinitely synchronous concurrent algorithms by McConnell and Tucker 1992, and on deciding arithmetic using SAD computers by Hogarth 2004. Beggs and Tucker (2006) even establish that infinitely parallel computations such as these can be embedded in Newtonian kinematics.

that the relevant statements are determinately true or false. The most common view among set theorists appears to be that they are indeterminate. Even if they are determinate in some cases, it is not out of the question that further axioms may settle the determinate truths, and that these axioms could be known either by us or by more ideal creatures. In addition, idealizations akin to the one above suggest that relevantly ideal reasoning may take us beyond what is provable from axioms. So at least, there is not much strong reason to believe that some statements of set theory are both determinately true without being a priori, though the matter deserves further investigation.

The case of second-order logic may be something of a mix. Here, in some cases we have clear intuitions of truth despite the lack of a proof from widely accepted axioms. For the reasons given above, we should not equate knowability in mathematics with provability. So in some cases it is reasonable to say, as in arithmetic, that the relevant sentences are both true and knowable a priori. In other cases in which we lack clear intuitions, the situation is more akin to that in set theory. In these cases it may be that the sentences are indeterminate, or it may be that an idealization would enable knowledge of these sentences. In any case there is little reason to believe in truths that are not a priori in this domain.

In the worst case, if some mathematical truths are not a priori, then some such truths need to be added to the scrutability base. The vocabulary expansion needed may be small or even nonexistent, as *PQTI* already includes some mathematical vocabulary (used in stating *P*, for example). The precise nature of the needed mathematical sentences and vocabulary will depend on delicate issues involving the correct treatment of the domains discussed above. But the general framework of scrutability will not be threatened.

3 Normative truths

What should say say about moral truths, and other normative and evaluative truths? One could ask the question: are moral truths scrutable from *PQTI*? But it is easier to ask the more general question: are moral truths scrutable from non-moral truths? If they are, then moral truths will not pose a distinctive problem for scrutability.

On the face of it, there are good grounds to hold that insofar as there are moral truths and they are knowable, then they are scrutable from non-moral truths. Certainly, given that moral truths are knowable at all, they appear to be inferentially scrutable and conditionally scrutable: given full enough knowledge of the non-moral properties of a situation, we are in a position to know its moral properties. Applying the argument from chapter 4, if conditional scrutability holds here, a

priori scrutability plausibly holds too. Moral truths will be inscrutable from non-moral truths only if some crucial principles or conditionals governing inferences from non-moral truths to moral truths are unknowable, or if any inference from non-moral truths to moral truths has an irreducibly empirical justification. But there is little reason to believe in unknowable moral principles here, and there is little reason to believe in such an irreducibly empirical justification.

Perhaps the best reason to deny that moral truths are a priori scrutable from non-moral truths arises from the possibility of moral disagreement even among ideal reasoners who agree on the non-moral truths. It is not obvious that this sort of disagreement is possible: perhaps apparent moral disagreement always involves empirical disagreement, or nonideal reasoning, or merely verbal disagreement. But it is also not obvious that this sort of disagreement is impossible. *If* this sort of disagreement is possible, it is not easy to see how the truth of moral claims that are the object of disagreement will be scrutable from non-moral truths.

However, if this sort of disagreement is possible, then it is natural to hold that there is no fact of the matter about who is correct. That is, one will then naturally embrace a form of moral anti-realism, according to which there are no moral truths. If so, there will be no inscrutable moral truths. Alternatively, one might embrace some form of moral relativism, so that moral sentences are adjudged true insofar as they are true according to an appropriate standard (that of a speaker, or an assessor); but then one can argue that according to that standard, a conditional from nonmoral truths to the moral sentence in question will also be adjudged correct (and a priori). Either way, there is no trouble for scrutability here.

What sort of meta-ethical views are incompatible with a priori scrutability of moral truths from nonmoral truths? Moral rationalism is clearly compatible with scrutability, as are those varieties of moral anti-realism on which moral sentences are not true. Forms of moral anti-realism where moral concepts are response-dependent in some fashion are also compatible: moral truths here will be scrutable from nonmoral facts about causes and responses. Many forms of moral empiricism are compatible: they will typically involve at least the conditional scrutability of moral truths from nonmoral truths, and need not deny the inference to a priori scrutability.

One tricky case arises from forms of moral realism that hold that there are a posteriori necessities of the forms 'goodness = X', where X is a nonmoral expression. This claim alone is quite compatible with a priori scrutability of 'moral' truths, just as the a posteriori necessity of 'water is H₂O' is compatible with a priori scrutability of 'water'-truths. For example, as Horgan and Timmons suggest, such a view is compatible with the a priority of 'If X actually regulates our positive moral responses, then X is goodness'. Then as long as the nonmoral truth that X regulates our pos-

itive moral responses is itself scrutable, 'X is goodness' will be scrutable. Of course scrutability is incompatible with a hardline form of this view on which there are no a priori entailments from nonmoral truths to moral truths. But the arguments in chapter 4 can themselves be seen as good reasons to reject such a view.

Another tricky case involves a moral sensibility theory, on which one must have a certain sensibility (certain emotional responses, say) in order to appreciate moral truths.³ This theory is compatible with scrutability: a proponent of the view may hold that moral truths are knowable a priori by those with the right sensibility, or at least that they are conditionally and a priori scrutable from nonmoral truths by those with the right sensibility. Still, it introduces a wrinkle into the picture. On this view, ideal reasoning will require the right sensibility, involving components that one might take to be emotional as well as traditionally rational. (One could say that an ideal reasoner needs a big heart as well as a big brain.) If this view is correct, the scrutability thesis is still fine, but it may have less of a rationalist upshot than one might have supposed.

The meta-ethical view that is most obviously incompatible with scrutability of the moral from the nonmoral is a hardline form of moral realism on which there are moral truths that are not knowable even on full knowledge of nonmoral truths and ideal reflection. Such a view is unattractive, though. The best reason for being a moral realist stems precisely from our apparent knowledge of moral truths. If that knowledge is denied, moral anti-realism seems much the more natural option.

What about truths in other normative domains, such as epistemic truths concerning what is rational, or what is justified, or what can be known a priori? Here, the situation is much the same as above. If there is a difference, it is there is less reason to believe in the possibility of significant ideal disagreement in this case. Epistemic disagreement typically appears to be more delimited than moral disagreement, and one can reasonably hope that much of it will disappear on ideal reflection. Of course the issue is delicate, as the notion of ideal reflection is itself an epistemic one, so that an epistemic anti-realist or relativist will hold that the notion of ideal reflection may itself be incoherent or subject-relative. But in any case, there is little reason to believe in inscrutable epistemic truths.⁴

³Thanks to Martine Nida-Rümelin for discussion here. Much the same applies to sensibility theories in aesthetics. More generally, most of what I say about the moral domain in this section can also be applied to the aesthetic domain.

⁴An interesting question: if epistemic anti-realism is true, where does that leave my epistemic conclusions of this book, such as the scrutability thesis? If an epistemic error theory is true, presumably most of these claims are simply false, although perhaps it remains possible that nearby reconstructed claims (such as claims about what is the case in a fiction) could be true. If epistemic expressivism is true, presumably this is compatible with my claims in the same way that meta-ethical expressivism is compatible with claims in normative ethics. I am an epistemic realist, though, and it

Something similar applies to other evaluative domains, such as the aesthetic domain. In this case, there is perhaps more reason to believe in the possibility of ideal disagreement, and so more reason to accept some sort of anti-realist treatment. But given that truths in this domain are knowable, there is little reason to believe in inscrutable aesthetic truths.

If someone disagrees, holding that there are inscrutable truths in one or more of these domains, then one will have to expand the scrutability base, most likely by including some normative or evaluative truths in the base. For example, certain fundamental moral, epistemic, or aesthetic principles might need to be included.⁵ Doing so will require a small expansion of the basic vocabulary, perhaps to include normative or evaluative expressions such as ‘ought’ or ‘good’. But this will not pose a threat to the overall scrutability thesis.

6

4 Ontological truths

Ontological claims are claims about the existence of entities. *PQTI* already includes many existence claims. It says there exist entities with various microphysical properties, entities with various phenomenal properties, and perhaps entities with various macrophysical properties too. In effect, it contains ontological claims about microphysical entities, subjects of experience, and possibly macrophysical entities.

Still, an important challenge to scrutability from *PQTI* concerns further ontological claims. One such challenge concerns the existence of abstract objects: for example, is the existence of numbers scrutable from *PQTI*? And if one is to move from *PQTI* to the stripped-down base *PQTI*–, macrophysical entities also pose a challenge: is their existence scrutable from a microphysical base? On some ontological views, even once the relevant base truths are fixed, it is

is arguable that epistemic realism gives the conclusions more interest and weight.

⁶Three residual worries here: (i) What if (for example) moral particularism is true, so that there is no finite set of moral principles that generates moral truths from nonmoral truths. Then one will have to include truths about the moral properties of individual cases in the base. (ii) What if moral principles make irreducible appeal to high-level nonmoral notions such as ‘kill’? If these are definable in base language, the principles can be stated without them, and if they are not scrutable from base language, they will need to be included in the base anyway. If they are scrutable but not finitely definable, one has the choice between including them in the laws (at cost of vocabulary expansion) or using an equivalent infinitary version of the principles (at cost of much longer base sentences in the original vocabulary). (iii) What about “thick” moral notions such as ‘brave’? In these cases I think it is plausible that truths involving these notions are scrutable from (although perhaps not definable in terms of) nonmoral truths and “thin” moral truths about values and norms.

epistemically possible that numbers or macrophysical entities exist, and epistemically possible that they do not. On such a view, the scrutability base will have to be expanded mildly to include existence claims about abstract or macrophysical objects, or to include general ontological principles covering them.

It is far from obvious that this sort of expansion is required. As before, it is plausible that insofar as ontological truths are knowable, they are conditionally and a priori scrutable from a limited base. This scrutability might be underwritten by a “lightweight realist” view of ontology on which existence claims can be analytic, a rationalist view on which basic ontological principles are a priori, or an anti-realist view on which there are no ontological truths at all.⁷ Still, there are also important “heavyweight realist” views on which the relevant ontological truths are sometimes not knowable at all. On such a view, ontological claims will not be analytic or a priori and will not be analytically or a priori entailed by the relevant base truths. There will also be other heavyweight realist views on which ontological truths are knowable only nonconclusively, so that even if they are a priori scrutable from relevant base truths, they are not conclusively a priori scrutable.

Carnap (1950b) rejected the heavyweight view of ontology. He held that ontological questions can be read in either an “internal” or an “external” sense: internal questions are settled analytically (as on a lightweight view), and external questions do not have objective answers (as on an anti-realist view). My view (Chalmers 2009) has a similar spirit: if existential quantification is understood in a lightweight way, then ontological truths are conclusively scrutable, and if it is understood in a heavyweight way, then ontological claims are indeterminate. Either way, I think that there is little reason to believe in inscrutable ontological truths of the sort that a heavyweight realist requires. If we embrace a view in this vicinity, there is no problem for scrutability.

What if heavyweight realism is true? In this case the scrutability base may need to be expanded slightly in order to cover positive truths about the existence of macrophysical or abstract objects. The most natural way for this to work is for certain general ontological principles to be added: perhaps principles saying that when certain microphysical conditions obtain, certain macroscopic entities exist. Even here it is not obvious that we need to expand the vocabulary. The crucial expression needed is the existential quantifier, but that is already included in our base.

The only potential expansion here is further vocabulary used to specify ontological principles. For example, the popular principle of universal mereological composition says that for any set of objects, there is a fusion of those objects, (an entity with those objects as parts such that any part

⁷See Chalmers 2009 for more on lightweight realism, heavyweight realism, and the sort of Carnapian anti-realism with which I am most sympathetic.

of the entity overlaps one of the objects). This principle uses mereological vocabulary, requiring especially the crucial notion of ‘part’. Now, it is plausible that truths about parthood are scrutable from more basic truths: perhaps *a* if part of *b* if *b* is located at every point where *a* is located, for example. Correspondingly, it is not out of the question that we can state ontological principles without an expansion of vocabulary: for example, for any objects, there is an object located at all and only points at which one of the original objects is located. But there are certainly views on which mereological or other vocabulary will be essentially required within these principles, in which case there will be a limited expansion. As long as the ontological principles are limited in their extent, however, there will be no threat to scrutability from a compact base.⁸

There are also questions about fundamental natural ontology: roughly, the fundamental concrete entities and properties that populate the world. These questions cannot easily be dismissed in the way that questions about numbers and macrophysical objects can be dismissed, because they clearly affect the character of a scrutability base. If Fundamental Scrutability is true, then there will be a scrutability base made up almost entirely of truths about fundamental entities and properties. And even if Fundamental Scrutability is false, one can expect that there will be a minimal scrutability case containing truths about these entities and properties along with other truths. Of course if physicalism is true, these entities may already be specified by *P*, and if dualism is true, they may already be specified by *P* and *Q*. But if more liberal ontological views are true, we may need to go beyond *P* and *Q*.

For example, if a god exists, then the base almost certainly needs to be expanded to specify the god’s existence and properties. If libertarian free will exists, the base may need to be expanded to include it. If there are angels or nonphysical ectoplasm in another realm, the same applies. If our physics is embedded in a more fundamental 26-dimensional protophysics, then the base may need to include protophysical truths. Of course if there is no god, libertarian free will, angels, ectoplasm, or protophysics, then no such expansion will be required. An especially promising case for expansionism concerns so-called *quiddities*: the intrinsic natures associated with microphysical properties such as mass and charge. I discuss quiddities at length in section 9 of chapter 7, so I defer further discussion until then. For now, I will take it for granted that our scrutability base needs to include all truths about the fundamental properties of fundamental entities.

There are also questions about relative fundamentality, or grounding. For example, metaphysicians debate whether the whole universe is more fundamental than its simple parts: monists hold

⁸Thanks to Robbie Williams for discussion here.

that whole grounds the parts, while pluralists hold that the parts ground the whole. Likewise, physicalists hold that the physical is fundamental and grounds the mental, while idealists hold that the mental is fundamental and grounds the physical, and dualists hold that both mental and physical are fundamental. One might hold that *PQTI* does not settle these questions: it is compatible with both monist and pluralist views, and with physicalist, dualist, and perhaps idealist views.

Here Carnap himself appeared to hold that there is no objective fact of the matter between the views in question: he holds that his basic truths are compatible with physicalism, dualism, and idealism. One might also hold that the only truths about relative fundamentality here are scrutable from truths about necessitation or scrutability. On that sort of view, idealism may be ruled out because *P* is not scrutable from or necessitated by *Q*, and physicalism is ruled out to the extent that *Q* is not scrutable from or necessitated by *P*, but the choice between monism and pluralism is left open, as there is mutual necessitation between properties of the whole and the properties of the parts. This view would reduce the scrutability of truths about relative fundamentality to the scrutability of modal truths, discussed below.

If one thinks that there are truths about fundamentality that go beyond truths about necessitation or scrutability (as I am inclined to), one can build these truths into the base. In fact, there is precedent for doing this already: we saw in the fifth excursus that perhaps the most attractive formulation of the that's-all truth involves saying "Those are all the fundamental truths". If one understands this claim as invoking metaphysical fundamentality, then these truths will be built into the base. *PTI* will in effect settle things in favor of physicalism (in either monist or pluralist form depending on whether *P* is phrased in terms of properties of the world or of its parts), while *PQTI* will settle things in favor of dualism. And even if one does not understand *T* in this way (perhaps understanding it in terms of conceptual fundamentality instead), then truths about metaphysical fundamentality may nevertheless be scrutable using *T*, at least on some views. At worst, one can include certain additional claims about metaphysical fundamentality in the base: perhaps, for example saying that *P* and *Q* obtains, that the truths in *P* are all the metaphysically fundamental truths, and that all positive truths are scrutable from *PQ*.

In any case, I think it is attractive for multiple reasons to endorse a scrutability base that includes some notion of fundamentality. A world-sentence using such a base might (at least) set out the fundamental truths about fundamental properties and entities, and say that they are the fundamental truths, the fundamental properties, and/or the fundamental entities. I return to the role of fundamentality in scrutability bases in chapters 7 and 8.

5 Other philosophical truths

What about truths in other areas of philosophy? An enormous number of philosophical questions are normative or ontological questions and will be covered by the discussion above: they will be either scrutable from *PQTI*, or at worst from an expanded base that expressions such as ‘ought’, ‘good’, ‘exists’, and/or ‘fundamental’. Still, there are other sorts of philosophical questions. To pick just two, the debate between Fregean and Russellian views in the philosophy of language or between internalist and externalist views in the philosophy of mind are not obviously questions about normativity or ontology.

I will not try to go through all philosophical questions here. But I can give a sense of the options. For any given philosophical debate, as before, the options include rationalism (one view or the other can be known to be true a priori), empiricism (neither view is a priori, but one is scrutable from underlying truths such as *PQTI*), indeterminism (there is no objective fact of the matter as to which view is correct), and expansionism (one side is right, and the scrutability base needs to be expanded or modified to reflect this).

In addition, where debates are concerned, a fifth option is *pluralism*: both sides are right concerning different notions. I argue in chapter 9 that pluralism is often the correct diagnosis of a philosophical debate. This diagnosis typically leads to two or more clarified debates, involving the two or more disambiguated notions. In principle, each of these debates will themselves be subject to the options above. But these debates are usually more tractable than the original debate, so that they will fall more easily under one of the four options above.

I think that all five options are applicable in some cases. On my own views, rationalism holds for fundamental normative questions, empiricism holds for questions about fundamental natural ontology (including questions about phenomenology), and indeterminism holds for ontological questions outside fundamental natural ontology. I think that pluralism holds for debates between the Fregean and the Russellian, the internalist and the externalist, and the compatibilist and the incompatibilist. Expansionism may be correct in some cases (such as the case of quiddities), but as before I would like to think that it is relatively rare.

Some may find the scrutability thesis particularly implausible where philosophy is concerned, given the extent of disagreement even among highly rational philosophers. Here it is worth keeping in mind that we are far from ideal, however, and that our experience of disagreement as nonideal reasoners is at best very weak evidence of what would happen on ideal reflection. And again, it is worth noting that perhaps the deepest debates in philosophy concern the normative and the

ontological. In these cases, at worst, if relevant normative and ontological truths are allowed in the base, the disagreements will plausibly be settled by base truths.

I cannot claim to have made a conclusive case here that all philosophical truths are scrutable from a compact base. Still, if all ordinary nonphilosophical truths are scrutable in this way, and if paradigmatic philosophical truths in normative and ontological realms are scrutable too, then the thesis that there are certain special truths in philosophy not scrutable from a compact base begins to look unattractive.

The scrutability thesis yields a distinctive metaphilosophical picture. If the thesis is correct, then all philosophical truths (like all other truths) can be settled by ideal a priori reasoning from base truths. We may be ignorant of these truths because we are ignorant of base truths, or because we are not ideal reasoners. But there is no third form of ignorance. Questions that cannot be settled by a priori reasoning from base truths have no determinate answer.

This principle is structurally analogous to Hume's thesis that claims not involving "abstract reasoning concerning quantity or number" and "experimental reasoning concerning matters of fact and existence" involve just "sophistry and illusion". Compared to Hume's thesis, the scrutability thesis allows a broader sort of reasoning from a broader base, and where Hume's thesis claim sophistry or meaninglessness, the scrutability thesis merely claims indeterminacy. So as with the logical empiricists' verification principle, the scrutability thesis is not as strong as Hume's thesis, but there remains a structural parallel.

Unlike Hume and the logical empiricists, I am not especially inclined to cast my metaphilosophical thesis in a prescriptive voice, at least initially. To the extent that I argue for the scrutability thesis by considering cases, it will have limited prescriptive force when applied to the same cases. And the force of the thesis will in any case depend on the character of the base. Still, once one has made a case for a principled scrutability thesis in many domains, its cumulative support might be seen as carrying at least some prescriptive force when applied to other domains. A sort of philosophical methodology that coheres well with this picture is developed in chapter 9.

6 Modal truths

What about modal truths: truths about what is possible or necessary? Here the analysis depends on what sort of modality is at issue. Where the nomological modality is concerned, fundamental laws of nature are already built into *PQTI*. We can stipulate that where *L* is a basic physical or psychophysical law, $\Box L$ is in *PQTI*, where the box stands for nomological necessity. It is

reasonably plausible that all truths involving nomological necessity can be derived from these, along with the rest of *PQTI*. On some Humean views, as I discuss in the next chapter, truths about nomological necessity will themselves be scrutable from non-nomic truths. But for present purposes there is no need to take a stand on the issue.

What about the epistemic modality, and in particular epistemic modality interpreted as apriority? So far we have usually talked of sentences being a priori, but here it is more natural to phrase relevant claims as ‘It is a priori that S’. I will assume that sentences of this sort are true (in a context) if and only if the embedded sentence S is a priori (in that context). Then it is plausible that true sentence of this sort is a priori. This follows from an S4 principle for apriority: If it is a priori that S, it is a priori that it is a priori that S. Here, the thought is that if one can come to know S with nonexperiential justification, it is a short step to knowing that there is nonexperiential justification for S, and this step does not require any further experiential justification. If this is right, the S4 principle follows, and all positive claims about apriority are themselves a priori.

What about negative claims about apriority, such as ‘It is not a priori that S’. Here, the thesis that all true sentences of this sort are a priori requires an S5 principle about apriority: if it is not a priori that S, it is a priori that it is not a priori that S. This principle is less clearly true than the S4 principle, but reasons for accepting it are given in the fifth excursus.

Finally, what about the metaphysical modality? Here, it is familiar that truths such as ‘It is necessary that water is H₂O’ and ‘It is necessary that Hesperus is Phosphorus’ can be true without being a priori. In these cases, though, it is highly plausible that these truths are a priori entailed by nonmodal truths: in particular, by ‘Water is H₂O’ and by ‘Hesperus is Phosphorus’. Given that the relevant expressions are rigid (*de jure*), sentences such as ‘If Hesperus is Phosphorus, it is necessary that Hesperus is Phosphorus’ is a priori. So as long as the relevant nonmodal truths are scrutable from *PQTI*, the modal truths are scrutable as well.

I think that what applies in these cases applies generally. It is plausible that all modal truths are a priori entailed by nonmodal truths. Some such truths are themselves a priori, while those that are a posteriori typically follow from nonmodal identity statements, or perhaps from statements about composition. For example, suppose with Kripke that the follow is a modal truth: ‘It is necessary that if Elizabeth II exists, she is the daughter of George VI’. This modal truth is arguably implied by the nonmodal truth that Elizabeth II is the daughter of George VI, and the a priori principle that people have their parents essentially. And even if the principle in question is not a priori, perhaps because there could be exceptions in odd hypothetical circumstances, then it remains very plausible that the modal truth at issue is a priori scrutable from *PQTI*.

Some philosophical views involve a stronger separation between the necessary and the a priori than one finds in the cases above. On these views, there may be modal truths that are not scrutable from nonmodal truths. For example, someone may hold that it is necessary that a deity exists, without its being a priori that a deity exists. Perhaps nonmodal properties of the deity will imply that it is omnipotent, omniscient, and so on, and so will imply that it is a deity. But these and all other nonmodal truths could be coherently combined with the claim that the deity exists only contingently. If so, the modal truth is not scrutable, and our scrutability base will have to be expanded by including modal expressions. I reject the existence of modal truths of this sort (see Chalmers 2002 for reasons), but the case in question at least provides an illustration.

7 Intentional truths

Our scrutability base *PQTI* builds in truths about phenomenal states. One might ask, why does it not build in truths about other mental states such as beliefs and desires? This is partly because I think such states are themselves scrutable from *PQTI*, and partly because building them in raises a worry about trivializing mechanisms. I discuss both of these matters below.

Are truths about belief scrutable from *PQTI*? On some reductive views, such as analytic functionalism or logical behaviorism, these truths may even be scrutable from *P*. But no such strong claim is required here. One does not even need the claim that the intentional is scrutable from the non-intentional, since it is plausible that the phenomenal is intentional through and through. This applies both to the phenomenology of perceptual experience and the phenomenology of believing itself, both of which will be specified in *Q*. So it is certainly not out of the question that a base consisting of physical and phenomenal truths will yield truths about beliefs and desires.

Here we can run a version of the Cosmoscope argument. Using the information in *PQTI*, the Cosmoscope will give us an image of a subject's behavior and underlying functioning in actual circumstances and many counterfactual circumstances. In addition, it will enable us to know just what it is like to be that subject, and what it would be like to be the subject in various counterfactual circumstances. In addition, it will give us complete information about the subject's environment. Will this be sufficient for us to determine what the subject believes? On the face of it, yes. Though one can reasonably question whether behavioral information alone suffices for "radical interpretation" of a subject's beliefs, once full phenomenal and environmental information is added, there is no clear reason to think that the subject's mental life is underdetermined. At least, it is not easy to see what sort of mental truth we might remain ignorant of, given all this information.

An opponent might appeal to Kripke's version of Wittgenstein's rule-following argument, suggesting that underlying states do not determine what a subject believes. Here, there are at least three relevant observations. First, Kripke understates the role of the phenomenology of cognition, briefly mentioning it and assimilating it to a "distinctive sort of headache". But the experience of addition is nothing like this, and arguably involves a cognitive phenomenology that is richly intentional (see Siewert 1998 and Horgan and Tienson 2002). Second, Kripke's argument centrally turns on the idea that there is no reductive analysis of intentionality in non-intentional terms, but no such analysis is needed here, both because the phenomenal is already intentional, and because scrutability does not require explicit analyses. Third, Kripke himself seems to appeal to a version of the scrutability thesis, holding that if truths about belief are not scrutable from the relevant base, then they are not strictly truths at all, at least in the ordinary sense.

One might also object that local physical and phenomenal information underdetermines belief content, due to its dependence on the subject's environment. But *PQTI* will specify the subject's environment, and assuming that ordinary truths about that environment are scrutable from *PQTI*, then insofar as such truths play a role in constituting the belief's content, there is no reason to doubt that the content is scrutable too. For example, knowing the physical and phenomenal truths about a subject might underdetermine whether their beliefs are about water or twin water, but once we know relevant truths about the H₂O in their environment, the matter will be settled.

What if I am wrong, and truths about what a subject believes are not scrutable from *PQTI*? Then one will have to add such truths to the scrutability base. The most obvious suggestion is to initially include arbitrary intentional states along with phenomenal states, yielding *PMTI* (with *M* for "mental"). This would get rid of any gap, but it introduces complications, in the guise of trivializing mechanisms.

The most obvious way to specify intentional states such as beliefs is to use a *propositional* vocabulary. If John believes that the cat is on the mat, then one might say that John stands in the belief relation to the proposition *The cat is on the mat*, or one might use more complex structures to specify Russellian or Fregean propositions that John stands in the belief relation to. But now the danger of trivialization arises. The first of these approaches will end up using arbitrary vocabulary elements, such as 'cat' and 'mat', in the scrutability base, albeit appearing only in certain restricted contexts. The other approaches will require arbitrary objects and properties to be specified in the base (in Russellian propositions), or else arbitrary Fregean senses (in Fregean propositions).

One danger is that once our base vocabulary include propositions, it is not a large step to the base truths including truths about the *truth* of these propositions, such as '*The cat on the mat* is

true', and so on. From such truths, it is plausible that arbitrary truths ('The cat is on the mat') follow a priori. If so, we have trivialization. To deal with this worry, one might initially stipulate that propositional vocabulary can only occur in base truths as objects of attitudes ('John believes p '), and not in any other way (' p is true'). One might also worry about sentences such as 'John believes p , and his belief is true'. Here one can also restrict the places that terms such as 'believe' and 'belief' can occur in base sentences. One can also simply bar terms such as 'true' and other devices of semantic evaluation ('false', 'refers', 'about') from the base vocabulary.

A related worry occurs whether or not we build intentional truths into the base. Suppose that statements about the truth of beliefs are scrutable from base truths, as they must be if the scrutability thesis is true. So 'John believes truly that p ' might be scrutable from statements about the success of John's belief that p , or the success of certain p -related behavior, or something much more complicated. Then one will be able to use the scrutability of these truths to yield the scrutability of all truths. In response, I think one can reasonably deny that this is a trivializing mechanism. The most natural way to obtain 'John believes truly that p ' is to obtain something like 'John believes p ' and ' p '. Even if there is some other way to obtain 'John believes truly that p ' without going directly through p , it is plausible that one will have to exploit the same sort of information about the world that one would need to exploit in order to know p . So there is no trivialization here.

A third worry arises if the base includes factive states such as knowledge. From 'John knows p ', it follows trivially that p is true. This does not automatically yield all truths, but it yields many, and others might follow using counterfactuals: 'If John considered p , he would come to know p ', and so on. In response, I think that factive states can reasonably be excluded from the base. It is plausible that truths about factive states such as knowledge are scrutable from truths about nonfactive states such as belief, along with truths about justification, external facts, and so on. So these need not occur primitively.

A fourth and related worry arises even for nonfactive states, if some sorts of externalism are true. For example, it is often held to be necessary that if a subject believes that water is wet, then water exists, and one might hold that something like this is a priori as well. This is particularly clear if beliefs are stated using Russellian propositions, as the existence of a Russellian proposition involving a certain object or stuff entails that that object exists. One may not be able to recover all truths using this method, but the mechanism still seems overly powerful as a way of recovering some external world truths.

To get around this worry, I think that if intentional states are to be specified in the base, they

should be specified in a non-externalist way. A natural suggestion is that one should specify a belief's primary intension, or something along those lines. If primary intensions are indeed narrow content, then this will not yield the worry about externalism. Furthermore, doing this will bypass any worries about vocabulary expansion, as primary intensions can be specified using the same vocabulary that is used to specify scenarios in general. And when primary intensions are combined with sufficient information about the external world, it is plausible that truths about wide content, truth, and so on are thereby scrutable.

A potential problem is that it is not clear that a primary intension can always be finitely specified: it will involve values at an infinite number of scenarios, most of which may have an infinite specification. Perhaps one could allow such long specifications in the scrutability base, but now it will be enormous, and this length will feed back into the size and number of scenarios in turn, which will then feed back into the length of the specification of beliefs. An alternative is to specify primary intensions directly, regarding them as properties or relations and using a rich vocabulary that can pick out such properties. Doing so will avoid the second complication, that of externalism, although of course it will face the first complication in a significant way. Still, one could argue as before that as long as the rich vocabulary for expressions denoting properties occurs only within belief contexts, and as long as we exclude 'true' and cognates from the base, then there is no danger of trivializing the thesis.

A second potential problem is a threat of dialectical circularity: I need the scrutability thesis to establish that there are primary intensions, but here I am appealing to primary intensions to support the scrutability thesis. I think that the situation is not really circular, though. If it turns out that that intentional states are not scrutable from *PQTI* and need to be built into the base, then we can initially build in an expansive class consisting of truths about whatever intentional states (narrow or wide) there are. Then there is no obstacle here to the claim that all truths are scrutable from the expanded base, and there is correspondingly no obstacle to the claim that all beliefs have primary intensions. At most, there is a worry about whether the scrutability base is compact. But once we have established scrutability from the expansive base, and the existence of primary intensions, then (via the scrutability of truths about all intentional states from truths about primary intensions and non-intentional truths), we can establish scrutability of all truths from a base whose intentional truths include at most truths about primary intensions.

Most of the above has been premised on the hypothesis that I am wrong in holding that all intentional truths are scrutable from *PQTI*. It should also be noted that even if I am right, a version of some of these issues may arise from the inclusion of phenomenal truths in the base. If

phenomenal truths are themselves intentional truths, then specifying them will involve specifying a certain intentional relation between subjects and intentional contents. And even if they are not themselves intentional truths, it is not implausible that they directly imply intentional truths, for example concerning the content of phenomenal states. Either way, there will be a trivialization worry in the background. For a given *S*, then if there is a phenomenal state *R* whose content is the same as *S*, then one could suggest that *S* might be near-trivially scrutable from a truth like ‘If I was to have *R*, it would be veridical’.

This issue is not obviously as broad in its scope as the issue concerning intentional states more generally, as it is far from obvious that for any sentence, there can be a phenomenal state with the same content. On some views, the content of phenomenal states is restricted to low-level properties such as colors and shapes, in which case the threat of triviality is minimized. But on others, phenomenal states can also represent a wide-range of high-level properties, perhaps represented in perception (Siegel 2006), or perhaps represented in cognitive phenomenology (Horgan and Tienson 2002, Siewert 1998). Even if this does not extend to all properties and all contents, the scope is broad enough to cause concern.

In any case, one can deal with the issue in the same way as for intentional states in general. If phenomenal states are to be characterized in the base in terms of their intentional content, one can require that these contents occur in base truths only as relata of relevant intentional relations. In addition, whether phenomenal states are so characterized or not, one can exclude notions such as ‘veridical’ from the base. It is plausible that phenomenal states are narrow states, in which case the worry about the environment does not arise. It is an open question how the narrow contents of phenomenal states are best characterized: perhaps in terms of intentional relations to certain privileged properties (Chalmers 2006), or perhaps in terms of something like primary intensions (Chalmers 2004). But as before, there is no obvious problem with building contents such as these into the base if necessary.

Another way to avoid trivialization for intentional truths and phenomenal truths is to appeal to psychophysical laws. It is plausible that intentional and phenomenal properties supervene on physical properties with metaphysical or nomological necessity. If so, there will be psychophysical principles or laws governing this supervenience. Then even if the relevant intentional or phenomenal truths here are not scrutable from physical truths, they will be scrutable from physical truths along with psychophysical laws. It is unlikely that these laws will have special clauses for each intentional content or phenomenal property. Rather, we can expect that the laws will be quite general. The laws might involve intentional or phenomenal notions and might quantify over concepts

or contents, but expressing the laws will not require using arbitrary concepts. If so, scrutability from physical truths and psychophysical laws will avoid trivialization and will retain a compact base.

8 Social truths

There are many expressions that are intimately linked to intentional concepts and terms, such that knowing their extension require knowing various intentional truths. This applies especially to *social* expressions such as ‘friend’, ‘money’, ‘law’, and so on. For John to be Fred’s friend, John and Fred must have certain attitudes. For a piece of paper to count as money, people in the community must have certain attitudes. To be a philosopher, one must engage in certain patterns of thought, and so on. Insofar as there is an epistemic gap between *PQTI* and intentional truths, there will be a corresponding epistemic gap between *PQTI* and truths involving some of these notions. But once intentional truths are within the fold, there is no residual problem.

Here we can picture an extended Cosmoscope that delivers intentional truths to us, allowing us to know what every individual believes, desires, intends, and so on. Given this sort of extended Cosmoscope, there would be no obstacle to determining who is whose friend, who has what sort of money, what the laws of a society are, and so on. All of these phenomena are constituted by attitudes and practices among the members of a society, and the extended Cosmoscope will give us full access to those attitudes and practices.

This does not require us to take a position on questions about holism and individualism in social science. A holistic view holds that truths about societies are not reducible to truths about individuals. But we have seen that scrutability does not require reducibility. Biological truths about life may not be reducible to microphysical or phenomenal truths, but they are scrutable from these all the same. The same goes for social truths and individual truths. Even holists typically endorse some form of supervenience: social truths cannot vary without varying truths about individuals. We can put this claim in an epistemological key: we cannot get an imaginative grip on holding the truths about individuals’ attitudes and practices (along with the rest of *PQTI*) constant, while varying the social truths. This strongly suggests that social truths are scrutable from *PQTI* and intentional truths. It follows that if the latter are scrutable from the former, then social truths are scrutable from *PQTI*.

9 Deferential terms

What about cases in which a speaker uses an expression with deference to one's linguistic community, intending to use it to mean what one's community means with the term? Tyler Burge discusses the case of Bert, who thinks he has arthritis in his thigh, having no idea that arthritis is a disease found only in the joints. Bert has at best a partial understanding of the term 'arthritis'. But because he uses the term with deference to his linguistic community, his utterance of 'I have arthritis in my thigh' is false, and if he were to say 'Arthritis is a disease of the joints', he would say something true. One might think that these cases pose problems for scrutability. After all, Bert does not even know what 'arthritis' means, so it is not obvious how he could come to infer these truths from *PQTI*.

I will say that an expression is used *deferentially* by a speaker when the referent of the speaker's use of the expression depends on how others in the linguistic community use the expression.⁹ In the case above, Bert uses 'arthritis' deferentially: the referent of his use depends on others' use. It is certainly possible to use an expression nondeferentially: one can coin an entirely new term (e.g. 'glub'), deliberately use an existing term with a new meaning (e.g. stipulate that 'horse' will pick out the number two), or use a term with its correct meaning but insensitively to the use of others (e.g. stipulate that 'bachelor' picks out unmarried men, regardless of how others use the term). There are also intermediate cases in which a subject has a full understanding of a term (e.g. 'bachelor') without deference, and cases where there is both full understanding and deference. One can test for deference in these cases by asking: if it were to turn out that others' use the term 'bachelor' for something other than unmarried males, would the speaker's utterance of 'Bachelors are unmarried men' be true or false? It is plausible that in ordinary language use, many expressions are used deferentially and many are used nondeferentially. Whether an expression is used deferentially or nondeferentially plausibly depends on the intentions and/or dispositions of the speaker, but we need not precisely characterize these intentions and dispositions here.

I think that deferential usage poses no problem for scrutability here, as long as one recog-

⁹In chapter 3 I stipulated that ordinary truths exclude deferential utterances; that restriction is now being relaxed. What about relaxing other stipulations here: that the truths in question must be uttered by the speaker and that the speaker must have normal capacities? On the former: we need in principle to cover any context-independent sentence *M*, whether or not the speaker utters it or even possesses the relevant concepts. On the latter: we need in principle to cover any subject who is capable of thought. In both cases we can invoke an idealization to make the case that *if* the subject were to entertain a thought expressible by $PQTI \rightarrow M$, the arguments in chapter 3 and 4 suggest that they would be in a position to justify this thought through idealized reasoning, yielding a priori knowledge.

nizes that scrutability must proceed through knowing metalinguistic truths about one's community. Once Bert knows what others in his community (and especially the experts in the community) refer to with their uses of the term 'arthritis', then he will be in a position to know that 'Arthritis is a disease of the joints' is true. That is: 'Arthritis is a disease of the joints' will be inferentially and conditionally scrutable for Bert from a base that includes "'Arthritis' is used by others in my linguistic community to refer to a disease of the joints". Likewise, 'I do not have arthritis in my thigh' will be inferentially and conditionally scrutable from a base that includes this metalinguistic truth along with underlying truths such as *PQTI*. Then the arguments in the last chapter suggest that there will also be a priori scrutability from this base. Assuming that there are no problem with the scrutability of the relevant metalinguistic truth from underlying truths, then the problem is dissolved.

More generally: let us say that a deferential truth is a true utterance involving an expression used deferentially. Then it is highly plausible that deferential truths are scrutable from nondeferential truths, where the latter class includes nondeferential truths about the use of language. If so, then any problem for scrutability distinctive to deferential truths is removed.¹⁰

This picture requires that the apriority of a sentence in a context can depend on whether a given expression is used deferentially or nondeferentially in that context. If Susan fully grasps the concept of *bachelor* and expresses it with a nondeferential use of the term 'bachelor', 'All bachelors are male' may be a priori for her. If Fred only partially grasps the concept *bachelor* and expresses it with a deferential use of the term 'bachelor' (perhaps he has picked up the word from conversation but has no idea that bachelors must be male), 'All bachelors are male' will not be a priori for him. In the latter case, Fred will not be in a position to know the sentence from the armchair at all: at best, he can come to know it by further exposure to their linguistic community. And he certainly will certainly not be in a position to come to know the sentence a priori.

As usual, the difference in the apriority of the utterances mirrors epistemological differences in the thoughts that the utterances expressed: Susan's utterance expresses a thought that constitutes potential a priori knowledge, while Fred's utterance does not. One can find a reverse pattern for

¹⁰Block (2006) discusses a case in which a speaker has acquired *two* terms 'chat' on different occasions and has identical beliefs associated with both, but in which both have different referents due to deference to different speakers or communities. It is not really clear that this case is possible, but if it is, one can handle it as above, with the proviso that scrutability of truths involving a particular token of 'chat' will require tracing the causal history of that token, or of the corresponding concept in thought. Cases in which both concepts are used in language or thought simultaneously can be handled as in the discussion of indexicals and demonstratives below, by adjoining indexicals that pick out relevant concepts, thoughts, or experiences.

other utterances. For Fred, a sentence such as ‘Bachelors are what others in my community refer to as “bachelors”’ may be a priori. For Susan, the same sentence may be a posteriori, as knowledge of the sentence depends on empirical knowledge that others use the term the same way. Again, this is just what we should expect when apriority of an utterance is tied to what the speaker is in a position to know a priori.¹¹

10 Names

What about names? Here Kripke’s epistemological arguments against descriptivism suggest a challenge. Kripke makes the case that someone could use a term such as ‘Feynman’ or ‘Gödel’ while being ignorant of or mistaken about any properties of the referent. Given this, it may seem hard to see how truths involving these names will be scrutable.

One can respond to this challenge in exactly the same way as in the case of deferential terms above. In the relevant cases, once the subject knows enough about what *others* refer to with their use of terms such as ‘Feynman’ and ‘Gödel’, perhaps along with truths about the causal history of their own use of the relevant names, then they will have no trouble identifying the referent of the term as they use it, and coming to know relevant truths.

As in Chapter 1, nothing here requires that names are semantically equivalent to descriptions. The scrutability here, as usual, need not be grounded in a description, and any associated descriptive content may vary between different users of a term. Perhaps in some extreme cases of deference, there will be an associated description. If Fred hears the term ‘Gödel’ for the first time and immediately asks ‘Who is Gödel?’, it is perhaps not implausible that for Fred, something like the following is a priori: ‘If Gödel exists, Gödel is the referent of the term ‘Gödel’ as used by the speaker from whom I acquired the term’. But other cases may have a more complex mix of deference and other information, so that no clean description like this is available. Either way, there is no problem for scrutability, and for the apriori entailment of sentences involving these names by *PQTI*.

¹¹The trickiest cases are those in which a speaker fully grasps a concept such as *bachelor* but nevertheless expresses it with a deferential use of ‘bachelor’. On the current model, these are best regarded as cases in which ‘All bachelors are male’ is not a priori, as certainty about the sentence will require certainty about usage elsewhere in the linguistic community. Still, there is certainly something a priori in the vicinity for the speaker, and one can develop models on which this sentence is associated with both deferential contents and with nondeferential contents. For present purposes we need not resolve the issue, as there will be no problem for scrutability either way, but I discuss the issue further in *The Multiplicity of Meaning*.

Some object that the relevant sentences involving names cannot be a priori, as they express singular propositions, and these singular propositions cannot be known a priori. For example, in the case of Fred above, the allegedly a priori sentence will express the singular proposition that a certain specific person is the referent of someone's use of a term, and one cannot know such propositions about concrete individuals a priori.

In response: Whether or not these sentences express singular propositions, on the current framework they express thoughts. These thoughts are clearly thoughts that Fred is in a position to know on uttering the sentence, no matter what his relationship is to the individual in question. These thoughts do not require empirical justification: Fred can suspend judgment about the external world, and these thoughts will still constitute knowledge. And it is very difficult to see what the empirical evidence justifying these sentences might be. Someone might suggest, as in the objection from acquaintance in Chapter 3, that one's causal acquaintance with the person in question plays an evidential role, but one can respond here as I responded to the objection there. So I think there is not much reason to deny that at least as I have defined apriority, these sentences are a priori. (Of course it remains possible that in some other sense of 'a priori', such as a sense tied to apriority of a corresponding singular proposition, these sentences are not a priori.) I discuss this issue and some related Russellian objections about the a priori scrutability of names at greater length in the eleventh excursus.

11 Metalinguistic truths

On a number of occasions above, I have appealed to knowledge of metalinguistic truths, and in particular to truths about the referents of expressions used by others, such as "'Arthritis' as used by *X* refers to *Y*", "'Feynman' as used on occasion *O* refers to *Z*", and so on. Is it clear that these truths are scrutable?

In response, I think it is extremely plausible that given (i) the sort of information delivered by a Cosmoscope and (ii) enough intentional truths about speakers in our community, we are in a position to know truths about reference. It is arguable that the first alone is enough. But if not, the first and the second are enough. The details will depend on one's view of the relation between language and thought, though.

Certainly it is plausible that knowing the intentional content of mental states, in combination with knowing base truths about the world, puts one in a position to know the extension of any concepts involved in those mental states. This is straightforward on a view where content is Rus-

sellian (here the referents are built in), and it is also straightforward on a view where contents are primary intensions (which in effect yield a function from base truths to extensions). If the content of an utterance always reflects the content of an associated thought, then the same will apply to utterances, and the problem is solved. If the content of an utterance is a more complex function of the contents of associated thoughts, associated intentions, and so on, then then it will plausibly be scrutable from these contents. Of course sometimes the content of an utterance depends on a surrounding community. But this plausibly holds only in cases of semantic deference, and in these cases, the primary intension of the utterance plausibly mirrors that of a corresponding thought (both involve functions that picks out what relevant others in the community refer to with ‘arthritis’, and so on). So the same story applies. But in any case there is little reason to deny that intentional content of mental states across a community, combined with other qualitative information, will yield scrutability of truths about reference.

In the cases of deference and names above, one also needs to know truths about which others in the community are using the same word as one, or perhaps truths about which others one acquired a word from. Again, it is highly plausible that one could come to know these truths with the aid of a Cosmoscope to examine patterns of linguistic usage and causal connections between speakers, perhaps supplemented by facts about speakers’ intentional states.

What about truths about truth? Given that truths about reference are scrutable, it is plausible that truths about truth are scrutable too, or at least that they are scrutable if all other truths are. For example, if in an utterance of ‘Bill is Fred’ one can know that ‘Bill’ refers to the D_1 and ‘Fred’ refers to the D_2 , then as long as one can know that the D_1 is the D_2 , there is no problem with knowing that the utterance is true. One can also proceed via the primary intension of the utterance (which as above will be determined by the primary intensions of corresponding thoughts, or perhaps thoughts, intentions, and so on), which then in conjunction with base truths will straightforwardly yield a truth-value. So as long as we have intentional truths in the scrutability base, and as long as there is no other problem with scrutability in general, then there is no problem with the scrutability of these metalinguistic truths.

All this applies to the truth and reference of expression tokens. What about expression types? This case must be restricted to cases where the expressions in question are not context-dependent. In these cases, scrutability is straightforward. A metalinguistic truth such as “‘Philosophers exist’ is true” is clearly scrutable from the non-metalinguistic truth ‘Philosophers exist’ and the Tarskian truth “‘Philosophers exist’ is true iff philosophers exist”. So as long as the Tarskian truth is scrutable from non-metalinguistic truths, metalinguistic truths will be scrutable from non-

metalinguistic truths.

What of the Tarskian truth, then? Such truths are often held to be a priori, in which case there is no problem. I think the matter is a little more complicated. A quotational expression such as “‘Philosophers exist’” can be understood as picking out an *orthographically* individuated item, in which case it is not a priori that the item means anything at all. Or it can be understood as picking out an *semantically* (and orthographically) individuated item, in which case it is a priori that it means what it does. Understood the latter way, the Tarskian truth is a priori. Understood the former way, the Tarskian truth is not a priori.¹² But it is nevertheless scrutable: the meaning of the orthographic item will be scrutable from intentional states of users of the item, much as with the case of tokens above. So one has scrutability either way.

12 Indexicals and Demonstratives

We have seen already that indexical truths such as ‘I am Australian’ and ‘It is now 3am’ are not scrutable from non-indexical truths. To handle these, we added two indexical truths ‘I am D_1 ’ and ‘Now is D_2 ’ to the base. Are there any others that we need to add?

One might add ‘here’, but it is arguably equivalent to something like ‘The location where I am now’. A complication arises if time travel is possible: then the same person might be in two places now, but intuitively ‘here’ refers to just one of those places. To handle this, one could add ‘here’, or perhaps better, have a single indexical that picks out a single local time-slice (or perhaps a single total conscious state) of a person. Then ‘I’, ‘now’, and ‘here’ will all be determined by that time-slice in the obvious way.

What about demonstratives such as ‘that’? It is plausible that demonstrative truths such as ‘That is a cup’ are inferentially and conditionally scrutable given sufficient information about one’s experiences and their external causes. So it is plausible that these truths are a priori scrutable as well.

A complication arises due to Austin’s “Two Tubes” puzzle and related cases. I have distinct tubes attached to each eye, and see a dot through each. I experience a symmetrical visual field with

¹²This is particularly clear if one thinks about the corresponding thought. Suppose that ‘philosopher’ and ‘wise person’ are synonyms. Then “‘Philosophers exist’ is true iff philosophers exist” and “‘Philosophers exist’ is true iff wise people exist” will express the same thought. At the level of thought, any linguistic trapping for the concept expressed by ‘philosophers’ and ‘wise people’ is irrelevant. So it is clear that this thought about an orthographic item cannot be known a priori, though it can be known easily by someone who knows the language.

two red dots. I might wonder whether the two dots are the same object, and I put this by asking ‘Is $that_1$ the same as $that_2$ ’? I might also wonder whether the two dots are really red. I might hazard two conjectures by saying simultaneously ‘That is red’ (ostending the first) and ‘That is red’ (ostending the second). The first utterance (S_1) is true but the second utterance (S_2) is false.

In this case, it is plausible that S_1 is not scrutable from *PQTI*. From *PQTI* one can determine that the subject at the center is seeing two dots, one of which is red and the other of which is not. However, one has no basis to tell which of the two dots is *this* one, for the demonstratives above. The situation is entirely symmetrical between the two: for example, one can’t appeal to facts about the dot that one is now seeing, or the dot that one is now attending to, as one is seeing and attending to both. One likewise cannot appeal to ‘The dot on the left’, and so on, at least in the extreme case in which one has a symmetrical history and one has no basis as picking out one side as left or right (one can also imagine a case where the dots do not seem spatially related in this way).

I take the moral to be that as with ‘I’ and ‘now’, one needs to add some demonstratives to the base to handle these cases. Which demonstratives? It is plausible that certain *experiential demonstratives*, picking out experiences (or instances of phenomenal properties) will do. In the case above, I will have available two experiential demonstratives ‘ $This_1$ experience’ and ‘ $This_2$ experience’. ‘ $This_1$ experience is caused by a red dot’ is true, while ‘ $This_2$ experience is caused by a red dot’ is false. Given this information and other contextual information, I will have no trouble determining that S_1 is true and S_2 is false. So far each such demonstrative, one can build into the base a truth of the form ‘ $This_n$ experience is D_n ’, where as with the base truths for ‘I’ and ‘now’, D_n is a maximally specific description in the language of *PQTI* that the experience satisfies. This will then enable inferential, conditional, and a priori scrutability.

One might wonder how a connection between two demonstratives could be a priori. But this is not uncommon. It is possible to use a demonstrative ‘that’ to pick out an object, and to use a demonstrative ‘there’ to pick out wherever that object is located (whether or not it is where it seems to be). In such a case, ‘That is there (if it exists and is located) is a priori. Likewise, one can use an experiential demonstrative ‘ $That_1$ experience’ to pick out an object, and an perceptual demonstrative ‘ $That_1$ object’ to pick out whatever object is perceived with that experience. In that case, ‘ $That_1$ experience is an experience of $That_1$ object (if both exist)’ is a priori. A priori links of this sort between demonstratives suffice to ground a priori scrutability.

Of course the experiential demonstratives such as ‘ $That_1$ experience’ need not be uttered or even thought. One can think of each such demonstrative as a unique demonstrative in our ideal

language, tied to a specific experience: for each experience E , there is a corresponding demonstrative ' $This_E$ '. The scrutability base for a subject and a time need only involve demonstratives for each of the experiences the subject is now undergoing. And usually, few if any of these demonstratives will be required. A sentence S involves no demonstratives, one will not need any of these E -sentences for it to be scrutable. And if a sentence involves only one or two demonstratives, then typically one will need only one or two such E -sentences. But arbitrary such E -sentences are available if necessary.

Having such demonstratives available helps with the scrutability of orienting expressions such as 'left' and 'right'. As we have seen, in some symmetrical cases truths about these may not be inferentially or conditionally scrutable from $PQTI$. But if one has experiential demonstratives available, the symmetry is broken, and there is no problem with scrutability: one can determine what is on one's left by determining what is connected to a certain marked experience in one's visual field, or more better, by determining what typically causes experiences that bear a certain psychological and phenomenological relation to that experience. If so, there is no need to build in orienting expressions such as 'left' and 'right' in addition.

It is arguable that a similar phenomenon can arise for thoughts as well as experiences. It is not obvious how the referent of expressions such as 'This thought' will be scrutable in general. Here one can construct relevant puzzles cases with multiple thoughts of this form at a single time. How one handles this depends on how one understands the relation between occurrent thoughts and experiences. If occurrent thoughts necessarily involve experience (presumably a phenomenology of thinking), then demonstrative truths about the thought will plausibly be scrutable from demonstrative truths about the corresponding experience. But if they do not, then one may need to build in demonstratives for thoughts into the scrutability base. The same goes for demonstratives for other occurrent mental acts: 'That urge', 'This remembering', and so on. If these acts, or even one's thinking of these acts, necessarily involves experience, then the referents here will plausibly be scrutable, but if not, one may need extra demonstratives for occurrent mental acts in the base. My own tentative view is that occurrent thought always involves experience, so I do not think that the base needs to be expanded with further demonstratives. But the matter is far from obvious, and adding these further demonstratives is not out of the question.

¹²Or at least, for each atomic experience, where atomicity is spelled out appropriately: the thought is that any demonstrative truths about composite experiences will be scrutable from truths about atomic experiences along with Q .

13 Vagueness

I have already discussed an objection from vagueness in chapter 2. There the objection was that if S is true iff S is scrutable from $PQTI$, then given that $S \vee \neg S$ is true for all truth-apt S , it seems to follow that for all such S , either S is scrutable from $PQTI$ or $\neg S$ is scrutable from $PQTI$. But in the case of borderline cases of vague expressions, this is perhaps implausible. The solution adopted there was to restrict the scrutability principle to *determinate* truths: S is scrutable iff S is determinately true. Then as long as borderline cases are indeterminate, the problem is removed.

Still, there are other issues concerning borderline cases. One potential objection is that the scrutability thesis requires that whenever S is indeterminate, it is scrutable that S is indeterminate, which might seem far from obvious. In response, however, we need only note that the thesis (at least as formulated in the first excursus) requires only that the indeterminacy of S is scrutable when it is *determinate* that S is indeterminate, and this seems quite plausible.

The biggest objection stems from the epistemic theory of vagueness, according to which vague sentences are true or false even in borderline cases. On this view, in any utterance of ‘John is tall’, ‘tall’ expresses a precise property, with a sharp cutoff between cases with the property and cases without it. The utterance will always be true or false, with no intermediate status even in borderline cases. It is just that in borderline cases, we are unable to know the truth-value of the utterance even though we know the subject’s height and other relevant facts, because we are unable to know where the sharp cutoff falls. If so, then a sentence such as ‘John is tall’ can be true even though we cannot know it. And on this view, adding the information in $PQTI$ will not remove the ignorance. If so, then a sentence such as ‘John is tall’ may be true and inscrutable.

One might try to respond by appealing to the thesis that a truth is scrutable iff it is determinately true, and by giving an epistemological reading of determinacy, for example where determinacy comes to unknowability. If we do this, then this case is not a counterexample. But this epistemological reading of determinacy differs greatly from the original notion, and it tends to trivialize the scrutability thesis, so that the resulting thesis is significantly less interesting. One might also respond by appealing to idealization, and suggesting that even if we cannot know where the sharp cutoff falls, an idealized version of ourselves could. It is true that the epistemic theory itself does not make any claims about what ideal reasoners could know.¹³ Still, it would not be contrary to the spirit of the view for a proponent of the theory to deny that the cutoff is knowable even by ideal reasoning. So henceforth I will stipulate this understanding of the epistemic theory.

¹³Williamson 1993 is explicitly agnostic about ideal reasoners.

I think one should accept that *if* this version of the epistemic theory is true, then the scrutability thesis is false. Many truths involving borderline cases of vague expressions will be true but inscrutable from *PQTI*. One might suggest expanding the scrutability base to include the truths in question, or truths about the cutoff points for vague expressions ('Someone is tall iff their height is greater than ...'). But almost all terms in natural language are vague, and any scrutability base that includes all vague terms will certainly not be compact. And it is not clear how one could get around adding an enormous number of such terms: adding truths about the cutoff for 'tall' would seem to leave the cutoff for 'bald' no easier to know than it was before, for example. So it appears that a base vocabulary will have to be enormous. If so, the scrutability thesis is false.

A proponent of the scrutability thesis should instead respond by denying the epistemic theory of vagueness. This theory is widely regarded as extremely counterintuitive, so one is certainly not biting a large bullet by denying it. In fact, one might suggest that the implausibility of the epistemic theory is tied in some fashion to the way it denies scrutability, for example in holding that there are truths about someone's tallness that cannot be known even when one knows their exact height and other relevant qualitative truths. Furthermore: if scrutability holds in all other cases, then this fact can be used to argue *against* the epistemic theory, by establishing that the epistemic theorist must deny a principle that holds everywhere else. So if the epistemic theory is the main potential threat to scrutability, the dialectical situation certainly favors the scrutability thesis.

I will not try to argue further against the epistemic theory at this point, but in Chapter 10, I will return to considerations in favor of the scrutability thesis that also yield arguments against the epistemic theory. For now, this version of the epistemic theory plays the very useful role of providing a view according to which the scrutability thesis (whether in inferential, conditional, or a priori form) is clearly false.

14 Secondary qualities

What about truths about secondary qualities, such as the colors of external objects? It is fairly plausible that any truths here are scrutable from truths about the sorts of experiences that these objects cause, along with truths about their physical properties. For example, as long as *PQTI* can tell us that a given apple typically causes a certain sort of experience as of red, and that it has a physical property that typically causes experiences of that sort. then one is plausibly in a position to know that the apple is red. This will be so on a wide range of views about color: physicalist

views, dispositionalist views, phenomenalist views, response-dependent views, and so on.

There is one or two views of color on which this scrutability is not as straightforward. On some primitivist views of colors (discussed at greater length in the next chapter), colors are primitive properties that are quite distinct from any physical or dispositional and that are only weakly connected to experience. On some other views, we have at least primitive *concepts* of color, although these may pick out physical or dispositional properties. On some of these views (as discussed at greater length in the next chapter), one cannot rule out a priori a skeptical scenario on which everything looks to one just as it does, but in which the objects that typically look red have been green all along. On such a view, even all the information in *PQTI* may not rule out such a scenario. So if there are truths about objects' colors, they will be inscrutable from *PQTI*, or at least they will not be conclusively scrutable.¹⁴

I think it is implausible that there are inscrutable truths of this sort. On my own view (Chalmers 2006), we have primitive color concepts, but these pick out primitive properties that are not instantiated by objects in our world. If so, there are no inscrutable truths about their instantiation. Still, if someone holds that there are inscrutable truths about the instantiation of color, then they can include these truths in the scrutability base. One will presumably have to add corresponding truths for all other secondary qualities, and perhaps for all secondary qualities that might be attributed in some form of perceptual experience. If so, the base will undergo an unattractive expansion, one that arguably reflects an independently unattractive aspect of these views of color. In any case, all these truths fall under the family of truths about secondary qualities, and there seems to be no associated danger of trivialization, so there is no threat to the scrutability thesis here.

15 Macrophysical truths¹⁵

As defined, *PQTI* contains macrophysical truths: truth about nonfundamental objects in the language of classical physics, including expressions for spatiotemporal properties, mass, and so on. It is natural to ask whether macrophysical truths are dispensable from the base, and in particu-

¹⁴Closely related is the view endorsed by Tye (2006) and Byrne and Hilbert (2007), on which there are unknowable truths concerning which objects instantiate unique blue, in light of the fact that the conditions eliciting experiences as of unique blue vary between subjects.

¹⁵Thanks to Kelvin McQueen for detailed discussion of all the issues in this section. His Ph.D. thesis addresses the scrutability of macrophysical truths in various scientific domains, including the scrutability of macroscopic mass and the scrutability of macrophysical properties under various interpretations of quantum mechanics, in much more depth than I have here.

lar whether they are scrutable from a smaller class *PQTI*— which includes microphysical but not macrophysical truths. This issue could be discussed in the next chapter on narrowing the base, but in some respects it is closer to the spirit of this chapter, and it will be convenient to discuss it here.

In considering this question there are really two sets of issues: those that already arise for classical physics and those that arise distinctively for nonclassical physics, especially quantum mechanics. I will consider classical physics first. For this purpose I will assume a conception of microphysics on which fundamental objects are characterized by their relative location in a Newtonian spacetime, along with their velocity, mass, and so on. (One could also assume a Bohmian interpretation of quantum mechanics.) I will then relax this assumption in order to consider issues tied to nonclassical physics.

Given a classical conception of microphysics, there is an obvious route to the scrutability of macrophysical truths. The location of a macroscopic object is just the spatiotemporal region consisting of the location of its fundamental parts. The mass of a macroscopic object is just the total mass of its fundamental parts. So to determine the location and mass of a macroscopic object, it suffices to know the location and mass of its fundamental parts. Something similar plausibly applies to velocity and the like. If so, macrophysical truths are scrutable from (classical) microphysical truths.

One can also make the point by appeal to the Cosmoscope. Suppose that *S* is a sentence about the location, shape, and mass of a macroscopic object. Armed with a Cosmoscope that contained only microphysical and not macrophysical information, one could straightforwardly come to know whether *S* is true. One simply needs to determine whether there is an appropriately located and shaped spatial region at the relevant time, occupied by particles with the appropriate total mass. If the macrophysical truth includes mass density distributions, one can come to know this in a similar way.

There are a few residual issues. First, there are questions about just what count as objects. These questions are in effect addressed in the discussion of ontology earlier. Whether we assume a liberal or a restricted view of objects, it is plausible that truths about the existence of macrophysical objects will follow from microphysical truths along with certain principles of composition for macroscopic objects. On my own view, these principles will themselves either be a priori or scrutable from *PQTI*—. If so, there is no problem here. If not, then certain principles of composition may need to be built into the base. With the aid of those principles, there will be no problem in determining just where there are objects.

Second, one could plausibly argue that an object's macroscopic shape is more complex than

just the region occupied by its parts (perhaps it includes internal regions of vacuum, for example). Still, however shape is understood, shape should be scrutable from a Cosmoscope.

One might argue that it is not a priori that the mass of a macroscopic object is the sum of the mass of its fundamental parts, or that its location is the region of locations of its fundamental parts. Perhaps one could imagine world with microphysics like ours, but such that an object O , with mass three grams, is composed of two particles with mass one gram each? Or perhaps one could imagine that O is located in Australia, although its parts are located in the US? I think that these scenarios are at least difficult to imagine. Just perhaps, there could be certain micro-macro laws that resulted in odd emergent phenomena, so that the mass of a whole differs from the sum of the mass of its parts. But it is plausible that these laws would require further fundamental laws of nature (perhaps fundamental micro-macro laws), over and above the laws of microphysics that are shared with our (putatively classical) world. And such laws will themselves be ruled out by T . So $PQTI-$ rules out such a hypothesis. One might think that there may at least be questions about how macroscopic objects affect mass-measuring instruments that are left often by P alone. But as with other measuring instruments, once Q is included, then even these facts are settled.

One occasionally hears of some domain in science where macrophysical phenomena are held to be “emergent” from lower-level phenomena in the sense that they are unpredictable. However, such cases typically do not involve violations of scrutability. The paradigmatic cases of emergence in complex systems are cases in which macroscopic phenomena are predictable in principle from a full specification of microscopic phenomena. It is just that they are hard to predict in practice, or that they are hard to predict from microscopic laws alone without microscopic boundary conditions. In some extreme cases (Nielsen et al), it has been held that macroscopic phenomena may be uncomputable from an infinitary microscopic base, but as in the case of mathematics, this is no obstacle to scrutability. Others are cases where macroscopic *laws* do not seem to follow from microphysical laws, but once one enhances the base with particular microphysical facts the problem goes away.

If one had truly “strong” emergence, of the sort that Broad (1925) believed applied to chemistry and biology, then perhaps there would be a danger for scrutability. But again, if there is actually no strong emergence of this sort, then it will be ruled out by $PQTI-$ (and especially by T). If, contrary to current belief, there are in fact such laws of strong emergence, on the other hand, then these will be fundamental laws and will simply have to be built into the base, in a place analogous to the fundamental laws of nature in P .

Finally, it should be noted that everything here applies regardless of whether spacetime is New-

tonian or relativistic. The relativistic location and mass of macroscopic objects remains derivable from the relativistic location and mass of their fundamental parts. And truths about the classical location and mass of macroscopic objects, insofar as there are such truths, are themselves derivable from truths about relativistic location and mass, perhaps along with truths about one's own location and velocity, to determine a reference frame.

Next: issues tied to quantum mechanics. These issues are certainly more complex. For a start, much depends on which interpretation of quantum mechanics is accurate. The three major options include a Bohmian interpretation on which particles have classical properties; a collapse interpretation on which wavefunctions evolve in a superposed way and then occasionally collapse into a more definite state; and an Everett (or many-worlds) interpretation on which wavefunctions evolve in a superposed way and then never collapse. The fundamental microphysical truths in P will look quite different depending on which of these interpretations is correct.

On a Bohmian interpretation there is no problem for scrutability of macrophysical truths: things go through as for the classical microphysics described above. On an Everett interpretation, though, things are complicated, as it is not clear that there *are* any truths about the locations and masses of macroscopic objects on this interpretation. And even on a collapse interpretation, it is not obvious how to move from a wavefunction state to classical macrophysical truths.

Of course the proper treatment of this issue depends on whether there are (positive) macrophysical truths on the relevant interpretation of quantum mechanics, and if so what they are. If there are no such truths, then they cannot pose a problem for scrutability. Perhaps their absence could then pose a problem for the scrutability of other macroscopic truths from $PQTI$ -. But it is then far from clear what the problematic macroscopic truths might be, such that they are determinately true where ordinary macrophysical sentences are not. Presumably if it is not determinately true that objects have certain locations, it is also not determinately true that Schrödinger's cat is alive. The status of most other macroscopic sentences will be similar.

In any case, for present purposes, I will assume versions of the Everett and collapse interpretations according to which there are some macrophysical truths that correspond roughly to apparent macrophysical truths. For example, for a given object such as a tree or a dog, there will be macrophysical truths with the character 'There is an object with such-and-such size, shape, mass, and location'. On the Everett interpretation, this might work by giving a special role to one's own current "branch" of the wavefunction in determining which of one's sentences are true. On a collapse interpretation, this might work by allowing that at least a collapsed wavefunction in which most of the amplitude is contained within a given region can determine macrophysical truths.

One can then ask: how far down do these macrophysical truths go? Where there are determinate macrophysical truths, are there also classical truths at the level of chemistry, or of atoms, or of subatomic particles? On the face of it, the interpretive strategies above apply even to relatively low levels. If they do, then we can get from classical low-level truths to classical high-level truths just as in the non-quantum case. The residual issue will then be that of getting from fundamental quantum-mechanical truths to classical low-level truths. In any case, we can concentrate on classical truths about location at one of the lowest levels at which there are such classical truths: perhaps a truth saying that that a particle (or a molecule or a cell) is in a certain location.

On the collapse interpretation, a natural interpretive strategy is to say that an entity is located in a certain region if a high enough fraction of the (squared) amplitude of its wavefunction is concentrated within that region.¹⁶ This will deliver classical truths at both the microscopic and macroscopic level. If the interpretive principle here were itself a priori, then classical truths at the fundamental level would themselves be scrutable from quantum-mechanical facts. It is implausible that the principle is a priori. But it remains open that the conditional from *PQTI*– to the principle is a priori. The key question is whether it is a priori that *if* we are in a quantum-mechanical world with collapse, the location of objects is determined in this way.

Something similar applies on the Everett interpretation, where the natural interpretive strategy gives a special role to one's own branch of the wavefunction. This will deliver classical truths at the macroscopic level, at which branches have undergone "decoherence" and can be analyzed separately, and likewise at more microscopic levels until one reaches a level at which decoherence gives out. Again, if the interpretive principle here were itself a priori, then classical truths at the macroscopic levels would themselves be scrutable from quantum-mechanical facts. It is implausible that the principle is a priori. But it remains open that the conditional from *PQTI*– to the principle is a priori. Here, the key question is whether it is a priori that *if* we are in an Everett world of the appropriate sort, then the location of objects is determined in this way.

In both cases, I think there is a good case for scrutability of the relevant principle. One natural way that this might go is via the view that in the next chapter I call spatial functionalism: roughly,

¹⁶Here I assume something like the GRW version of a collapse interpretation. On this interpretation, collapses happen randomly, and while the post-collapse wavefunction has most of its amplitude concentrated in a small region, there are still infinite low-amplitude tails extending throughout space; hence the appeal to a "sufficiently high fraction". On other collapse interpretations, such as those on which collapse happens upon measurement by a conscious observer, collapse may leave a more determinate state without tails. In this case one has the option of saying that an entity is in a region if all of the associated amplitude is concentrated in that region, instead of a high enough fraction (although if one takes that route, objects may not remain in regions for long).

that what it is to be a spatial property is to play the appropriate causal role. In the case of macroscopic spatial properties, it is plausible (as I argue there) that spatial properties can be picked out as that manifold of properties that serve as the causal basis for spatial experience, in much the way that color properties are picked out in the previous section 13. To simplify, the property of being two meters away from one might be picked out as the spatial relation that normally brings about the experience of being two meters away from one. Something similar goes for properties involve relative length, position, shape, and the like.

One can then argue that on a collapse interpretation, the properties and relations that normally bring about the relevant sort of spatial experiences are precisely properties and relations requiring the wavefunctions amplitude to be largely concentrated in a certain area. Likewise, on an Everett interpretation, the relevant spatial experiences will normally be brought about when there is an object in the same branch of the wavefunction with relevant properties. The details can be argued about, just as in the case of color. But the general point is that *PQTI-* will put one in a position to determine the causal and counterfactual bases of spatial experiences, and that doing that will put one in a position to determine principles connecting quantum-mechanical properties to macrophysical spatial properties of objects. Something similar goes for mass.

One might resist if one is a spatial primitivist (analogous to the color primitivists above, and discussed more in the next chapter), according to which we have primitive concepts of spatial properties, and do not pick them out via the role that they play. On this view, it is much harder to see how apparent macroscopic spatial properties will be scrutable from *PQTI-*. But if one is a spatial primitivist, it is hard to see how macroscopic objects have the spatial properties that they appear to have at all, if collapse and Everett interpretations are true. If those interpretations are true, I think the best thing for a spatial primitivist to say is that objects do not strictly have determinate spatial properties (the molecule is not determinately at a location), but instead have more complex properties best describable in quantum-mechanical terms. But then there is no challenge to scrutability, as the putative macrophysical truths will not be truths at all, and the residual truths will certainly be scrutable.

Of course, even a spatial primitivist might allow that if we discover that a collapse or Everett interpretation is true, we could come to use spatial vocabulary in a less strict way. We might say for ordinary purposes that an object has a certain location when enough of its amplitude is concentrated there, for example. This would parallel a color primitivists' attitude to our discourse on discovering that objects do not really have primitive colors. Spatial expressions in the new vocabulary might in effect work much as spatial functionalists hold that our existing spatial expressions

work, and any spatial truths in the new vocabulary will remain scrutable.

There will be a problem for scrutability only on a spatial primitivist view that holds that spatial claims in the primitive vocabulary express truths even if the Everett or collapse interpretations are correct. This view is analogous to the realist color primitivist discussed above. As with that view, it involves an uncomfortable combination of claims, and I think it should be rejected. (See also the discussion of solidity and semantic conservatism in the eighth excursus for some relevant observations.) But if it is accepted, then we need only expand the scrutability base either to include certain claims about macroscopic location, or perhaps better, to include general principles linking quantum-mechanical properties to spatial properties of macroscopic properties. This will require a more liberal version of *P*, but the vocabulary need not go beyond that of microphysical properties along with space, time, and mass.

16 Counterfactual truths

PQTI contains true counterfactuals in a microphysical, macrophysical, and phenomenal vocabulary. Still, it is attractive to hold that one can dispense with counterfactuals, on the grounds that they are scrutable from truths about laws of nature. This result along with that of the previous section would allow us to strip down *PQTI* to its cousin *PQTI-*, which dispenses with macrophysical and counterfactual truths.

To start with, one can make a plausible case that macrophysical counterfactuals are scrutable from microphysical counterfactuals. Here one can use the reasoning in the last section to make the case that the antecedents and consequents of these counterfactuals are scrutable from specifications of this world and others using a stripped-down vocabulary that does not use macrophysical notions. One can then reason as in chapter 3 (on the objection from counterfactuals) to make the case that the truth-value of these counterfactuals will be derivable from the truth-value of counterfactuals in the stripped-down language. The same goes for any counterfactuals connecting the macrophysical with phenomenology.

What about microphysical counterfactuals? It is natural to hold that these are scrutable from microphysical laws and that the latter are scrutable from psychophysical laws connecting microphysics to phenomenology. When a microphysical counterfactual has a maximally specific antecedent, true in exactly one nomologically possible world, then one need only apply the laws to the antecedent to determine the status of the consequent. When the antecedent is true in more than one nomologically possible world, we can still apply the laws to each world to determine

the status of the consequent in that world. Then reasoning about similarity between the relevant worlds and our own, as in section 9 of chapter 3, will render the original counterfactual scrutable.

Likewise, counterfactuals connecting microphysics and phenomenology will be derivable from psychophysical laws. We can assume, as is plausible, that phenomenal truths supervene at least nomologically on microphysical truths. Then there will be many nomologically necessary conditionals from microphysical sentences to phenomenal sentences. Assuming that there is some systematicity to these conditionals, there will be a smaller class of nomologically necessary psychophysical conditionals from which all of these conditionals follow. On a dualist view, all the conditionals will follow from fundamental psychophysical laws that have a status similar to fundamental laws of physics. On a materialist view, these conditionals will have a different status, but they remain nomologically necessary (even if they are also metaphysically necessary), and for present purposes we can count them as laws. Then as long as these laws (in a version saying that they are nomologically necessary) are incorporated into *PQTI*–, reasoning of the sort in the previous paragraph suggests that all psychophysical counterfactuals are scrutable from here.

17 Conclusion

We have seen that there is a reasonable case that all truths are scrutable from *PQTI* and from its stripped-down cousin *PQTI*–. In almost all of the hard cases, there are independently attractive reasons for embracing a view on which scrutability holds.

In a couple of cases, there is a relatively strong case for expansion even on my own views. These include the domain of ontology, where there is a case for including any further truths about fundamental natural ontology (such as truths about quiddities, perhaps), and the domain of indexical truths, where there are good reasons to include certain truths involving phenomenal demonstratives. But these expansions are quite compatible with the spirit of the scrutability thesis.

Other significant challenges come from mathematical truths, ontological truths, intentional truths, truths involving vague expressions, and truths about secondary qualities. Handling the mathematical case requires a heavy idealization. In the other cases, there are important philosophical views on which expansion of the base is required. I reject these views, and I think that in each case the rejection can be independently motivated. Still, the views are worth noting.

For most of the hard cases, even if we take an expansionist line, the compact scrutability thesis will not be threatened. Adding mathematical, normative truths, ontological truths, secondary-quality truths, and macrophysical truths will leave the base compact: in fact, in only some of these

cases will any new vocabulary be required. Adding intentional truths raises tricky issues because of the threat of trivializing mechanisms, but we have seen that even if these truths are added, the threat can be avoided.

The most significant threat to the compact scrutability thesis in this chapter arises from the epistemic view of vagueness. If an appropriate version of it is true, there might be failures of scrutability for arbitrary expressions. Still, rejecting the epistemic view of vagueness cannot be counted as an implausible move. Some might add further challenges that threaten large classes of expressions, perhaps arising from names, or from expressions used deferentially, or from expressions expressing recognitional concepts. In each of these cases, however, we have seen that the threat can be naturally answered. I conclude that the compact scrutability thesis remains extremely plausible.

What of stronger scrutability theses such as Fundamental Scrutability, holding that all truths are scrutable from metaphysically fundamental truths and primitive indexical truths? The first two expansions above leave this thesis intact, as these expansions concern fundamental truths and primitive indexical truths respectively. If one were to make any of the later expansions, one might have to reject the thesis, at least if one does not take the relevant normative, ontological, or intentional truths (and so on) to be metaphysically fundamental. And of course there remains a question about whether the base must include phenomenal truths and whether these are metaphysically fundamental. Still, I think that where the cases we have considered so far are concerned, Fundamental Scrutability remains attractive and plausible. I return to this matter in chapter 8.

What of a stronger A Priori Scrutability thesis that invokes the conclusive a priori, in effect requiring that all truths be a priori scrutable from base truths with certainty? In the discussion in this chapter I have mainly considered ordinary a priori scrutability, not the conclusive variety. Still, in many of the hard cases, the arguments for scrutability can straightforwardly be extended to arguments for conclusive scrutability. One might worry about arguments from knowability to scrutability in the analysis of normative and ontological truths. Even if it is plausible that we can know these truths, it might be denied that we can know them with certainty (even given base facts and insulated ideal reasoning), in which case the arguments in question cannot easily be extended into arguments for conclusive scrutability. I think that there are good grounds for holding that certainty is possible in these domains (given base facts and insulated ideal reasoning), at least to the extent that there are truths in these domains at all. But making this case requires more than I have argued above. In any case, if we allow a somewhat expanded base including normative truths, ontological truths, and the like, the conclusive scrutability thesis will remain plausible.

Eleventh Excursus: The Fregean and the Russellian

In *Meaning and Necessity*, Carnap used his account of intensions and extensions to support a Fregean approach to meaning. In particular, he argued that his intensions could play the role of Fregean senses, for names and for other expressions. In later work on possible-world semantics, following Kripke, intensions and worlds have tended to be strongly separated from epistemological notions, and the intensions that result are far from Fregean in character. The scrutability framework allows us to reconnect intensions to epistemological notions, by defining intensions over epistemically possible scenarios. This allows them to once again play some of the roles of Fregean senses.

Developing a Fregean account of meaning and content is the main focus of *The Multiplicity of Meaning*, so I do not focus on these issues much in this book. But here I will provide a brief sketch of how A Priori Scrutability can serve as a foundation for this sort of theory. I will also address potential objections from Russellian opponents to the argument for A Priori Scrutability.

Russellian theories hold that the meaning of simple expressions such as ordinary proper names in natural language is exhausted by their referents. On this view, the meaning of the names ‘Hesperus’ and ‘Phosphorus’ is the same, the planet Venus. Fregean theories of meaning hold that the meaning of all such expressions involves a further element—sense—that is tied to cognitive significance: so because ‘Hesperus is Phosphorus’ is cognitively significant, ‘Hesperus’ and ‘Phosphorus’ have different senses.¹⁷

A Priori Scrutability bears on this debate in part by making the case that most expressions in natural language, at least as uttered by speakers, have substantive and nontrivial a priori connections to other expressions. In particular, they have substantive a priori connections to expressions in a compact base language (the language of *PQTI*, say). Substantive connections of this sort strongly suggest that there is a Fregean aspect of content that is reflected in these connections. In

¹⁷The terminology here is a little awkward in that Russell himself may come out as a Fregean, in that he held that ordinary proper names such as ‘Hesperus’ is not exhausted by their referents, as these names are really disguised descriptions. On Russell’s own view, the only truly simple expressions are names for sense-data and other entities with which we are directly acquainted. This contrasts with contemporary Russellian views that extend Russell’s view to all ordinary proper names. It has become standard to classify Russell’s view as a broadly Fregean view, with content of an associated description playing the role of sense, and I follow that practice here. That being said, there is room for a distinction among broadly Fregean views between “Russellian Fregean views” that put special weight on acquaintance and broadly descriptive content, and “non-Russellian Fregean views” that do not. In some respects, my own view (or at least the view that derives from Super-Rigid Scrutability) is akin to a Russellian Fregean view. I do not take a stand on where Frege’s own views fell.

fact, one can use these connections to define a sort of Fregean content.

In Chapter 1, we saw that A Priori Scrutability provides a natural response to Kripke's epistemological arguments against descriptivism. Kripke made the case that there is no nontrivial description D such that 'Gödel is the D ' is a priori. In response, a Fregean may suggest that the Fregean content of 'Gödel' is not captured by a description, but by an intension governing its application to epistemically possible scenarios. For example, the intension of 'Gödel proved the incompleteness of arithmetic' may be true at a scenario corresponding to the actual world, while being false at a scenario in which a professor named 'Gödel' stole the proof from someone named 'Schmidt'. As we have seen, Kripke's argument gains no purchase against such a view, because the argument itself turns on intuitions about truth-values and reference at specific scenarios, which can always be captured by an intension.

A Priori Scrutability, and its stronger cousin, Generalized A Priori Scrutability, vindicate this response by ensuring that these intensions are well-defined. First, these theses can be used to make the intuitive notion of a scenario precise, as we saw in the tenth excursus. With scenarios understood this way, we can say that as before that the intension of a sentence S (in a context) is true at a scenario w iff S is a priori scrutable from D (in that context), where D is a canonical specification of w . There are related definitions for falsity at a scenario, indeterminacy at a scenario, Generalized A Priori Scrutability ensures that these intensions will be well-defined, and that whenever a sentence S is epistemically possible (not ruled out a priori), there will be some scenario at which its intension is true.

For example, suppose I utter the true sentence 'Gödel proved the incompleteness of arithmetic'. A Priori Scrutability entails that this sentence token S is a priori scrutable from a canonical specification D of my actual scenario, where D is something along the lines of $PQTI$. So the intension of S is true at that scenario. The Gödel-Schmidt hypothesis is itself epistemically possible, and by Generalized A Priori Scrutability will be a priori scrutable from some complex sentence in $L-D'$, perhaps—which will be a canonical specification of some scenario. The truth-value of S will be a priori scrutable from D' : presumably, D' will imply the negation of S . So the intension of S will be false at this scenario.

These intensions behave in a highly Fregean fashion. Consider the intension of 'Hesperus is Phosporus', as uttered in my current context. Because this sentence S is true, its intension will be true at the my actual scenario, but given that S is not a priori, its intension will be false at some scenarios. For example, if D' specifies a scenario in which the relevant bright object in the morning sky differs from the relevant bright object in the evening sky, then 'Hesperus is not Phosporus' will

be scrutable from D' , so S will be false at that scenario. Intensions behave compositionally in the usual way, so that in a scenario (like the one above) where 'Hesperus is Phosphorus' is false, the intensions of 'Hesperus' and 'Phosphorus' will pick out different referents: the different bright objects in the morning and evening sky. So as used by me, 'Hesperus' and 'Phosphorus' will have different intensions, just as they have different Fregean senses.

This behavior is quite general: whenever ' $A = B$ ' is not a priori (in a context), the names A and B will have different intensions (in that context). The same goes for other expressions: for example, if two general terms are such that 'Everything is an A iff it is a B ' is not a priori, they will have different intensions. So intensions serve as Fregean semantic values for general terms too.

Intensions can also play many of the roles that descriptions play in descriptive accounts of meaning. Intensions will not in general correspond to descriptions, for reasons given in chapter 1, but they can at least be approximated by descriptions. For example, the intension of 'water', in a scenario, might very roughly pick out the same thing as 'The clear, drinkable liquid found around here' in that scenario. Of course there will be counterexamples, leading to ever-longer descriptions, but an approximate description can at least capture some of the relevant behavior. And like descriptions, intensions can be seen as capturing criteria that an entity in the environment must satisfy in order to qualify as the referent of a term.

If an opponent wants to resist the case for Fregean theories of meaning, they must do one of three things: (i) resist the case for the scrutability theses, (ii) resist the move from scrutability theses to the existence of intensions, or (iii) resist the claim that intensions so-defined have the key properties required for a successful Fregean theory of meaning. In what follows I will say something about all three strategies, in reverse order.

Regarding (iii): There are a number of potential objections to the claim that intensions so-defined have the properties required for a successful Fregean theory of meaning. One natural objection is that this account cannot cope with Kripke's modal argument: 'Hesperus is Phosphorus' is necessary, so 'Hesperus' and 'Phosphorus' must have the same intension, while for descriptions D such as those above, 'Hesperus is the D ' is not even close to necessary. One can deal with this objection by invoking two-dimensional semantics, on which names are associated with two intensions: a primary intension, governing application to epistemically possible scenarios, and a secondary intension, governing application to metaphysically possible worlds. The primary intensions of 'Hesperus' and 'Phosphorus' are distinct, mirroring the fact that 'Hesperus is Phosphorus' is not a priori, but their secondary intensions are the same (both pick out Venus in all worlds), mirroring the fact that 'Hesperus is Phosphorus' is necessary.

In a more refined account, one can construe the Fregean sense of a term such as ‘Hesperus’ as an *enriched intension*: an ordered pair of a primary intension and an extension. Then ‘Hesperus’ and ‘Phosphorus’ will have distinct enriched intensions that share an extension. Once we combine this with the thesis that modal operators (such as “Necessarily”) are sensitive only to the extension component of the enriched intensions of embedded expressions, whereas epistemic operators are sensitive in addition to primary intensions, this yields the desired result. This hypothesis also has the advantage of yielding a broadly Fregean treatment of indexical expressions such as ‘I’ and ‘now’, and a broadly Fregean treatment of attitude ascriptions (see my ‘Propositions and Attitude Ascriptions: A Fregean Account’ for more).

A second objection suggests that intensions are too coarse-grained to serve as Fregean senses, because of the idealization in the notion of apriority. For example, ‘ $77 \times 33 = 2541$ ’ is a priori, but it is still cognitively significant. So ‘ 77×33 ’ and ‘2541’ should have different Fregean senses, but they will have the same primary intensions. One can deal with this objection by holding that the Fregean sense of a complex expression should be seen as a *structured intension*, composed of the (primary or enriched) intensions of its parts, according to the expression’s logical form. So ‘ 77×33 ’ and ‘2541’ will have different structured intensions.

Almost all potential problem cases can be handled this way. The only potential problem cases will arise if there are pairs of *simple* expressions ‘A’ and ‘B’ such that ‘ $A=B$ ’ is a priori while also being cognitively significant. It is unclear that there are any such pairs, however. Someone might suggest that we can stipulate a simple name such as ‘N’ whose reference is fixed to be something complex such as 77×33 ; but in such a case it is at least arguable that ‘N’ should be understood to have complex structured content. If there are exceptions here, however, then we might try moving to a finer-grained space of scenarios and finer-grained intensions, as discussed at the end of the tenth excursus. Alternatively, we can simply allow that while intensions are very much like Fregean senses, they are not quite as fine-grained as Fregean senses. This will be consistent with the intensions playing all sorts of explanatory roles.¹⁸

A third objection says that intensions of this sort are unsuitable to serve as aspects of *meaning*, because meanings are properties of expression types in a public language, whereas intensions are context-dependent: the same expression can have different intensions in different contexts. In response: it is certainly true that the intension associated with an expression such as ‘Gödel’

¹⁸As I discuss in chapter 9, I am inclined toward a semantic pluralism on which there are many notions of meaning or content playing many different roles. The intensions discussed here are not all there is to meaning and content: they can play many of the relevant roles, but they do not play all the roles.

or ‘water’ or ‘tall’ or ‘that’ can vary between contexts. So *if* “meaning” is understood to be incompatible with this sort of variation, then intensions are not meanings. However, intensions can still play many of the explanatory roles that meanings are supposed to play—for example, they can serve as truth-conditions for utterances, as referents for ‘that’-clauses, and as arbiters of cognitive significance. If they play these roles, little rests on the verbal question of whether they are called “meanings” or whether we use some other term such as “content”.¹⁹

It is worth noting here that Frege himself allowed that the sense associated with an expression in natural language can vary between different speakers. Likewise, when Kripke argues against descriptive theories of meaning, he puts no weight on a criterion of constancy between speakers, instead allowing that descriptions might vary between speakers and casting his arguments in terms of what is “a priori for a speaker”.

Of course intensions should still be public at least in the sense that they can in principle be shared across speakers, even if they need not in practice be shared across speakers. This sort of publicity is ensured by having a common space of scenarios for all speakers. The common space is ensured in turn by the use of a context-independent scrutability base, as discussed shortly.

A fourth objection says that these intensions do not determine referents as senses should. One version of the objection is tied to indexicality: two speakers in the same world can have the same intension but a different referent, perhaps for ‘I’ and ‘now’. But although this rules out an absolute determination of referent by intension, it remains compatible with a weaker determination of reference by intension plus the speaker’s scenario of utterance, and this sort of determination is good enough for most Fregean purposes.

Another version of the objection notes that scenarios are linguistic constructions, so that intensions yield only word-word relations, not word-world relations. But it is easy to extend these to word-world relations. Most obviously, given that one scenario is actualized for any given speaker, the intension of a sentence at that scenario will determine its truth-value. Of course, the notion of actualization of a (linguistically constructed) scenario in effect assumes word-world relations for words in a scrutability base. But as we saw in the tenth excursus, one can also construct scenarios

¹⁹A related objection is that these intensions are speaker meanings rather than semantic meanings. Again, this depends on how one draws these divisions. If the division is drawn so that semantic meanings are associated with expression types while speaker meanings are associated with utterances, or so that semantic meaning of an expression cannot vary with context while speaker meanings can, then clearly intensions fall on the side of speaker meanings. On the other hand, if the division is drawn so that semantic meanings of an utterance are literal truth-conditions for that utterance while speaker meanings are not, then intensions fall on the side of semantic meanings. See “On Sense and Intension” for more discussion of this issue.

nonlinguistically, for example as centered worlds, where worlds are understood as complexes of properties and relations. For any given speaker, there will then be a nonlinguistic fact of the matter as to which of these scenarios is actualized for any given speaker, so that intensions will yield a word-world relations.

Even without such a construction, the scrutability framework in effect allows word-world relations for all expressions to be fixed by word-world relations for base expressions along with scrutability relations. A Fregean sympathetic with this picture might say that the base expressions have primitive senses, while other expressions have less primitive senses that derive from these. Given scrutability from a base of super-rigid and indexical expressions (as discussed in the tenth excursus and at more length in the thirteenth), we can construe these primitive senses as super-rigid senses and indexical senses, both of which pick out their referents in an especially direct fashion. If the Fregean is friendly to acquaintance, it might be held that these are special acquaintance-based senses. Even without an appeal to acquaintance or super-rigidity, then as long as we have the materials to account for reference in the scrutability base, scrutability then allows intensions to do the rest of the work.

Regarding (ii): the move from scrutability theses to defining intensions is fairly straightforward. Given Generalized A Priori Scrutability, one can construct scenarios from epistemically complete sentences in the compact base language and define intensions from there. An opponent could in principle resist the move from the scrutability thesis to the existence of epistemically complete sentences, but I do not think this is an especially promising place to resist. There is a good case that every epistemically possible sentence is a priori entailed by an epistemically complete sentence in an ideal language (see Chalmers 2011 for argument), and even if there were not, one could get much of the force of the intension framework out of intensions defined over ever-more-complete partial scenarios.

Another way to resist here might be to claim that intensions are sometimes ill-defined because of ineliminable context-dependence in base sentences. If a base sentence G exhibits a certain sort of context-dependence, it can happen that a token S is scrutable from G via some uses of G and not others. For example, if G is 'Fred is a medium-sized number', then a token of 'Fred is less than 100' may be a priori scrutable from G via some uses of 'medium-sized' and not others. Whether 'Fred is less than 100' is then scrutable from G depends on the fine details of how scrutability from sentence types is defined.²⁰ On the definitions used in chapter 2, all of which are designed for context-independent base sentences, one gets a priori scrutability in this case, and one gets a well-defined intension. But this just shifts the problem. Now it may turn out that a token of 'Fred

is greater than 100' is scrutable from G too (via a different usage of 'medium-sized' in G), so that intensions are very ill-behaved.

To avoid worries of this sort, we have stipulated that base sentences must be epistemically invariant, so that their a priori relations to sentence tokens do not depend on context. This requirement makes the scrutability thesis a stronger thesis: it plausibly excludes most names and natural-kind terms from the base, for example. In addition, the microphysical terms in $PQTI$ are arguably not epistemically invariant: one might competently use 'charge' under somewhat different guises with different a priori entailments. Fortunately, as discussed earlier, the differences do not seem to yield any cases where both S is a priori scrutable from $PQTI$ under some uses but not others: any local differences in the role of 'charge' wash out at the level of scrutability from the whole network in $PQTI$. But to eliminate epistemic variation altogether, we can simply stipulate a certain specific guise for microphysical terms, perhaps by requiring a certain theoretical definition of 'charge' and the like; or we can move straight to a Ramsey-sentence definition as in chapter 7. It is plausible that the minimal vocabularies canvassed in chapters 7 and 8 can all be understood in epistemically invariant ways (with a minor exception for phenomenal demonstratives that does not greatly affect the application to intensions²¹), so that intensions defined using them will be well-defined.

Regarding (i): the remaining strategy for a Russellian is to resist the arguments for the scrutability theses. While it might suffice to resist only Generalized A Priori Scrutability, in practice it is likely that a Russellian who resists that thesis will also resist A Priori Scrutability. So I will focus mainly on resistance to A Priori Scrutability, with occasional attention to Generalized Scrutability. I will also attend to Inferential and Conditional Scrutability, because of their roles in establishing A Priori Scrutability. Of course a Russellian might resist these via any of the objections I have considered in the last two chapters, but I will focus on forms of resistance that are might be

²⁰The definitions used in chapter 2 yield a priori scrutability but not conditional scrutability in this case, with the status of inferential scrutability being unclear. The main relevant difference in fine details is whether the definition appeals to properties of all thoughts apt to be expressed by G (as in the definition of conditional scrutability), some thought (as in the definition of a priori scrutability), or a nearby thought (as in the definition of inferential scrutability). When base sentences are all context-independent, this difference does not make much difference, but it makes a difference when context-dependent sentences are involved.

²¹For phenomenal demonstratives, here the key expressions will arguably be either epistemically variant ('that') or unshared across speakers ('that_E', where E is a token experience). This makes for a very limited lack of alignment between scenarios, but one that still supports a correspondence relation. See Chalmers 2002 and 2011 (OSI, NES) for discussion.

distinctively motivated by the Russellian view.

As far as I can tell, there is no distinctive reason why a Russellian should resist the Inferential Scrutability thesis. It appears perfectly consistent with a Russellian view, for example, to hold that if one knows *PQTI**, then one will be in a position to know the truths of one's utterances of sentences such as 'water is H₂O' and 'Gödel proved incompleteness', and there is no form of resistance here that is distinctively motivated by a Russellian view.

As for Conditional Scrutability: there is also no clear reason why a Russellian should resist this thesis, and or why they should resist instances of it such as the claim that *cr**('water is H₂O' | *PQTI*) is high. There is a general issue about how Russellians should understand credences, and about whether utterances of sentences such as 'Hesperus is Hesperus' and 'Hesperus is Phosphorus' should be associated with the same credence (because they express the same singular proposition) or different credences (because a subject might be much more confident of the first utterance than the second). If we took the first view, this might lead to doubts about some generalized scrutability claims, such as the claim that there exists some *PQTI** from which the first but not the second is conditionally scrutable. But given that we have defined the credence associated with a sentence token in terms of the credence associated with a corresponding mental state, the second view is the natural view to take. Even on a Russellian view, it seems clear that the first utterance may be associated with a much higher-credence mental state than the second. This claim is quite consistent with holding that both mental states involving grasping the same singular proposition. We need only hold that the two states involve grasping the same proposition in different ways, as most Russellians allow.

Now, some hard-line Russellians reject all talk of ways of grasping propositions, and correspondingly may reject the claim that 'Hesperus is Hesperus' and 'Hesperus is Phosphorus' express different mental states. But as discussed in Chapter 2, this leads to obvious problems. For example, there is an obvious intuitive sense in which a speaker might express knowledge by uttering the first sentence, while not expressing knowledge by uttering the second sentence. And there is an obvious intuitive sense in which a speaker might have higher credence associated with an utterance of the first sentence than with an utterance of the second. Any Russellian needs to accommodate these intuitions somehow. As discussed in chapter 2, the obvious way to do so is to make distinctions between the token mental states associated with the sentences, distinctions that are compatible with sameness of content. Then however the difference between unconditional credences associated with these utterances is accommodated, we can use the same method to understand the conditional credences used in stating Conditional Scrutability. Once this is done,

any obstacle to Conditional Scrutability will be removed. If a Russellian cannot provide a way to understand these credences, on the other hand, then this is a good reason to reject their view.

As for A Priori Scrutability: some of the same issues that come up with Conditional Scrutability also come up here. For example, it might be held that apriority of a sentence token depends on the singular proposition it expresses, so that tokens of ‘Hesperus is Hesperus (if it exists)’ and ‘Hesperus is Phosphorus (if it exists)’ will both be a priori, and there will be no *PQTI** from which one but not the other is a priori scrutable. But again, as discussed earlier in this chapter, we have defined the apriority of a sentence token in terms of the epistemic status of an associated mental state, and it is clear that there may be a difference in the epistemic status of the mental states associated with utterances of the two sentences above.

Perhaps the most likely form of Russellian resistance will be to accept Conditional but not A Priori Scrutability. It may be allowed, for example, that cr^* (‘water is wet’ | *PQTI*) is high or 1, but denied that ‘*PQTI* → water is wet’ is a priori. A Russellian might allow that the latter material conditional is knowable from the armchair, but hold that this knowledge always involves empirical justification. This parallels a familiar line among Russellians (e.g. Salmon 1986) who reject Kripke’s claims about contingent a priori knowledge: they hold that even if the term ‘one meter’ is introduced to stand for the length of stick *S*, ‘the length of stick *S* is one meter’ is knowable from the armchair but involves empirical justification. Something similar holds for cases involving descriptive names, such as ‘Newman 1 is the first child born in the 22nd century’, and ‘Julius is the inventor of the zip’, where the names were introduced using the associated descriptions as stipulations. It would be natural for a Russellian to take a similar line about related conditional sentences, such as ‘A businessman invented the zip → Julius is a businessman’. So it would be natural for such a Russellian to take a similar line on ‘*PQTI* → Julius is a businessman’, and ‘*PQTI* → water is wet’. According to this line, all these sentences are knowable from the armchair, but require empirical justification.

For simplicity, let us focus first on the case of ‘Julius’, introduced by stipulating that it is a name for the inventor of the zip, whoever that may be. The standard Russellian line is that ‘Julius invented the zip’ expresses a singular proposition about William C. Whitworth, who actually invented the zip, and that no-one could know this proposition a priori, as no-one could know a priori of Whitworth that he invented the zip.²² Instead, anyone who knows this proposition knows it a posteriori, justified in part by their direct or indirect acquaintance with Whitworth. Now, of course on the current understanding of apriority, we cannot simply equate the apriority of the sentence with the apriority of the singular proposition. Still, the claim that the knowledge expressed by

'Julius invented the zip (if anyone did)' is empirical is a more general one that deserves attention.

I have already discussed a version of this line under "the objection from acquaintance" above, but it is worth applying my response there to the case of descriptive names. As before, I think the argument from suspension of judgment gives good reason to hold that the relevant knowledge is not empirical. Even if one has the relevant sort of acquaintance with Whitworth, one can engage in Cartesian suspension of belief about the external world, including suspension of one's beliefs in the deliverances of perception. Under such suspension of belief, one cannot be justified in believing 'There are zips', 'Someone invented zips', 'Whitworth exists', or 'Julius exists'. But on the face of it, assuming that one uses 'Julius' in a way governed by the original stipulation, one is still perfectly justified in believing 'Julius invented the zip (if anyone did)', and indeed, this belief constitutes knowledge. One's justification for this claim cannot be constituted by perceptual evidence concerning zips, and indeed by any perceptual evidence, as the justificatory role of evidence of this sort is undermined by the Cartesian suspension of belief. And introspective justification does not seem to the point here (certainly the putative acquaintance-based justification was not supposed to be introspective). So *prima facie*, one's justification here is *a priori*.

One can flesh out the situation by comparing this case to one in which a term such as 'Newman 1' is used without any empirical acquaintance with the individual who is the referent. In this case, an utterance of 'Newman 1 is the first child born in the 22nd century' nevertheless appears to be true, appears to be justified, appears to express knowledge, appears to express armchair knowledge, and in this case there is no obvious empirical justification anywhere in the vicinity. So there is a very strong *prima facie* case that this utterance expresses *a priori* knowledge. At this point, the Russellian may hold that there is a discontinuity between descriptive names used with and without acquaintance:²³ for example, perhaps the former are disguised descriptions rather than names (so that utterances involving them do not express singular propositions, and may well be *a priori*), while the latter are truly names (so that utterances involving them express singular propositions and are not *a priori*). But the argument from suspension of judgment suggests that even if Russellians are right about this semantic claim, it does not yield a difference at the epistemic or mental level. As discussed earlier, we need to distinguish the role of acquaintance in possessing words and concepts with its role in justifying belief. The arguments in this chapter strongly suggest that even if acquaintance with Julius enables us to possess a certain sort of concept of him, it is not essential

²²This line was put forward by Donnellan (1978), and is also endorsed by Salmon (1988), Soames (2004), and many others.

to justifying our belief in ‘Julius invented the zip (if anyone did)’.

Other Russellians hold that descriptive names are always disguised descriptions and are never names, whether used with or without acquaintance.²⁴ These Russellians may hold that cases involving descriptive names such as ‘Julius’ are no guide to cases involving ordinary proper names such as ‘Gödel’. Even if they are right, however, the bulk of what I have said about the case of ‘Julius invented the zip’ also applies to cases such as ‘*PQTI* → Gödel discovered incompleteness’. In particular, the argument from suspension of judgment suggests that this case is a priori, just as before. One can also run the argument from frontloading, and one can use the diagnostic test as discussed under the objection from acquaintance in chapter 4. So once again, there is a strong case against empirical justification, and a strong case that these conditionals are knowable a priori.

I am inclined to think that the best thing for a Russellian to say is that some singular propositions, such as those expressed by the conditionals above, really can be known a priori, at least under the right mode of presentation. If a theorist insists that singular propositions cannot be known a priori, then I think that this theorist might reconsider their support for the thesis that these propositions are the contents of belief. Such a theorist could still in principle hold on to Russellianism about language, while accepting a non-Russellian view of the contents of thought.

Because of the consistency of views like these, A Priori Scrutability is not outright inconsistent with Russellianism. Still, A Priori Scrutability leads naturally to Generalized A Priori Scrutability, which leads naturally to defining intensions for sentence tokens, which serve as a Fregean semantic value. So once A Priori Scrutability is accepted, it is hard to deny that there is some sort of Fregean content in the vicinity of language and thought.

It remains open to the Russellian to deny that this Fregean content is truly an aspect of meaning: perhaps it is only a content of thought, or a “guise” rather than a content, or a sort of pragmatic content that is nonsemantically associated with utterances. As before, I suspect that claims of this sort will largely turn on verbal issues about what counts as “meaning” and “content”. What matters at the end of the day is the explanatory work to which these Fregean contents can be put. That work is one of the main foci of *The Multiplicity of Meaning*.

²³Reimer (2004) takes a version of the line on which acquaintance transforms content of a name from something descriptive to something referential (although her line holds that if a descriptive name relies on description rather than acquaintance to fix reference even after acquaintance, as in the case I am discussing, the name is still semantically descriptive). Salmon (1998) and Soames (1995) take related views on which although the content of a descriptive name pre-acquaintance is still its referent, the acquaintanceless user of the relevant name is unable to fully grasp propositions involving this referent.

²⁴This sort of view is also associated with Evans (1979), although Evans’ version of the view is not Russellian.

Twelfth Excursus: Scrutability and the Unity of Science.

The unity of science was one of the central concerns of the Vienna Circle. Otto Neurath edited the huge, never-completed *Encyclopedia of Unified Science* (see especially Neurath, Carnap, and Morris 1971). In his 1932 article “The Physical Language as the Universal Language of Science” (translated into English as the 1934 book *The Unity of Science*), Carnap wrote:

The opinion is generally accepted that the various sciences named [philosophy, formal sciences, natural sciences, social sciences] are fundamentally distinct in respect of subject matter, sources of knowledge, and technique. Opposed to this opinion is the thesis defended in this paper that science is a unity, that all empirical statements can be expressed in a single language, all states of affairs are of one kind and are known by the same method.

There is no single thesis of the unity of science. An *imperialist* unity thesis (embraced explicitly at the end of Carnap’s article) holds that all sciences are part of a single science such as physics. A *reductive* unity thesis holds that all correct scientific theories in different domains are somehow reducible to or grounded in a single theory, such as a fundamental physical theory. A *coherence* thesis holds that correct scientific theories in different domains are coherent, in that there are mutually supporting connections between them. A *similarity* thesis holds that correct scientific theories in different domains have some similarity in methods or form. A *consistency* thesis holds only that correct scientific theories should be consistent with each other.

The logical empiricists are often associated with imperialist or reductive versions of the unity of science thesis, although coherence and similarity theses are also prominent in their writings. In recent years, the trend among philosophers of science has been to reject strong unity of science theses in favor of weaker theses such as coherence theses, or to argue that science is not unified at all. It is widely accepted that attention to the scientific practice reveals far more diversity between the sciences than unity.²⁵

Scrutability has at least some bearing on the unity of science. The thesis that all truths are scrutable from base truths naturally suggests that all scientific truths are grounded in certain base truths. And the scrutability theses discussed here give a central role to microphysical truths in the base. So this might suggest a commitment to a strong, reductive version of the unity of science

²⁵For analyses of the logical empiricists on the unity of science, see Symons, Pomdo, and Torres 2011. For recent work favoring the disunity of science, see Dupre 1993, Galison and Stump 1996, and Cartwright 1999.

thesis. This might be read as a point in favor of the scrutability of the framework: it reveals a sense in which science is unified. Or it might be read as an objection to the framework: it is committed to a reductive thesis that the philosophy of science has revealed to be implausible.

The issues here are subtle, but it is worth exploring just what sort of unity thesis might follow from the scrutability thesis. For ease of discussion, I will start by assuming Microphysical Scrutability: the thesis that all truths are scrutable from the microphysical truths P . I will later consider how things are affected by a change to scrutability from $PQTI$. I will discuss both the constructive point (whether scrutability yields some form of unity) and the defensive point (whether objections to unity are objections to scrutability).

I will concentrate mainly on reductive unity theses. For any relation of reduction between theories, there is a corresponding unity thesis: all scientific theories are reducible (in that sense) to a single theory. But many notions of reduction and so of unity can be distinguished. One dimension of variation concerns which aspects of theories we are concerned with: their languages, their laws, their methods, their explanations, their truths? Another dimension concerns the character of the reduction relation: it might be logical (e.g. entailment), metaphysical (e.g. identity), epistemological (e.g. evidential grounding), or semantic (e.g. meaning equivalence). Another dimension concerns the structure of the reduction relation: it might be conditional (yielding one-way conditionals from the reducing theory to the reduced theory) or biconditional (yielding two-way conditionals from one theory to the other).

The Microphysical Scrutability thesis can be seen as a unity thesis involving epistemological conditional relations among truths: that is, the truths in the languages of all correct theories are epistemologically deducible from the truths of fundamental physics. This thesis has a strong reductive flavor. But it does not entail the traditional unity theses that are now widely rejected.

One aspect of the classical conception of reduction is definitional reduction. Definitional reduction concerns semantic biconditional relations among language: the key claim is that the expressions of the reduced theory can be defined in terms of the expressions in the reducing theory. The corresponding classical unity thesis, found explicitly in Carnap's work on unity, is a definitional unity thesis: the expressions of all correct theories are definable in terms of those of a single theory (such as physics).

The scrutability thesis does not entail the definitional unity thesis for a familiar reason: scrutability does not require definitions.²⁶ A definitional reduction from economics to physics would require that economic expressions be definable using microphysical expressions, which in turn requires biconditionals connecting economics and physics. By contrast, scrutability requires only

much weaker one-way conditionals from physical truths to economic truths.

This allows the scrutability thesis to escape perhaps the most well-known objection to classical unity theses: the objection from multiple realizability (e.g. Fodor 1974). On the face of it, economics could be realized in physics or in ectoplasm. Definitional reduction of economics to physics appears to rule out the possibility that economics is realized by anything other than physics. Furthermore, even in a physical world, different instances an economic kind such as money might be grounded in a heterogeneous and open-ended class of physical realizations, suggesting that any physical definition would be wildly disjunctive. By contrast, the scrutability of economic truths from microphysical truths is quite consistent with the the multiple realizability of economic kinds. In fact, the scrutability thesis can allow that in other scenarios, economic truths are scrutable from ectoplasmic truths.

Another aspect of the classical conception of reduction is deductive-nomological reduction, often called Nagelian reduction after Ernest Nagel (1962). Nagelian reduction concerns logical conditional relations among laws: the key claim is that the laws of the reduced theory are entailed by the laws of the reducing theory, perhaps along with bridge laws. We might call the corresponding classical unity thesis a Nagelian unity thesis: the laws of all correct theories are entailed by the laws of a single theory such as physics, along with bridge laws.

The scrutability thesis does not entail the Nagelian unity thesis for a couple of reasons. First, the scrutability relation is weaker than logical entailment. Second, and perhaps more importantly, Microphysical Scrutability does not say that all truths are scrutable from microphysical *laws*: it says that they are scrutable from microphysical *truths*, including the distribution of microphysical items throughout space and time. It follows that any true laws in chemistry, economics, and so on are scrutable from microphysical truths, but not that they are scrutable from microphysical laws.

This allows the scrutability thesis to escape another objection to classical unity theses: the objection from contingency. On the face of it, there is contingency in biology or sociology that goes beyond the contingency of physics. The principles of neuroscience could easily have been different, even keeping physics fixed. Certain key constants of social network theory appear quite arbitrary. So these principles do not seem to be derivable from the laws of physics alone. To

²⁶When generalized scrutability of B-truths from A-truths obtains, there will be at least approximate definitions of B-expressions using A-expressions. One might think that this is enough for multiple realizability to cause problems. In the case of economics and physics, however, we have scrutability but not generalized scrutability, precisely because there are scenarios in which economics is not grounded in physics. Scrutability alone does not support even approximate definitions of economic expressions in physical terms, at least if definitions are required to be a priori.

handle this problem, a Nagelian reductionist needs to allow initial conditions and not just laws in the reduction base. But in any case, there is no analogous problem for the scrutability thesis, which has microphysical truths in the reduction base. The microphysical truths underlying brains and societies will themselves be contingent and arbitrary: even holding physical laws constant, they could have been different. And it is plausible that this contingency matches up well with the contingency of neurobiology and sociology. So for all this objection says, it remains plausible that neurobiological and sociological principles will be scrutable from all the microphysical truths in the vicinity of brains and societies.

Still, the scrutability thesis shares something of the spirit of the Nagelian unity thesis. A priori entailment has something of the spirit of logical entailment: both might be seen as a sort of deducibility. Microphysical truths go beyond microphysical laws, but only so far. If physics is deterministic, microphysical truths are themselves entailed by and scrutable from microphysical laws along with microphysical boundary conditions (the state of the universe at the big bang, perhaps). And even if physics is nondeterministic, microphysical truths will be scrutable from these things along with the values of probabilistic variables. So microphysical scrutability might be seen as sharing some of the attractions of this classical unity thesis, without some of its costs.

In one respect, the scrutability thesis is stronger than the Nagelian unity thesis. The classical thesis allows bridging laws in the entailment base: chemical truths are entailed by physical truths plus physical-chemical bridging laws. Scrutability does not allow bridging laws in the base: chemical truths are a priori entailed by microphysical truths. Where logical entailment is concerned, bridging laws play the helpful role of connecting vocabularies, but where a priori entailment is concerned, this role is not needed: truths in one vocabulary can be a priori entailed by truths in a quite different vocabulary. One might think of the framework as akin to one that requires the bridging laws to be a priori, except that as we saw in chapter 1, a priori entailment does not require explicit bridging laws or definitions at all.

This difference is a benefit rather than a cost of scrutability. As Kim (1998) has pointed out, allowing bridging laws makes the Nagelian conception of reduction much too weak. To see this, note that many mind–body dualists (including myself) allow that there are laws connecting physical properties to mental properties, so that mental truths will be logically entailed by physical truths plus psychophysical bridge laws. The Nagelian model appears to predict that on this view, the mental is reducible to the physical. But such a claim obviously mischaracterizes the dualist view. The underlying trouble is that there can be laws connecting entirely distinct domains, each of which is irreducible to the other. So for a connection that deserves to count as reducibility, mere

bridging laws do not suffice.

Scrutability invokes the much stronger requirement of a priori entailment, which brings with it a sort of epistemological deducibility of higher-level truths from lower-level truths. It is arguable that something like this is required to satisfy one key desideratum of reducibility: that any brute facts in the the higher-level domain be grounded in brute facts in the lower-level domain. Allowing bridging laws subverts this desideratum: bridging laws can themselves introduce brute facts, as the case of mind–body dualism suggests. By contrast, scrutability favors the desideratum, at least if we allow that there are no brute facts in the a priori domain.

Of course there are many different notions of reduction, and there is no point getting into a verbal dispute over what counts as ‘reduction’. But the desideratum outlined above corresponds to at least one key notion of reduction, or one key constraint on such a notion. We might call it *transparent bottom-up explanation*: once one has spelled out the lower-level facts, the higher-level facts are transparent. This sort of explanation is a goal of many reductive projects in science. Reductive projects naturally have aims such as the following: once one has spelled out all the physical facts about an organism, the chemical facts will be transparent. Once one has spelled out all the chemical facts, the biological facts will be transparent. We may not have explained why all the physical facts obtain, but given that they obtain, there will be no residual mystery about why and how they give rise to the chemical or the biological facts.

Where scrutability fails, transparent bottom-up explanation fails. This is borne out by the mind–body case: even after spelling out all the physical facts, the mental facts are not transparent, so there is a residual mystery about how the physical gives rise to the mental. Of course there are options that are intermediate between a priori scrutability and bridging laws. For example, one could appeal to a posteriori identities or a posteriori necessities connecting low-level and high-level domains, for example. But even these leave an element of bruteness in an explanation. If one “explains” consciousness by saying that it is identical to a certain neural state and leaves it at that, then one has not given a transparent bottom-up explanation. In effect, the identity claim plays the same sort of explanatory function as a bridging law in the case of mind–body dualism. When scrutability fails, there will be a priori coherent scenarios in which the low-level facts are as they are and the higher-level facts are different. These scenarios cannot be ruled out by the low-level facts alone, so the low-level facts do not transparently explain the higher-level facts. Instead, one needs primitive interlevel bridging principles in one’s explanatory theory.

Of course many cases of reduction involve interlevel identities: the reduction of water to H₂O is one such. But in this case, the identity claim ‘water is H₂O’ is itself scrutable from lower-level

truths.²⁷ In this sort of case, the high-level truths are scrutable from and transparently explainable in terms of lower-level truths. But when the identity claim is not scrutable in this way (as in the consciousness case), it effectively functions as a primitive claim in a bottom-up explanation, playing the same epistemological role as a brute bridging law. To remove this element of bruteness and achieve transparency, something stronger is required. Scrutability can naturally play that role.

It might be argued that scrutability is too weak for transparent bottom-up explanation, on the grounds that a priori entailment can connect distinct domains. For example, if mathematical truths are a priori, then they are a priori scrutable from physical truths (or by any other class of truths), but they need not be reducible to physical truths in any reasonable sense. Likewise, two sets of truths can be a priori scrutable from each other, but it seems odd to hold that they can be reducible to each other. I think that this is a reasonable criticism, and suggests that scrutability needs to be strengthened to yield the relevant sort of reduction. Here one might strengthen the requirement by moving from a priori entailment to the stronger sort of in-virtue-of claims discussed in chapter 1 and the nineteenth excursus (especially the conceptual grounding relation discussed there), or by moving to more specific models of scrutability-based explanations such as the mechanistic model that follows. In any case, scrutability will still plausibly be a necessary condition for a relevant sort of reduction.

Scrutability is also a weak constraint insofar as good reductive explanations require the low-level phenomena doing the explaining to have a certain internal unity. Scrutability could be satisfied even if microphysical truths were entirely chaotic, non-law-governed, and disunified; but in that case microphysical truths at best explain macrophysical truths in a weak sense. In the actual world, microphysical truths have a certain internal simplicity and autonomy that makes for better explanations than this, but the degree of simplicity will vary from case to case: a reductive explanation of the Second World War might be a poor one, precisely because the microphysical base would be so complex. Still, even when the base is arbitrarily complex, scrutability allows a sort of transparent bottom-up explanation: *given* the low-level truths, high-level truths fall out. Good reductive explanation requires something more, but scrutability will again be a necessary

²⁷Jackson (1998) gives a nice model of the water/H₂O case, arguing that that 'Water is H₂O' can be derived from microphysical facts using the a priori premise 'Water is what plays the water role' and the empirical premise 'H₂O plays the water role', which is itself derivable from microphysical facts. This in effect invokes a definition of 'water' (although a functional rather than a microphysical definition) to ground the derivation. As always, the scrutability framework can dispense with the definition here, but 'water is H₂O' will nevertheless be scrutable insofar as it is scrutable that H₂O plays the key roles associated with water.

condition.

A model of reduction that is quite compatible with scrutability while imposing further constraints is one grounded in *mechanistic explanation*.²⁸ On this model, high-level phenomena are explained in terms of the orchestrated functioning of a mechanism: a structure performing a function in virtue of its components parts, component operations, and their organization (Bechtel 2005). For example, DNA and RNA molecules might serve as a mechanism by which the transmission of hereditary characteristics is enabled, thereby explaining genetic phenomena. Mechanistic explanation typically proceeds via functional analysis of high-level phenomena, casting high-level explananda in terms of functional roles. For example, the genetic phenomena that need to be explained are the functional roles of transmitting hereditary information. One then shows how lower-level mechanisms can play those roles. For example, one shows how DNA can transmit hereditary information. In this way, one achieves a sort of transparent bottom-up explanation.

Employing the scrutability model, we can divide this picture into three parts. First, high-level explananda are expressed using functional concepts, or concepts involving functional roles: for example, the concept of a gene can be seen as a concept of an entity that transmits hereditary characteristics in a certain way. Second, one tells a story about how low-level mechanisms play the relevant roles: for example, about how DNA transmits hereditary characteristics. Third, given that the roles in the mechanistic story and the functional analysis match up well enough, high-level truths will be scrutable from the mechanistic story. In effect, functional analysis grounds scrutability from underlying mechanisms.

I do not say that reductive or mechanistic explanation in science requires scientists to demonstrate an a priori entailment from low-level truths to high-level truths. That claim would be much too strong. Still, I think there is an important sort of reductive explanation in science for which scrutability is at least a tacit constraint. That is, it is a tacit desideratum that in principle, a given reductive story could be fleshed out with further lower-level truths, such that higher-level phenomena would be scrutable from there. If it turned out such such scrutability were impossible in principle, then the reductive explanation could reasonably be regarded as defective, or as failing to satisfy an important desideratum of transparency. In practice, a reductive explanation will proceed by giving just enough detail to make plausible that some sort of fleshed-out story of this sort could be obtained.

²⁸For my own version of a mechanistic picture of reductive explanation, see section 2 of Chalmers 1995. The scrutability model is also compatible with other sorts of reductive explanation, including structural explanation as well as functional, but functional explanation is certainly the most common kind.

Some will think that scrutability is too strong, on the grounds that the connections between physics and biology, say, are empirical rather than a priori. I have already answered this sort of objection in arguing for scrutability. But it is worth keeping in mind again that scrutability does not require definition of biological notions in microphysical terms, and allows us to appeal to all microphysical truths and not just microphysical laws. And as before, even though bridging principles such as “water is H₂O” are empirical, this is no bar to the a priori scrutability of the principles themselves from low-level truths.²⁹ It remains plausible that someone using a Cosmoscope armed with all microphysical truths (along with phenomenal and indexical truths) could ascertain all the biological truths and all bridging principles connecting physics and biology, for example. There are tricky cases here, such as the interface between the quantum and classical domains, but these cases can be handled as in the discussion of macrophysical truths above.

30

Some may worry that other standard worries for Nagelian accounts of reduction will apply to scrutability. It is clear that standard problems tied to definability, to multiple realizability, and to bridge laws will not arise. Nor will problems tied to logic: some versions of a Nagelian account require that all theories be formulated in first-order logic, but scrutability does not. Another problem for Nagelian reduction concerns the “reduction” of an old theory to a new one: the old theory contains falsehoods, which cannot be entailed by truths. The falsehoods in the old theory will not be scrutable from truths either, but various nearby truths will be, including claims that those falsehoods are approximately true, or true in certain circumstances.

One might also worry about the autonomy of the high-level sciences. Cellular biology, cognitive psychology, economics, and paleontology are all enormously different from physics and from one another. It would be crazy to do cognitive psychology by doing physics. These fields have their own methods and their own conceptual and ontological frameworks. Perhaps most importantly, they all have a sort of explanatory autonomy: economic explanations are different in kind from microphysical explanations, and cannot be begin to be replaced by microphysical explanations.

Scrutability is quite consistent with explanatory autonomy. If an economic truth (say, about the financial crisis in 2008) is scrutable from physical truths, then a weak sort of explanation of the economic truths in terms of physical truths will be possible. Given that the physical truths are as they are, we will be able to derive the existence of the financial crisis and so reductively “explain” it. But for most purposes this will be a much poorer explanation than an economic explanation

³⁰Marras (2005) argues against models of reduction in terms of a priori entailment by arguing that bridge laws are empirical and known inductively. I hope it is clear by now that this argument involves a non sequitur.

(in terms of credit mechanisms, for example). The “explanation” will presuppose an enormously complex set of physical truths. Even if these truths are grounded in laws and boundary conditions, the boundary conditions and perhaps the laws will have much irrelevant complexity. This “explanation” may have little predictive power and little practical use. By contrast, an economic explanation may be far simpler, more systematic, more predictive, and more useful.

In general, I favor explanatory pluralism: there are multiple explanations of most phenomena, and which explanation we choose depends on our purposes. There are causal explanations, historical explanations, reductive explanations, and many others. Reductive explanations are useful for some purposes, especially in trying to get a sense of how the world as a whole hangs together (how could there be economic phenomena in a physical world?). These explanations help to give us a unified picture of the world. But for most purposes, they cannot take the place of other explanations.

Overall, we can see scrutability as a weak sort of reduction, one that is compatible with various sorts of irreducibility that are manifest in science. One might label it (as I do in Chalmers 1996) a sort of reductive explanation without reduction, where the relevant variety of reductive explanation involves transparent bottom-up explanation in terms of underlying truths. At least, there is plausibly a notion of reductive explanation here for which scrutability is a necessary condition.³¹

Correspondingly, the scrutability thesis can be seen as a weak sort of unity thesis, consistent with the various manifest respects in which science is disunified. It avoids the most prominent objections to classical unity theses, but at the same time it shares something of their spirit, and it can do at least some of the work that we might want a reductive unity thesis to do.

Of course microphysical scrutability is false, at least on my view. I think that phenomenal truths, indexical truths, and a that’s-all truth are not scrutable from microphysical truths. Correspondingly, I think that these are not explainable in terms of physical truths. But we can add these to the scrutability base, yielding the thesis that all truths are scrutable from *PQTI*. How does this alter the foregoing?

If phenomenal truths are not scrutable from the microphysical, this brings out a certain disunity of the sciences. If we equate scrutability with reductive explanation, then phenomenal truths will not be reductively explainable in terms of microphysical truths. Nor will truths whose scrutability requires phenomenal truths: perhaps mental truths, social truths, secondary-quality truths, and others. These truths will be explainable in terms of physics and phenomenology, but not in terms

³¹For more on the relationship between a priori entailment and reductive explanation, see Chalmers 1996, chapter 2 (sections 2 and 3), and Chalmers and Jackson 2001.

of physics alone. Something similar goes for truths whose scrutability requires indexicals. For example, objective physical truths may leave open whether water is H_2O or XYZ, so that a fully transparent explanation of the truth that water is H_2O requires an appeal to indexical claims about our location within the world. Likewise for the that's-all truth: in effect, positive truths can be reductively explained in terms of microphysical truths alone, but a full explanation of negative truths requires something more.

Despite this expansion, certain stripped-down analogs of many of these truths will be reductively explainable. As Searle has noted, physics can explain the "objective" aspects of heat and color if not the "subjective" aspects. Here we might think of a subjective truth (in the relevant sense) as one with a relevant dependence on phenomenal or indexical truths, and objective truths as those without such a dependence. One might define an objectivized notion of 'heat' solely in terms of an objective causal role (perhaps in terms of expanding metals and the like), leaving out any connection to experience. Then objective truths involving this notion might be reductively explainable in terms of microphysical truths. The same goes for an objectivized notion of 'water', which will apply equally to H_2O and XYZ. Given that all truths are scrutable from *PQTI* and that scrutability entails reductive explainability, it follows that all positive objective truths will be explainable in terms of microphysical truths, and that all positive truths will be explainable in terms of microphysical truths and a that's-all clause. Here the objective truths might be seen as one version of Sellars' "scientific image" (where subjective truths are part of the "manifest image"). We will then have at least a sort of strong unity of the scientific image, so conceived.

Still, the scientific image so conceived may be a pale reflection of actual science. Consciousness, mentality, sociology, secondary qualities, and other subjective aspects of the manifest image are all subject matters for science. A unity thesis that covers all of these will need to have more than microphysics in the base: it will need phenomenal and indexical truths too. The role of indexical truths is relatively minor. The most important addition will be certain psychophysical bridging principles: laws, identities, or necessities linking physical properties to phenomenal properties. As long as phenomenal properties supervene on physical properties, this addition will bring phenomenal truths into the fold.

We might put this picture by saying that all scientific truths are grounded in physics and psychophysics. The dual base here is less unified than a purely microphysical base, but it still allows a good deal of unification. We still have a unified scientific picture of the world grounded in a few fundamental properties linked by a few fundamental laws, albeit with more properties and laws than on the physicalist picture. This reinforces the moral of the prior discussion: while various

strong unity theses fail, the scrutability framework supports at least a moderate and attenuated conception of the unity of science.