

Chapter 5: Revisability and Conceptual Change

David J. Chalmers

1 Introduction

Perhaps the most famous attack on Carnap's logical empiricism is W.V. Quine's article "Two Dogmas of Empiricism". In the article, Quine argues against the analytic/synthetic distinction that Carnap employs, as well as against his sometime verificationism. The article is widely regarded as much more than an attack on logical empiricism, however. It is often seen as the most important critique of the notion of the a priori, with the potential to undermine the whole program of conceptual analysis.

In this chapter, I address Quine's most influential arguments in "Two Dogmas". I do this in part for defensive reasons and in part for constructive reasons. Defensively: Quine's arguments might be thought to undermine my frequent appeals to the a priori, so addressing these arguments helps to support those appeals. Constructively: addressing Quine's arguments in the spirit of the scrutability framework helps to bring out some of the power of that framework. For example, it helps us to analyze notions of meaning in a broadly Carnapian manner, and it helps us to understanding the relationship between rationality and conceptual change.

I will address Quine's article construed as a critique of the notions of analyticity and apriority. I am more inclined to defend the notion of apriority than the notion of analyticity, so I will focus more on the former, but the response that I will develop can be used to defend either notion from Quine's arguments. I will focus especially on the most influential part of Quine's article: the arguments in the final section concerning revisability and conceptual change.

In addressing these arguments, I will adopt a line of response grounded in Carnap's underappreciated article "Meaning and Synonymy in Natural Language". Carnap's article offers an approach to issues about meaning that is highly congenial to the scrutability framework. I will argue that an analysis inspired by this article, when conjoined with tools drawn from the scrutability framework and from Bayesian confirmation theory, provides just what is need to reject Quine's argument.

Along the way, I will motivate a Carnap-style analysis of meaning within the current framework. The analysis of meaning is developed further in the tenth and twelfth excursuses. In the ninth excursus I address objections to scrutability from conceptual change.

2 The Arguments of “Two Dogmas”

In sections 1 through 4 of “Two Dogmas of Empiricism”, Quine argues that if one tries to make sense of the notion of analyticity, one ends up moving in a circle through cognate notions (synonymy, definition, semantic rules, meaning), and one cannot break out of the circle. Many philosophers have not been greatly moved by this worry, as it seems that one finds a similar circle for all sorts of philosophically important notions: consciousness, causation, freedom, value, existence. So I will set these criticisms aside here.

In section 5 of the article, Quine makes points that are addressed specifically at Carnap’s logical empiricism, criticizing his construction of physical concepts from phenomenal concepts in the *Aufbau*, and his verification theory of meaning. I will set these points aside here, as I am not concerned to defend Carnap’s phenomenalist construction or the verification theory of meaning.

The extraordinary influence of Quine’s article can be traced in large part to the short final section of the article. Part of this influence stems from the positive picture that Quine offers in the first paragraph of the section, characterizing the totality of our knowledge as a “man-made fabric which impinges on experience only along the edges”, in which “no particular experiences are linked with any particular statements in the interior of the field, except indirectly through considerations of equilibrium, affecting the field as a whole”. This picture serves as a powerful alternative to the verificationist picture provided by some logical empiricists. It does not contain any direct argument against the analytic/synthetic distinction or related notions of apriority, however.

The most influential arguments against an analytic/synthetic distinction are found in the second paragraph, which I quote here in full:

“If this view is right, it is misleading to speak of the empirical content of an individual statement—especially if it be a statement at all remote from the experiential periphery of the field. Furthermore it becomes folly to seek a boundary between synthetic statements, which hold contingently on experience, and analytic statements which hold come what may. Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system. Even a statement very close to the

periphery can be held true in the face of recalcitrant experience by pleading hallucination or by amending certain statements of the kind called logical laws. Conversely, by the same token, no statement is immune to revision. Revision even of the logical law of the excluded middle has been proposed as a means of simplifying quantum mechanics; and what difference is there in principle between such a shift and the shift whereby Kepler superseded Ptolemy, or Einstein Newton, or Darwin Aristotle?"

I will focus on these critical arguments here. There are two crucial points.

(Q1) "Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system."

(Q2) "No statement is immune to revision."

If (Q1) and (Q2) are read as mere psychological claims, saying that as a matter of fact someone might hold onto or revise any statement, then they are highly plausible, but not much of interest will follow from them. But Quine is saying something more than this. We can understand (Q1) as saying that any statement can be *rationally* held true come what may, and (Q2) as saying that that no statement is immune to *rational* revision. These points have interesting consequences.

Many have taken these points to suggest either that no sentences are analytic, or that no distinction can be drawn between analytic and synthetic sentences. One possible connection goes via the theses that analytic sentences are those that can be rationally held true come what may and that all analytic sentences are immune to rational revision. If so, (Q1) suggests that by the first criterion, all sentences will count as analytic. And (Q2) suggests that by the second criterion, no sentence will count as analytic. Either way, there is no useful distinction between analytic and synthetic sentences to be had. Similarly, if we assume the theses that a priori sentences are those that can be rationally held true come what may and that all a priori sentences are immune to rational revision, (Q1) and (Q2) suggest that there is no useful distinction between a priori and a posteriori sentences to be had.

One common way of responding to the argument from (Q2) is to suggest that revisability is quite compatible with apriority (or analyticity), on the grounds that a priori justification (or the justification we have for believing analytic sentences) is *defeasible*.¹ For example, I might know a mathematical claim a priori, but my justification might be defeated if I learn that a leading mathematician thinks that the claim is false. I think that this response is correct as far as it goes, but to rest entirely on this response to Quine would be to concede a great deal. On a common traditional

conception (not far from the conception of the conclusive a priori outlined in the last chapter), at least some a priori justification (and some justification for believing analytic truths) is *indefeasible*. One might reasonably hold that some a priori justification (in logic or mathematics, say) yields not just knowledge but certainty, at least on ideal reflection. These claims are not obviously correct: for example, a defense of them would have to address worries about metacognitive skepticism. But they are also not obviously incorrect, and I do not think that Quine's argument establishes that they are false. To see why not, we need another line of response.

The response I will develop takes off from the response given by Grice and Strawson at the end of their article "In Defense of a Dogma". This response holds that (Q1) and (Q2) are compatible with an analytic/synthetic distinction, for a reason quite different from the one given above. Here is a passage addressing the argument from (Q2):

Now for the doctrine that there is no statement which is in principle immune from revision, no statement which might not be given up in the face of experience. Acceptance of this doctrine is quite consistent with adherence to the distinction between analytic and synthetic statements. Only, the adherent of this distinction must also insist on another; on the distinction between that kind of giving up which consists in merely admitting falsity, and that kind of giving up which involves changing or dropping a concept or set of concepts. Any form of words at one time held to express something true may, no doubt, at another time, come to be held to express something false. But it is not only philosophers who would distinguish between the case where this happens as the result of a change of opinion solely as to matters of fact, and the case where this happens at least partly as a result of a shift in the sense of the words. Where such a shift in the sense of the words is a necessary condition of the change in truth-value, then the adherent of the distinction will say that the form of words in question changes from expressing an analytic statement to expressing a synthetic statement. ... And if we can make sense of this idea, then we can perfectly well preserve the distinction between the analytic and the synthetic, while conceding to Quine the revisability-in-principle of everything we say.

¹See, for example, Field (1996). Kitcher (2000) defends a conception of the apriority that requires indefeasibility, while Peacocke (2004) defends a conception that does not. For present purposes I will remain neutral on whether apriority entails some sort of ideal indefeasibility; the observation about testimony in section 7 contains some relevant discussion.

Here the central point is that our judgments about any *sentence*, even an analytic sentence, will be revisable if the meaning of the words change. For example, if ‘bachelor’ changes from a term for unmarried men to a term for sociable men, then we will no longer judge that ‘All bachelors are unmarried’ is true. But this observation is just what an adherent of the analytic/synthetic distinction should expect. Analytic sentences should instead be understood as those sentences that are immune to rational revision *while their meaning stays constant*. More precisely, they are those that are immune to rational rejection while their meaning stays constant. There is a sense in which an analytically false sentence might be immune to rational revision, but from here onward I will understand “revision” as requiring rejection.

Following standard practice, we can say that when the meaning of a sentence changes, there is *conceptual change*: some expression in the sentence at first expresses one concept and later expresses another. When the meaning of a sentence stays the same, there is *conceptual constancy*: the expressions in the sentence will express the same concepts throughout. Then Grice and Strawson’s point could be put by saying that an analytic sentence is one that is immune to revision *without conceptual change*. More cautiously, the point could be put by saying that even if a sentence is analytic, it may still be revisable under conditions of conceptual change. Something similar applies to apriority.²

At this point, Quine has two obvious replies. The first reply is to say that the appeal to meaning in characterizing the class of analytic sentences is circular, as the notion of meaning as poorly understood as the notion of analyticity. The same could be said for the appeal to concepts and to propositions. This reply would be in the spirit of the first four sections of “Two Dogmas”. But then this argument will not be much of an advance on the arguments in the first four sections, and anyone who is not moved by those arguments will not be moved by this one.

The second, more interesting reply is to challenge Grice and Strawson to provide a *principled distinction* between cases of revision that involve conceptual change and those that involve conceptual constancy. Quine might argue that cases that are purported to be on either side of this division are in fact continuous with each other, and that there is no principled distinction to be

²We might also allow that there is conceptual change in this sense when the proposition expressed by an utterance of a sentence changes because of a shift in context. For example, ‘Someone is bald iff they have no hairs’ might be accepted in one context and rejected in another. It is not clear that a mere contextual shift could change the status of a sentence as analytic, as arguably the meaning of such a sentence stays constant throughout. But if we say that a sentence is a priori if it expresses a proposition that is knowable a priori, then it is natural to hold that in these conditions a sentence might be a priori in one context but not in another.

had. Something like this thought might even be read into the last sentences of the paragraph from Quine quoted above.

Now one might suggest that Grice and Strawson are not obliged to provide a *reductive* characterization of the distinction—that is, one that does not use ‘meaning’ and cognate notions—any more than they are required to provide a reductive definition of meaning or analyticity to answer the challenge in the first four sections. Again, this suggestion seems correct as far as it goes. But nevertheless, if Quine’s opponent cannot say much to characterize the principled distinction here, he or she is at least in the awkward dialectical position of leaving a challenge unanswered, and of leaving doubts about the distinction unassuaged.

My own view is that there is much that can be said to flesh out a principled distinction here. In particular, the scrutability framework provides a way to motivate a principled distinction here, building on tools first set out in Carnap’s “Meaning and Synonymy in Natural Language”.

3 Carnap on intensions

Carnap is Quine’s major target in “Two Dogmas of Empiricism”. It is not always appreciated that “Meaning and Synonymy in Natural Language” can be read as a sustained response to Quine, perhaps because Carnap spends little time discussing Quine in the article. Nevertheless, Carnap says enough to make clear that a response to “Two Dogmas” is intended.

Carnap’s article sets out to provide an analysis of the notion of meaning and of related notions such as synonymy. His aim is to provide a “scientific procedure” by which meaning and synonymy can be analyzed in broadly naturalistic terms. Importantly, he aims to explicate not only the notion of extension, but the notion of intension (the “cognitive or designative component of meaning”), which he notes has been criticized by Quine as “foggy, mysterious, and not really understandable”.

Carnap’s key idea is that we can investigate the intension that a subject associated with an expression by investigating the subject’s *judgments about possible cases*. To determine the intension of an expression such as ‘Pferd’ for a subject, we present the subject with descriptions of various logically possible cases, and we ask the subject whether he or she is willing to apply the term ‘Pferd’ to objects specified in these cases. If we do this for enough cases, then we can test all sorts of hypotheses about the intension of the expression.

In this article Carnap takes the term ‘intension’ as a primitive, and does not build possible cases into the very nature of intensions. But for our purposes it is useful to adopt a suggestion that Carnap makes elsewhere, and simply define an intension as a function from possible cases to

extensions. For a term like 'Pferd', the intension will be a function from possible cases to objects characterized in those cases. For a sentence such as 'Grass is green', the intension will be a function from possible cases to truth-values. Then Carnap's procedure above can be regarded as a way of directly ascertaining the values of the intension that a subject associates with an expression, by presenting the subject with a possible case and noting the extension that the subject associates with the case.

Of course one cannot actually present a subject with all possible cases to determine every aspect of an intension. But Carnap suggests that the intension that a speaker associates with an expression is determined by the speaker's *linguistic dispositions*. For a given expression E used by a given speaker, the speaker will have the disposition to associate a given extension with E , when presented with a possible case. For example, if S is a sentence, the speaker will have the disposition to judge the sentence as true or false of a possible case, when presented with that case. The intension of an expression can then be seen as a function that maps possible cases to the extension that the speaker is disposed to identify, when presented with that case.

In this way, Carnap defines an expression's intension in naturalistic and even operational terms. We can go on to define synonymy: two expressions are synonymous (for a speaker at a time) when they have the same intension (for that speaker at that time). And we can define analyticity: a sentence is analytic (for a speaker at a time) when its intension has the value "true" at all possible cases (for that speaker at that time).

With this definition in hand, we can go on to provide a principled criterion for conceptual change over time. An expression E undergoes change in meaning between t_1 and t_2 for a speaker iff the speaker's intension for E at t_1 differs from the speaker's intension for E at t_2 . If we accept Carnap's dispositional account of intensions, it follows that E undergoes change in meaning between t_1 and t_2 iff there is a possible case such that the speaker is disposed to associate different extensions for E when presented with the case at t_1 and t_2 .

Of course there are many immediate questions about Carnap's account. What is a possible case? In what vocabulary are these cases specified? How can we determine whether the meaning of this vocabulary has changed? Cannot speakers make mistakes about intensions? Cannot they change their mind about a case without a change in meaning? Can meaning really be operationalized this easily? And so on. Carnap's account may need to be modified or at least refined to answer these questions.

Before addressing these matters, though, I will illustrate how Carnap's account might be used to address the challenge in section 6 of "Two Dogmas" directly. In my view, the essential aspects,

if not the specifics, of the resulting response are sound. These essential aspects can then be carried over to more refined analyses couched within the scrutability framework (section 5) and in terms of Bayesian confirmation theory (section 6 and 7).

4 A Carnapian response

In “Meaning and Synonymy in Natural Language”, Carnap does not mention the arguments in section 6 of “Two Dogmas”, and he does not address considerations about revisibility or conceptual change. Nevertheless, his framework can be used to give a response to these arguments that is broadly in the spirit of Grice and Strawson’s response, fleshed out with a principled criterion for conceptual change.

We can start with Quine’s observation that any statement can be held true come what may. This seems correct. Even a paradigmatic synthetic sentence such ‘All bachelors are untidy’ can be held true in the face of apparently countervailing evidence, if we allow sufficient adjustment of ancillary claims. The question is whether such adjustments will involve conceptual change, and whether we have a principled criterion for determining this.

We might as well start with a case. At t_1 , Fred asserts ‘All bachelors are untidy’. At t_2 , Fred is presented with evidence of a tidy unmarried man. Fred responds: ‘He’s no bachelor! Bachelors must be over 30, and he’s only 25’. At t_3 , Fred is presented with evidence of a 35-year-old with a spotless apartment. Fred responds: ‘He’s not tidy! Look at the mess in his sock drawer.’ In this way, Fred holds the sentence true throughout, and through similar maneuvers he may hold it true come what may.

Does this case involve conceptual change? We can apply Carnap’s analysis to see whether the Fred’s intension for ‘All bachelors are untidy’ (call this sentence B) changes over the relevant timespan. Suppose that c is a detailed possible case in which there is an unmarried 25-year old with a tidy apartment. At t_2 , when Fred is presented with the information that c obtains, he responds that ‘All bachelors are untidy’ is true with respect to c . By Carnap’s criterion, Fred’s intension for B is true with respect to c at t_2 .

What about Fred’s intension for B at t_1 ? The key question is: if Fred had been presented with a description of c at t_1 , before he had evidence that the case was actual, would he have judged that ‘All bachelors are untidy’ was true with respect to c ?

If the answer is yes, then Carnap’s criterion suggests that need be no conceptual change involved in the change of judgment between t_1 and t_2 . In this case, Fred will simply have had an

unusual intension for ‘bachelor’ all along.

If the answer is no, then Carnap’s criterion suggests that there is relevant conceptual change between t_1 and t_2 . The intension of ‘All bachelors are untidy’ will have changed during this time, probably because the intension of ‘bachelor’ has changed during this time.

The same applies more generally. If a speaker’s judgment concerning a case at t_2 is reflected in the speaker’s dispositions to respond to such a case at t_1 , we can say that the speaker’s judgment concerning that case is *prefigured*. If a speaker’s judgment concerning a case at t_2 is not reflected in the speaker’s dispositions at t_1 , we can say that the speaker’s judgments concerning the case is *postfigured*. On Carnap’s account, postfigured judgments but not prefigured judgments involve conceptual change.

In any case, we have what is needed. Carnap’s framework allows us to see how any sentence can be held true come what may, while at the same time allowing a principled way to distinguish between those cases of holding true that involve conceptual change and those that do not. Something similar applies to cases of revisability, though I will not go through the details here.

5 Refining Carnap’s account

Carnap’s account of meaning is remarkably simple, and one might reasonably wonder whether such a simple account can be correct. The questions raised at the end of section 3 still need to be answered. I think that while some of these questions raise problems for the account, they can be addressed in a way that preserves something of the spirit of the account, if not the letter. In particular, the scrutability framework can be used to characterize intensions in a way that meets many of Carnap’s needs.

The central idea here is that Carnap’s possible cases correspond to scenario specifications such as $PQTI$, and his intensions will be defined by conditional or a priori scrutability with respect to these scenarios. Suppose our subject uses a sentence S . To evaluate the intension of S at a scenario w , specified by a sentence $PQTI^*$, we need only consider whether S is scrutable from $PQTI^*$ for the subject.

We can suppose that our subject is presented with a Cosmoscope carrying the information in $PQTI^*$, and uses it in conditional mode. Under the supposition that the Cosmoscope accurately describes reality (that $PQTI^*$ is true), the scrutability thesis suggests that ideal reasoning should lead the subject to a verdict about S . If this verdict is positive (that is, if S is scrutable from $PQTI^*$ for the subject), then the intension of S is true at w . If this verdict is negative (that is, if $\neg S$ is

scrutable from $PQTI^*$ for the subject), then the intension of S is false at w . And so on.

This framework allows us to answer the central questions for Carnap's account. The first two questions concern possible cases: what are they, and how are they specified? For our purposes, possible cases should be epistemically possible scenarios. Scenarios can be specified by $PQTI$ -like sets of sentences: $PQTI$ for one's actual scenario, variants on it for scenarios close to home, and sentences using different vocabulary for scenarios of very different sorts. For a full vocabulary for specifying scenarios, we can invoke the Generalized Scrutability thesis. Any scenario will be specified by a maximal epistemically possible set of sentences in a generalized scrutability base, where a set of sentences is epistemically possible when its conjunction cannot be ruled out a priori. Then Generalized Scrutability implies that given any truth-apt sentence S and any scenario w , a truth-value for S will be scrutable from the set of sentences specifying w .

A third common worry about Carnap-style accounts of meaning is that on the contemporary understanding, intensions are often not accessible to a subject, even by ideal reasoning. For example, if Kripke (1972) is right, the intension of 'water' picks out H_2O in all possible worlds, even for subjects who do not know that water is H_2O . Such subjects will not be disposed to identify H_2O as the extension of 'water' when presented with a possible case, so Carnap's definition will get the intension wrong.

The scrutability framework handles this issue by distinguishing epistemic possibilities from metaphysical possibilities, and by distinguishing epistemic and modal profiles as in chapter 1. Kripke's observation concerns the modal profile of 'water': the way the expression applies to metaphysically possible worlds. Where the modal profile is concerned, we think of possible cases counterfactually: if XYZ *had been* the liquid in the oceans and lakes (as in Putnam's Twin Earth), would water have been H_2O ? Kripke and Putnam suggest that we should answer positively here, suggesting that the modal profile of 'water' always picks out H_2O . In the two-dimensional semantic framework, modal profiles correspond to secondary intensions: so the secondary intension of 'water is H_2O ' is true at a Twin Earth world. To evaluate a secondary intension at a world, one often first needs empirical information about the actual world (e.g. the knowledge that water is actually H_2O).

By contrast, the sort of intensions we are analyzing here are primary intensions. Primary intensions are more relevant to the a priori than secondary intensions, because they in effect capture an expression's epistemic profile (chapter 1): the way the expression applies to epistemically possible scenarios. Here subjects consider the possible cases as epistemic possibilities: if XYZ *is* the liquid in the oceans and lakes, is water H_2O ? If we ask subjects to suppose that XYZ is actually

in the oceans and lakes, they should conclude conditionally that water is XYZ (their credence that water is XYZ, conditional on being in a Twin Earth scenario, should be high). This suggests that in a Twin Earth scenario, with XYZ in the oceans and lakes, the primary intension of 'water is XYZ' is true, and the primary intension of 'water' picks out XYZ.

This is just the sort of conditional reasoning familiar from Conditional and A Priori Scrutability. One supposes that a certain scenario actually obtains, and one considers what follows. The arguments of the previous chapters suggest that even if we need empirical information to evaluate secondary intensions, we do not need it to evaluate primary intensions. So the Kripkean worry for Carnap-style account does not arise where scrutability is concerned.

A fourth problem for Carnap's account as it stands is that subjects can make mistakes. A subject might miscalculate and judge that $36+27=73$, and they might even be disposed to judge this to be true with respect to all possible scenarios. On Carnap's account, it will follow that '36+27=73' is analytic for the subject. But this seems the wrong result: on the face of it, the sentence is not even true. Similar mistakes seem possible for non-ideal subjects in all sorts of domains.

The scrutability framework handles this problem in a familiar way: by invoking a normative idealization (chapter 2). Instead of appealing to what the subject *would* say in response to the case, we appeal to what the subject ideally *should* say. Or in modal terms: the intension of E maps a possible case c to the extension that the subject would identify for E , if they were to be presented with c and were to reason ideally. In effect, the idealization built into the notions of conditional or a priori scrutability allows these notions to define idealized intensions.

Construed in this idealized way, the account will no longer yield an operational definition of meaning, at least unless we can find an operational criterion for ideal reasoning. But this is not a bad thing for those who are inclined to reject behaviorism in any case. It is also far from clear that this account provides a naturalistic reduction of meaning: it will do so only if we already have a naturalistic reduction of ideal reasoning. But the account need not be a naturalistic reduction to be useful.

Someone might suggest that in these cases, it is facts about meaning that determine facts about ideal reasoning rather than vice versa: it is precisely because we mean such-and-such by 'Pferd' that we should say such-and-such. We need not take a stand on these questions about metaphysical priority here. All that we need is that in these cases, there are some facts about what subjects should say or about what ideal reasoning dictates, and that we have some sort of pretheoretical grip on these facts. Then these facts can be used to help us explicate a corresponding notion of

meaning, regardless of which of these notions is metaphysically prior. In effect, we are using an antecedent grip on normative notions to help explicate semantic notions. Of course it remains open to a Quinean opponent to reject normative notions entirely. I discuss opposition of that sort later in the paper.

A fifth worry for Carnap's account is that subjects might change their mind about a possible case without a change of meaning. Here, one can respond by appealing to Generalized Scrutability as above: then judgments about a sentence are determined by a scenario specification and by ideal reasoning. If so, then if the subject is given such a specification and is reasoning ideally throughout, then there will not be room for them to change their mind in this way. Changes of mind about a fully specified scenario will always involve either a failure of ideal reasoning or a change in meaning. I will return to this issue later.

The model we then reach is something like the following. The (primary) intension of an expression for a subject is a function that maps scenarios to extensions. Given a sentence S and a scenario w specified by a set of sentences D , the intension of S is true at w if S is scrutable from D , false at w if $\neg S$ is scrutable from D , and so on. In effect, the intension of S maps a scenario w to what a subject should ideally judge about the truth of S under the supposition that w is actual (or the supposition that all sentences in D are true). Likewise, we can think of the intension of a subsentential expression E as mapping a scenario w to what a subject should ideally judge to be the extension of E under the supposition that w is actual. This is not a perfect definition, but it is good enough for our purposes. This remains very much in the spirit of Carnap's definition, although the invocation of rationality makes it a sort of normative version of Carnap's account.

Importantly, we can use this account to give a version of the Carnapian response to Quine's arguments given in the previous section. Conceptual change (of the relevant sort) will occur precisely when an expression's primary intension changes across time. This will happen precisely when the subject's dispositions to judge the expression's extension in a possible case (given ideal reasoning) changes. As in the last section, we can find cases of holding-true where the dispositions change in this way, and cases where they do not. What matters is that we have a principled distinction.

A residual issue concerns the meaning of the basic vocabulary. If cases are specified in this vocabulary, then we need to ensure that the basic vocabulary does not change in meaning throughout the process. If we do not require this, the resulting condition for meaning change will be inadequate: a subject's dispositions to judge that S obtains with respect to a case specified by D might change over time, not because the meaning of S changes but because the meaning of terms

in D change. If we require this, however, then it appears that we need some further criterion for meaning change in the basic vocabulary items used in D , as the dispositional method would yield trivial results here. So it appears that the dispositional method for determining meaning change, even when idealized, is incomplete.³

A second residual issue concerns the role of the a priori in characterizing this account. I have so far been vague about whether intensions are defined using conditional or a priori scrutability. For many purposes, it is most natural to appeal to the latter: the primary intension of a sentence S is true at a scenario w is true iff a material conditional ‘If D , then S ’ is a priori, where D is a canonical specification of S . If we appeal to a priori scrutability, however, then we have arrived at a principled distinction only by helping ourselves to the contested notion of apriority along the way.

As before, it is not clear how bad these residual problems are. One might still see the intensional analysis as demonstrating that the Quinean phenomena of holding-true and revisability are quite compatible with the intensional framework and have no power to refute it. Even if one has to assume some independent grip on the notion of apriority, and on the meaning of expressions in the basic vocabulary, one can still use the framework to provide a reasonably enlightening analysis of relevant cases. Still, we have not broken out of the Quinean circle. It would be nice to be able to characterize the relevant distinctions without such a direct appeal to the contested notions.

I think that such a characterization can be found. At least for the purposes of answering Quine, we can define intensions in terms of conditional scrutability, rather than in terms of a priori scrutability. For example, one can say that the intension of a sentence S is true at a scenario w , for a subject, if $cr^*(S|D)$ is 1 for that subject, where D is a canonical specification of w .

If we do this, then we will have a principled criterion for conceptual change that does not appeal to apriority. On this criterion, a subject’s intension for S will change between t_1 and t_2 iff there is a scenario w with canonical specification D such that $cr^*(S|D)$ changes from 1 to 0 or vice versa. One could then run the arguments of the previous section once again using this notion. This will provide a reply to Quine’s challenge that gets around the second residual issue above

³This objection is related to Quine’s argument from the indeterminacy of translation in *Word and Object*. Quine took Carnap’s account to be a serious challenge to his arguments in “Two Dogmas”, and the indeterminacy argument can be seen in part as a response to it. Here, Quine argues that no dispositional analysis can settle facts about meaning, because multiple assignments of reference will always be compatible with a subject’s behavioral dispositions. This applies even to Carnap’s account, if we allow multiple potential assignments of reference to the basic vocabulary. In effect, Carnap’s account assumes that the meaning of the basic vocabulary is fixed, but it is not clear why such an assumption is legitimate, and it is not clear how this meaning might itself be grounded in dispositional facts.

(regarding apriority), though it may still be subject to a version of the first issue (regarding the basic vocabulary).

At this point, however, I think an alternative analysis involving conditional probability is available. This analysis is closely related to the one just mentioned, and is a descendant of the Carnapian analysis in the previous section, but it does not require this semantic apparatus or the full scrutability framework. Instead of appealing to possible cases and intensions, it proceeds using only standard Bayesian considerations about evidence and updating. In addition to the advantage of familiarity, this approach has other significant advantages in responding to Quine's challenge. By avoiding the need for canonical specifications of complete possible scenarios, it avoids the large idealization needed to handle enormous specifications. As I discuss in section 8, it also has the potential to avoid or minimize the residual issues about apriority and the basic vocabulary discussed above.

6 A Bayesian analysis of holding-true

Let us assume a standard Bayesian model, on which sentences are associated with unconditional and conditional credences for subjects at times. That is, for a given subject and a given time, a sentence S will be associated with an unconditional credence $cr(S)$, and a pair of sentences S and T will be associated with conditional credence $cr(S|T)$. (These ordinary credences $cr(S|T)$ should be distinguished from the idealized rational credences $cr'(S|T)$ and $cr^*(S|T)$ defined in chapter 2.) Credences are standardly taken to be real numbers between 0 and 1, but for our purposes exactitude is not required. It is enough that some credences be high and others low.

I will also assume a version of the principle of conditionalization: if a subject has credence $cr_1(S|E)$ at t_1 , and acquires total evidence specified by the evidence sentence E at between t_1 and t_2 , then the subject's credence $cr_2(S)$ at t_2 should be equal to $cr_1(S|E)$. I will give a more precise version of this principle below. I will discuss evidence sentences further later in this chapter, but for now we can think of them either as specifying that certain experiences obtain, or as specifying that certain observable states of affairs obtain.⁴

We can start with a typical case whereby an apparently synthetic sentence is held true in face of apparently countervailing evidence, by appeal to appropriate ancillary theses. As in section 1, suppose that at t_1 , Fred asserts 'All bachelors are untidy'. At t_2 , Fred acquires evidence indicating

⁴The arguments I present here can also be run using the principle of Jeffrey conditionalization (Jeffrey 1984), which allows conditionalization on evidence of which a subject is not certain.

that there is a tidy, unmarried 25-year old man, and responds by denying that the man is a bachelor, as bachelors must be over 30.

Let B be ‘All bachelors are untidy’, and let E be Fred’s total relevant evidence acquired between t_1 and t_2 . Let $cr_1(S)$ stand for Fred’s credence in S at t_1 , and $cr_2(S)$ stand for Fred’s credence in S at t_2 . Then $cr_1(B)$ and $cr_2(B)$ are both high.

The crucial question is: What is $cr_1(B|E)$, Fred’s conditional credence in B given E at t_1 , before Fred acquires the evidence in question?

If $cr_1(B|E)$ is high, then Fred’s judgment at t_2 reflects a conditional credence that he already had at t_1 . In this case, the judgment at t_2 is *prefigured*, in a sense analogous to the sense discussed earlier. Here, Fred’s accepting B in light of E can be seen as according with the principle of conditionalization.

If $cr_1(B|E)$ is low, then Fred’s judgment at t_2 fails to reflect the conditional credence that he already had at t_1 . In this sort of case, the judgment at t_2 is *postfigured*, in a sense analogous to the sense discussed earlier. Here, Fred’s accepting B in light of E appears to violate the principle of conditionalization.

Now, on standard Bayesian assumptions, there are two central ways in which one can obtain apparent violations of conditionalization for sentences. First, this can happen when the subject is not fully rational throughout the process: perhaps at t_1 they have not thought things through properly, or at t_2 they make some sort of reasoning error. Second, the content of the key sentence B can change between t_1 and t_2 . This may happen in cases involving indexicals, which are not relevant here, or in cases of conceptual change. In these cases, it remains possible that the subject’s credences in relevant *propositions* obey conditionalization, but that their credences in associated sentences do not, because the association between sentences and propositions changes over time.⁵

⁵A potential third way that conditionalization can be violated arises on views where sentences express certain sorts of relativistic contents: for example, a view on which utterances of the sentence ‘It is raining’ always express the same temporal proposition *It is raining*, which can be true at some times and not at others. On Saturday, I might have a low conditional credence in *It is raining* given *The weather forecast says rain on Sunday*, then on Sunday I might acquire evidence that the weather forecast says rain on Sunday, resulting in high credence in *It is raining*, without irrationality. On a more standard view on which the content of ‘It is raining’ uttered at t is *It is raining at t*, this will be classified as a change in content, but on the temporal view the content stays the same. For present purposes, we can either count these as changes in content in an extended sense, or we can require in principle (CS) that the content in question is non-relativistic content.

A potential fourth sort of violation arises from cases of self-doubt (e.g. the Shangri-La case of Arntzenius 2003). One could handle these cases by invoking insulated idealizations in the relevant notion of rationality, or simply by noting that the relevant sort of self-doubt is not playing a role in paradigmatic Quinean cases of revisability and holding-true.

We might formulate this as a version of the principle of conditionalization for sentences, making explicit the requirement of meaning constancy that was left implicit in the last chapter:

(CS) If a subject is fully rational, and if the subject acquires total evidence specified by E between t_1 and t_2 , and if the content of sentence S does not change between t_1 and t_2 , then $cr_2(S) = cr_1(S|E)$.

Perhaps the most familiar version of the principle of conditionalization is cast in terms of propositions: if a fully rational subject acquires total evidence specified by proposition e between t_1 and t_2 , then $cr_2(p) = cr_1(p|e)$. (CS) follows from this claim in conjunction with the plausible claims that when sentence S expresses proposition p for a subject at that time, $cr(S) = cr(p)$ at that time, and that the content of a sentence is the proposition it expresses.

It follows that *if* Fred in the postfigured case above is fully rational, then this is a case of conceptual change. Of course it might be that Fred is not fully rational, but this is of no help for Quine. It is unremarkable that irrational subjects might hold on to any sentence or reject any sentence, and this observation has no consequences regarding analyticity or apriority. For Quine's observations about revisability and holding-true to have any bite, rational subjects are required. So we may as well assume that Fred is fully rational.

If we assume that the relevant subjects are fully rational, we now have a principled criterion for conceptual change in a case of holding-true. Suppose that our subjects accepts S at t_1 , acquires apparently countervailing evidence E between t_1 and t_2 , and continues to accept S at t_2 . Then we can say

- (i) If $cr_1(S|E)$ is low, this is a case of conceptual change.
- (ii) If $cr_1(S|E)$ is high, this need not be a case of conceptual change.

One can now ask: is it true that a subject can hold on to any given sentence S come what may, in light of any evidence, *without irrationality or conceptual change*? By this analysis, this claim requires that for any given sentence S and any possible evidence E , $cr(S|E)$ is high (or at least is not low). But this claim is obviously false. For a rational subjects and most sentences (including most paradigmatic empirical sentences), there will be evidence sentences E such that $cr(S|E)$ is low.

The moral here is that in the general case, Quinean holding-true-come-what-may requires widespread violation of conditionalization, which requires irrationality or conceptual change. But

the fact that an irrational subject might reject a sentence is no evidence that it is not analytic or a priori,⁶ and the fact that a subject might reject a sentence after conceptual change is no evidence that it is not originally analytic or a priori. So Quine's argument from holding-true fails.

7 A Bayesian analysis of revisability

For our central example of revisability, we can use a familiar case from Putnam (1962). Let C be 'All cats are animals'. This might seem paradigmatically analytic or a priori. But let E specify evidence confirming that the furry, apparently feline creatures that inhabit our houses are actually remote-controlled robots from Mars, while the other creatures that we see are all organic. Putnam argues that if we discovered that E obtains, we would reject C . So let us suppose that Sarah accepts C at t_1 , acquires total evidence as specified by E , and rejects C at t_2 .

Here, the diagnostic question is: What is Sarah's initial conditional probability $cr_1(C|E)$?

If $cr_1(C|E)$ is low, then Sarah's judgment at t_2 reflects a conditional credence that she already had at t_1 . In this case, the judgment at t_2 is *prefigured*. Here, Sarah's accepting C in light of E can be seen as according with the principle of conditionalization.

If $cr_2(C|E)$ is high, then Sarah's judgment at t_2 fails to reflect the conditional credence that she already had at t_1 . In this sort of case, the judgment at t_2 is *postfigured*. Here, Sarah's accepting C in light of E appears to violate the principle of conditionalization.

For exactly the reasons given before, the postfigured case requires either that Sarah is not fully rational, or that her use of C undergoes conceptual change between t_1 and t_2 . Cases of this sort are of no help to Quine. Again, the fact that an irrational subject might reject a sentence is no evidence that it is not analytic or a priori, and the fact that a subject might reject a sentence after conceptual change is no evidence that it is not originally analytic or a priori.

For Quine's argument to succeed, he needs to exclude cases of this sort. That is, he needs to make the case that any sentence can in principle be rationally revised without a violation of conditionalization. This requires that for all rational subjects and for all sentences S , there exists an evidence sentence E such that $cr(S|E)$ is low.

This claim is not so obviously false as the corresponding claim about holding true come what may. For this reason, one might regard the argument from revisability as a stronger argument than

⁶Perhaps there are certain strong conceptions of analyticity on which an analytic sentence cannot be rejected by any subject, rational or irrational. But I do not think that these conceptions are standard, and in any case no such constraint applies to apriority.

the argument from holding-true. Indeed, supporters of Quine such as Putnam (1962) and Harman (1994) have concentrated on the argument from revisability, and have made claims not far from the claim in question.

Still, it is not clear just what the grounds are for accepting the key claim. At this point, a number of observations can be made.

First, Quine's official grounds for the revisability claim involves involves the ability to revise ancillary claims when necessary. These grounds are the same as for the holding-true claim, and it is clear that Quine sees the two as continuous. These grounds suggest that *after* obtaining evidence, a subject could use these features to revise a given sentence. But we have seen that revisions of this sort typically involve violations of conditionalization. These grounds do very little to suggest that *before* acquiring the relevant evidence, a subject's conditional credence $cr(C|E)$ will be low.

Second, it is certainly the case that almost any claim could be rationally rejected given testimony of an apparent epistemic superior. But this claim has no bearing on apriority: that a claim could be rejected in *this* way is no evidence that it is not a priori. It also does not establish that any claim is revisable under ideal reflection, as it is far from clear that this sort of revisability applies to ideally rational thinkers. Perhaps their grounds for accepting a mathematical claim, for example, will always defeat any evidence concerning an apparent epistemic superior.

Third, even if this sort of consideration applies to many apparent cases of a priori truths, there are a number of cases on which it has no purchase. Some such cases include *material conditionals* of the form 'If D , then S ' (like those discussed in the previous section), where D is a lengthy specification of an arbitrary scenario, and where S is a sentence such as 'Water is H_2O ' such that $cr(S|D)$ is high. Assuming a fully rational subject, it follows that $cr(D \rightarrow S|D)$ is high, so that $cr(D \rightarrow S)$ is also high. We can stipulate that D includes or entails a full specification of evidence that obtains in the scenario, so that D entails E for any evidence sentence E that obtains in the scenario and D entails $\neg E$ otherwise (setting vagueness aside).⁷ Now a quick two-case argument suggests that no evidence E could lead us to rationally reject $D \rightarrow S$. First case: if E does not obtain in the scenario, then D entails $\neg E$. In this case, $cr(\neg D|E) = 1$, so $cr(D \rightarrow S|E) = 1$.

⁶It may be useful to distinguish a pragmatist reading of the arguments in "Two Dogmas", which stresses the freedom to adjust ancillary hypotheses as one chooses, and an empiricist reading, which stresses the role of unexpected evidence in driving us to revise our beliefs. Roughly, where the pragmatist reading turns on the claim that one *may* accept or reject certain statements, the empiricist reading turns on the claim that one *should* (or perhaps that one *would*). A pragmatist reading will put equal weight on the argument from holding-true and revisability, while an empiricist reading will put more weight on the latter. I think that the pragmatist strand is more central in Quine's text, but the empiricist strand has been more influential among later Quineans.

Second case: if E obtains in the scenario, then D entails E . Now $cr(D \rightarrow S|E)$ must lie between $cr(D \rightarrow S|E \& \neg D)$ and $cr(D \rightarrow S|E \& D)$. But the former is 1, and the latter is just $cr(D \rightarrow S|D)$, which we have seen is high. So $cr(D \rightarrow S|E)$ is high. Putting the two cases together, $cr(D \rightarrow S|E)$ is high for all E . Importantly, material conditionals very much like these are the a priori truths that are most important in the scrutability framework.

Fourth, once one notes that this argument allows some truths S such that $cr(S|E)$ is high for all E , then it is clear that there is no longer a sound principled argument that for all S , there is an E such that $cr(S|E)$ is low. As a result, we may expect to find many more exceptions to this claim. Indeed, many Quineans have conceded such objections, for example in the domains of mathematics and logic, and there is no reason not to expect many more.

Fifth, it is worth stressing that even if this line of argument succeeded, it would be much more conservative than Quine's original line. It leads naturally to a view on which there is an analytic/synthetic distinction. At worst, it would be the case that most or all sentences previously regarded as analytic (a priori), such as 'All cats are animals', will be reconstrued as synthetic (a posteriori).⁸ But one could still use the current framework to characterize intensions, once one acknowledges that the intensions for sentences such as 'All cats are animals' will be false at some scenarios. One will still have a principled distinction between cases that involve conceptual change and cases that do not. In this way, the advocate of analyticity, apriority, and conceptual analysis will have much of what they want.

In any case, the Bayesian analysis has given us what we wanted: a principled criterion for identifying cases of conceptual change. It has only given us a sufficient condition, rather than a necessary and sufficient condition, but this is good enough for our purposes. With this analysis in hand, it is clear that Quine's arguments from revisability and holding-true fail.

8 Quinean Objections

(1) *The Bayesian analysis begs the question.*

⁷One can understand entailment here in a variety of ways. For present purposes we need only the claim that if D entails E , then $cr(E|D) = 1$ and $cr(\neg D|\neg E) = 1$ for a fully rational subject.

⁸Here we can distinguish between radical Quineans, who hold that there is no analytic/synthetic distinction, and moderate Quineans, who hold that there is such a distinction but that very few sentences are analytic. If the analysis given here is right, we would expect this distinction to correlate with the distinction between pragmatist and empiricist Quineans in footnote 8, and this seems to be what we find in practice. For example, Quine in "Two Dogmas" takes the pragmatist and radical lines, while Putnam (1962) takes the empiricist and moderate lines.

It might be suggested that the Bayesian principle (CS) that I have appealed to simply assumes a notion of conceptual change without argument, and therefore begs the question against the Quinean skeptic about this notion. I do not think that this is quite right. (CS) is itself a consequence of the principle of conditionalization for propositions and of two other weak assumptions, none of which say anything about conceptual change. Still, this line of argument assumes a notion of proposition, about which a Quinean might be skeptical.

Now, Quine's doubts about propositions have been much less influential than his doubts about the analytic/synthetic distinction. But in any case, I think it is clear that Bayesian accounts of confirmation require *either* something like propositions or something like the notion of conceptual change to get off the ground. Bayesian credences will be assigned either to abstract entities such as propositions, events, or sets, to linguistic items such as sentences, or to mental items such as beliefs. If we take the first route, then we can use these entities just as we used propositions to ground a notion of conceptual change. And if we take the second or the third routes, we need to require something like conceptual constancy in order to avoid counterexamples to principles such as conditionalization.

Of course a Quinean might simply reject Bayesianism altogether, along with the associated principle of conditionalization. This would seem rash, however, as Bayesianism is an extremely successful theory with widespread empirical applications. So by a Quinean's own lights, it is hard to reject it. Furthermore, even if one rejects Bayesianism, a successor theory is likely to have corresponding principles of diachronic rationality, governing how beliefs should be updated over time in response to evidence. And precisely the same issues will arise for these principles: if they apply to abstract items we can use these to define conceptual change, and if they apply to linguistic items or mental items, we will require a notion of conceptual change.⁹

⁹A related objection is that the very idea of a credence or a conditional credence presupposes conceptual constancy. After all, one's credence associated with a sentence is arguably determined by one's dispositions to make certain judgments and decisions involving the sentence: for example, the odds one would take on a bet on that sentence if it were offered. But in considering such dispositions, we have to assume that the meaning of the sentence stays constant from the initial moment to the bet. Likewise, a conditional credence $cr(M|E)$ is arguably determined by one's dispositions to make judgments about M conditional on the supposition of E . This requires conceptual constancy: if meaning changes between the initial moment and the judgment, a high initial credence might go with a negative judgment. Still, any conceptual constancy needed here is at best very local, within an episode of consciousness. In any case, if it turns out that the notion of apriority is as secure as the notion of a credence, so that the Quinean can reject the former only by rejecting the latter, that should be good enough for the defender of apriority. Thanks to Ned Block and Kelvin McQueen for discussion here.

I think the deeper moral is that there is a constitutive link between *rational inference* and conceptual constancy. Issues such as those floated here will arise for any principle of diachronic rationality at all. If it is a principle that from A and $A \rightarrow B$ one should infer B , and if the premises and conclusions here are sentences or mental items, then to avoid obvious counterexamples, the principle should require that A and B have the same meaning on each occasion when they occur. And if the principle applies to abstract objects such as propositions, these can themselves be used to define conceptual change. So if we are not skeptics about principles of diachronic rationality, a notion of conceptual change will be hard to avoid.

(2) *Rationality presupposes apriority.*

It might be suggested that in appealing to the notion of rationality, the notion of apriority is tacitly smuggled in. For example, someone might hold that all principles of rational inference depend on underlying principles about the a priori: for example, perhaps an inference from some premises to a conclusion is rational precisely if it is a priori that if the premises obtain, the conclusion is likely obtain. Or perhaps the distinctive idealization made by the Bayesian involves some tacit assumptions about the a priori. For example, perhaps the Bayesian requirement that rational subjects should have credence 1 in logical truths depends in some way on the belief that logical truths are a priori. If so, the appeal to rational principles here presupposes one of the key notions at issue.

The reply here is straightforward. Whether or not the objector is correct that rationality depends in some way on apriority, the appeal to rationality is innocuous in the current dialectical context. The relevant class of opponents are those who accept the notion of rationality, but who question the notion of apriority. My argument is intended to establish that *if* one accepts certain principles concerning rationality, then one should reject Quine's argument against the a priori. If this objector is correct, then the opponent should either give up on the principles concerning rationality, or accept the notion of the a priori. Either outcome is sufficient for my purposes. I am happy to concede that if an opponent rejects the notion of rationality, or rejects all relevant principles of diachronic rationality, then the current argument has no purchase against them.

It is also worth noting that the principles of rationality that I appeal to are principles that many or most opponents of the a priori accept. Conditionalization has no obvious connection to the a priori, for example. I do not know whether the special status that the Bayesian gives to logical truths has a special connection to the a priori, but in any case this status plays no role in my argument. That is, the argument does not require the Bayesian claim that rationality requires credence 1 in logical truths. In fact, the picture I have sketched appears to be compatible with a view

on which logical truths deserve rational credence less than 1, and on which they can be revised given relevant evidence. All that is required is that such a revision should obey conditionalization. Nothing here smuggles in any obvious presuppositions about the a priori.¹⁰

(3) *A principled line between conceptual change and irrationality cannot be drawn.*

A Quinean may suggest that our concept of rationality is not fully determinate, and that as a result a clear division between cases of irrationality and cases of conceptual change cannot be found. Some hard cases, such as revising logic in light of quantum mechanics, are not easily classified as either.

However, my reply to Quine's argument does not require drawing a line here. It suffices for the purposes of the argument that the violations of conditionalization involve *either* irrationality or conceptual change, and we do not have to classify these violations further. In any case, as long as there are clear cases of rational judgment, the existence of unclear cases entails at worst a vague distinction, not a nonexistent distinction.

(4) *The argument requires constancy in evidence sentences.*

Recall the first residual issue for the framework of intensions discussed earlier: the framework assumes conceptual constancy in the base vocabulary, so that the framework cannot explain this conceptual constancy. One might think that an analogous issue arises here, with respect to the *evidential* vocabulary: the vocabulary used to specify evidence sentences such as E . After all, conditionalization concerns what to do when one has a certain credence $cr(S|E)$ and then learns E . The conditional credence is in part an attitude to a sentence E , and what one learns is also a sentence E . One might think it is required that the sentence have the same meaning on both occasions. If so, then any apparent failures of conditionalization in a rational subject could be blamed on a change in the meaning of terms in E , instead of a change in the meaning of terms in S , and it is not clear that we have a principled way to choose.

As it stands, this picture is not quite right. Learning E does not typically involve the *sentence* E at all. Perhaps if learning was always by testimony, and if E is a sentence used in testimony, then the issue arises. But for our purposes we can assume that the relevant learning is by perception or by introspection. Here, E will be a sentence characterizing the evidence that one learns, and the learning process need not involve this sentence at all. So there is no use of E at t_2 that needs to be aligned with the use of E at t_1 . At best we need to require that E as used at t_1 correctly applies to

¹⁰Of course, my argument appeals to logical claims at various points, but this does not require that logical truths are a priori, or that they are unrevisable. It merely requires that they are true. Likewise, my argument does not require that the principle of conditionalization is itself a priori or that it is unrevisable. It simply requires that the principle is true.

the evidence acquired at t_2 . But this is a much weaker requirement, concerning only the extension of E as used at t_1 , with no role for any use of E at t_2 .

Still, it can be argued that acquiring evidence requires having certain *attitudes* to the evidence. For example, the rationality of Bayesian conditionalization on new experiences arguably requires not just that one has the experiences, but that one is certain that one has them. If so, one might suggest that the framework tacitly requires that at t_2 , one is certain of the evidence statement E (which says that certain experiences obtain). This issue is starker in alternative frameworks such as Jeffrey conditionalization, which accommodate uncertainty about evidence by giving an explicit role to one's credence in evidence statements such as E at t_2 . Does this not require some sort of constancy in the meaning of E after all?

The issue is delicate. For the reasons given above, I think that the sentence E as used at t_2 plays no essential role here. However, it is arguably the case that subjects must be certain of (or have other appropriate attitudes to) certain evidential *propositions*, such as the proposition that certain experiences obtain, which were expressed by E at t_1 . Or without invoking propositions: the subject must be certain that the relevant evidence obtains (that they are having certain experiences, say), where this is the same evidence concerning which they had conditional credences at t_1 . Without this alignment, one could always respond to an apparent failure of conditionalization by saying that although the subject's initial credence was conditional on evidence e obtaining, and although evidence e later obtained, the subject in fact became certain that some *other* evidence e' obtains. If this were so, there would be no violation of conditionalization (the subject would not acquire the evidence e), and there would arguably be no irrationality.

This requirement of alignment provides some room for the Quinean to maneuver, but the room is extremely limited. To eliminate this room altogether, we need only suppose that we have a grip on what it is for a subject to accept or suppose that certain evidence obtains. With this much granted, we can simply stipulate that for our purposes, the conditional credences $cr(S|E)$ relevant at t_1 are credences in S conditional on the evidence that is actually obtained at t_2 . This removes any loophole, and does so without making any assumptions about constancy in the meaning of language across time. At most, we have to assume an understanding of certain beliefs and suppositions about evidence.

The required assumptions can be made even smaller by noting that for our purposes, evidence can be limited to experiences or at least to observational states of affairs. While there is a sense in which non-observational states can serve as evidence for other claims, it is plausible that knowledge of these states of affairs is itself grounded in evidence concerning experiential

or observational matters. On a Bayesian view, our credences in these states of affairs must then match those determined by conditionalization on experiential or observational matters. I think it is also plausible that credences in observational states of affairs should themselves match those determined by conditionalization on experiential matters.¹¹ If the latter claim is granted, then for present purposes we can restrict the relevant evidence in cases of revisability and holding-true to experiential states. And even without it, we can restrict the relevant evidence to observational states. So to answer the Quinean worry, we need only suppose that we have a grip on what it is for a subject to accept or suppose that certain experiential or observational states of affairs obtain. And this is something that Quine's arguments in "Two Dogmas" do not give us any reason to doubt.

The upshot of all this is that the residual issues about a base vocabulary are not eliminated altogether on a Bayesian approach, but they are minimized, in a way that brings out the severe costs of the Quinean position. A Quinean who rejects the notions of analyticity and apriority along present lines must also insist that there is no objective fact of the matter about whether a subject accepts or supposes that a given observational state obtains. This view would presumably go along with a generalized skepticism about the contents of thought, perhaps in the spirit of Quine's skepticism about meaning developed in his arguments concerning radical translation. It would likewise require a certain skepticism about diachronic rationality, for reasons discussed earlier.

Quine himself argues both for skepticism about meaning (in *Word and Object*) and for a sort of skepticism about norms of rationality (in "Epistemology Naturalized"). But few have been prepared to follow him here, and even those who sympathize with the Quine of "Two Dogmas" have tended to reject these later views. Of course Quine's arguments for these views deserve attention in their own right, but it is clear that the arguments in "Two Dogmas" do not provide much direct support for them. Still, the current analysis suggests a deep linkage between these views. Defending the arguments of "Two Dogmas" against a certain sort of appeal to conceptual change leads naturally to skepticism about diachronic rational principles and about the content of language and thought. Contrapositively, once even minimal claims about rationality and about thought are accepted, the arguments I have considered against analyticity and apriority dissolve.

¹¹For example, if one is fully rational, one's credence that there is a red square in front of one, should match one's antecedent conditional credence that there is a red square in front of one given that one is having an experience as of a red square. (If norms of rationality do not ensure certainty about the experiences one is having, one can move to a Jeffrey-conditionalization analog.) Theses of this sort have been denied by some dogmatists about perception (e.g. Pryor 2006), and might also be denied by some who think that perceptual knowledge is more secure than introspective knowledge (eg. Schwitzgebel 2008).

(5) *There can be rational revision by resetting priors.*

Quineans of a pragmatist stripe often appeal to the underdetermination of theory by evidence: multiple theories are consistent with the same evidence, and we have considerable latitude in choosing between them. In the Bayesian framework, where theory is determined by evidence along with prior probabilities, this underdetermination comes to underdetermination of probabilities that are prior to any evidence. This underdetermination yields a potential way that fully rational subjects might violate conditionalization without conceptual change.

The relevant method here is that of *resetting priors*. This method stems from the observation that most Bayesians allow that there is some flexibility in one's ultimate priors: the prior probabilities that a subject should have before acquiring any empirical evidence. (Of course these priors are something of a fiction.) For example, on Carnap's framework for inductive logic, equally rational subjects may have different values for λ , the parameter that guides how quickly the subjects adjust their beliefs in light of inductive evidence, and this difference can be traced to a difference in ultimate priors. Two such subjects might acquire exactly the same evidence over time, while being led to quite different posterior probabilities. If G is the thesis that a certain sort of global warming is occurring, for example, one subject might be led to a high credence in G , while another might be led to a low credence in G .

Now, a subject with a high credence in G might reflect and observe that their high credence is tracable entirely to the value of λ in their ultimate priors, and that this value was quite arbitrary. They may note that it would have been equally rational to start with a lower value of λ , and to end up with a lower credence in G . At this point, a bold subject might choose to change their credences wholesale. At least if they have a good enough record of their evidence, they can "unwind" back to the ultimate priors, reset λ to a lower value, and reintegrate all the evidence by conditionalization. The subject will end up with a new set of credences, including (among many other differences) a much lower value for G .

A Quinean might suggest that there is nothing irrational about doing this, and that this method might be exploited in order that a subject can hold on to almost any sentence "come what may" and to revise almost any sentence. After all, for most non-observational empirical sentences S and most paths of evidence, there is some ultimate prior that will lead to a high credence in S , and some ultimate prior that will lead to a low credence in S . None of this requires conceptual change. So violations of conditionalization in a rational subject do not provide a sufficient condition for conceptual change, after all.

This position requires a rejection or at least a revision of orthodox Bayesianism. On the ortho-

dox view, conditionalization is a constraint on diachronic rationality, and this sort of revision will be irrational. Furthermore, the view tends to lead to an anything-goes view of rational belief. If there are no constraints on ultimate priors, the view entails that at any moment, if $cr(p) < 1$, then one's credence can be rationally revised so that $cr(p)$ is arbitrarily close to zero. And even if there are constraints on ultimate priors, these constraints must be weak enough to vindicate the large violations of conditionalization that the Quinean argument requires, leading naturally to a view on which most beliefs can be rationally revised at any moment into disbelief. Given this much, it is not easy to see how my beliefs can constitute knowledge at all.¹²

Furthermore, it is far from clear that all beliefs can be revised in this way. For example, given that logical beliefs, mathematical beliefs, and evidential statements are constrained to have credence 1, this method will not yield revisability for these beliefs. More generally, there is not much reason to hold that it will yield revisions to those beliefs usually classified as a priori ('All bachelors are unmarried', say), most of which do not appear to depend on ultimate priors. So this response is weakest where it needs to be strongest.

Most fundamentally: as long as we have a conceptual distinction between cases in which beliefs are revised by this process and cases in which they are not, we still have enough to draw a distinction between those violations of conditionalization that involve conceptual change and those that do not. The Quinean will have to insist that we do not have a grip on this conceptual distinction, so that there is no distinction to be drawn between cases of resetting priors and cases of conceptual change. I do not think there is much reason to accept this. Furthermore, even if this line were accepted, it would once again lead to an across-the-board skepticism about principles of belief updating and other forms of diachronic rationality. So if principles of diachronic rationality are allowed at all—even the liberal principles suggested by the current approach—then the distinction between conceptual constancy and conceptual change remains intact.

(6) *Subjects need not have conditional credences.*

It might be objected that the Bayesian analysis requires the assumption that for every sentence S used by a subject and every possible evidence sentence E , the subject has a conditional credence $cr(S|E)$. But this is an unrealistic idealizing assumption.

In response: The idealization is not enormous. For most S and most E , the subject will have some relevant dispositions involving S and E , for example involving their willingness to accept

¹²In addition, this method is a sort of belief revision that is not driven by evidence at all. So although this line of reasoning is perhaps the best way of preserving the pragmatist reading of Quine's arguments in light of the present analysis, it does not sit easily with the more influential empiricist reading.

various bets involving S and E . In many cases, these dispositions will line up in a clear enough way that $cr(S|E)$ will be high. In other cases, they will line up in a clear enough way that $cr(S|E)$ will be low. In other cases, the dispositions may be enough of a mix that it is hard to say.

A Quinean might suggest that if $cr(S|E)$ is indeterminate in this way, and the subject later rejects S upon learning E , this should not count as a violation of conditionalization. If so, they might then suggest that for any S , there is some E such that $cr(S|E)$ is indeterminate in this way, and such that the subject could later reject S on learning E without violating conditionalization. Perhaps this sort of revisability is enough for their purposes?¹³

I do not think that this is enough, however. Cases of this sort seem to turn essentially on the subject's not being fully rational. If the subject is fully rational, then the subject's dispositions to accept S on *supposing* E and on *learning* E should be the same, assuming no conceptual change. That is, if a fully rational subject rejects S on learning E and thinking things through, then if the subject were to have been initially presented with the *supposition* that E and had thought things through, the subject should have rejected S conditional on that supposition. To fail to meet this condition is a failure of full rationality, just as is an ordinary violation of conditionalization. So at best the Quinean has presented us with a kind of revisability that can only be exploited by subjects who are less than fully rational. Like the sort of revisability that can be exploited only by irrational subjects, this sort of revisability has no bearing on matters of apriority.

9 Conclusion

Quine is right that any statement can be held true come what may, and that no statement is immune to revision. But as Grice and Strawson observe, these phenomena are quite compatible with a robust analytic/synthetic distinction and a robust notion of meaning. A Bayesian analysis reveals that Quine is not right that any statement can be held true come what may *without conceptual change or irrationality*, and likewise for revision. We can pin down the distinction between cases that involving conceptual change and cases that do not using either the scrutability framework or Bayesian analysis.

¹³It is especially likely that ordinary subjects will lack credences $cr(S|D)$ involving the scenario specifications D discussed earlier, due to the enormous size of these specifications. This observation does not affect the use of conditional credences involving D to define intensions, as these credences used there are always idealized rational credences $cr'(S|D)$, for which the current issue does not arise. And where nonidealized credences are concerned, these cases will not yield cases of revisability along the lines in the text, because the subject will be incapable of learning that D .

The method of intensions characterizes intensions in terms of scrutability relations, and uses them to draw a distinction between cases that involve conceptual change and cases that do not. When intensions are grounded in a priori scrutability, this method assumes the notion of apriority, so it does not provide an independent grounding for that notion. Still, it shows how a framework involving apriority can accommodate all of Quine's data. And for the same reasons that most philosophers reject Quine's arguments in sections 1-4 of "Two Dogmas", no independent grounding is required.

The Bayesian analysis takes things a step further and defends the a priori on partly independent grounds. This analysis assumes the notion of conditional probability and the normative notion of rationality to provide conditions for conceptual change, but it does not assume the notion of apriority. In effect, constitutive connections between rational inference and conceptual change are used to make inroads into the Quinean circle.

The conclusion should not be too strong. While I have responded to Quine's arguments against the a priori and the analytic, I have not provided a positive argument for the analytic/synthetic distinction or the a priori/a posteriori distinctions, and I have not tried to ground these notions in wholly independent terms.

One might be tempted to take things a step further still, attempting to define apriority in terms of conditional probability and rationality. For example, one might suggest that a sentence S is a priori for a subject precisely when the ideal conditional probability $c_{r^*}(S|D)$ is 1 for all scenario specifications D . But there will be residual issues. For a start, it is not clear that one can define the class of scenario specifications without using the notion of apriority.¹⁴ So much more would need to be said here.

Still, we have seen that these notions can at least help us in diagnosing issues regarding meaning, conceptual change, and the a priori. And we have seen enough to suggest that Quine's arguments in the final section of "Two Dogmas of Empiricism" do not threaten the distinction between the analytic and the synthetic, or the distinction between the a priori and the a posteriori.

¹⁴One will also need to appeal to an insulated idealization (chapter 2) to handle cases of self-doubt: e.g., the case where S is a mathematical truth, and D specifies a scenario in which one is a poor mathematical reasoner.

Ninth Excursus: Scrutability and Conceptual Dynamics

It is sometimes argued (e.g. Wilson 2006) that “classical” models of concepts, stemming from early analytic philosophy and the logical empiricists, are inadequate because they cannot properly accommodate conceptual dynamics: the way that the concepts associated with our expressions develop and change over time. The scrutability framework is at least a relative of these classical models, and similar charges have occasionally been brought against it (e.g. by Block and Stalnaker 1999, Yablo 2002, and Schroeter 2006). The analysis given in this chapter can be used to address these arguments. I will discuss a number of cases of conceptual dynamics that might be thought to pose problems for scrutability.

There is no question that conceptual development is rife and intricate. A paradigmatic case, reflecting a phenomenon common with scientific theories, is given in Joseph Camp’s “The Ballad of Clyde the Moose”. Fred hears something going shup-shup in the night, and introduces the name ‘Clyde’ for whatever it is. He sets out to track down Clyde, and finds many other strange noises and odd events that he attributes to Clyde. The Clyde theory becomes increasingly complex, until a year later Fred discovers Clyde. There is a moose who was responsible for all these events, with one exception: the moose did not go shup-shup in the night. So the sentence ‘If anyone went shup-shup in the night, it was Clyde’ (*S*), which may seem to be stipulative and a priori at the first stage, is rejected at the second stage.

Someone might suggest that although *S* initially seems to be a priori scrutable from *PQTI* (or a relevant base), it is not. The current framework handles this case by saying that ‘Clyde’ undergoes conceptual change. Because of this change, *S* is a priori at the first stage but not at the second. Suppose that *E* is the evidence acquired by Fred between the first stage and the last. At the earlier stage, Fred presumably has a high conditional credence in *S* given *E*, but after acquiring evidence *E* he rejects *S*. This is a violation of conditionalization. Given that Fred is rational, the violation suggests conceptual change. Indeed, it is plausible that an utterance of *S* at the first stage would be true, but an utterance at the second stage would be false. If so, then even without invoking the principle of conditionalization, there is certainly be conceptual change in this scenario. The case brings out the important point that conceptual change can be well-motivated by theoretical change and not at all anomolous or arbitrary. But all this is consistent with the claim that at each stage, true utterances are scrutable from *PQTI* (or an underlying base), and false utterances are not.¹⁵

There are also cases in which an original stipulation does not turn out to be false, but moves from a priori to a posteriori. For example, ‘Neptune perturbs the orbit of Uranus, if it exists’ may

initially seem a priori to Leverrier, but after acquiring much further knowledge of Neptune, the same sentence may no longer seem a priori. Suppose that D specifies a scenario in which planet X perturbs the orbit of Uranus, but planet Y plays the rest of the roles that come later to be associated with Neptune. Then if $T1$ is ‘Neptune is planet X’ and $T2$ is ‘Neptune is planet Y’, Leverrier may initially take ‘If D then $T1$ ’ to be a priori and later take ‘If D then $T2$ ’ to be a priori’. In this case, one’s conditional credences $cr(T1|D)$ will move from high to low.

This sort of case yields an indirect violation of conditionalization. Assume that E is Leverrier’s total evidence between the stages and that it is consistent with D . Then at stage 1, $cr(T1|D\&E)$ is high, but after gaining total evidence E , $cr(T1|D)$ is low. It is easy to see that this yields a violation of conditionalization. The ratio formula for conditional probabilities entails that at stage one $cr(T1|D\&E) = cr(T1\&D|E)/cr(D|E)$, which by conditionalization should be identical to $cr(T1\&D|D)$ at stage two, which is just $cr(T1|D)$.¹⁶ Assuming that Leverrier is fully rational throughout, this is also a case of conceptual change.

In “Against A Priori Reductions” (2006), Laura Schroeter discusses a hypothetical case of this sort. We first believe that the watery stuff in the oceans is a natural kind before discovering that it is wildly disjunctive. Say that W specifies a world in which the dominant watery stuff is H_2O , and S is ‘If W , then water is H_2O ’. Schroeter argues in effect that $cr(S|W)$ is high at stage one and low at stage two, because at the second stage we no longer think it important that water is a natural kind. The reasoning in the paragraph above shows that this case involves a violation of conditionalization, so that it must involve irrationality or conceptual change.

Schroeter suggests that because we refer to the same thing (a disjunctive kind) by ‘water’ throughout and the changes are natural, there is no conceptual change. She suggests that we should individuate concepts in an externalist way so that when reference changes, the concept changes, but when conditional judgments change without reference changing, it need not. I am inclined to individuate concepts differently, but the scrutability thesis and the arguments in this chapter do not depend on how we individuate concepts. If we individuate them Schroeter’s way, we will simply conclude that the a priori liaisons (including the scrutability conditionals) associated with an expression can change while the concept stays the same. Importantly, the case gives no reason

¹⁶The ratio formula cannot be used if one’s initial credence in $D\&E$ is zero. But in the relevant cases, D and E will at least be mutually consistent, and we will usually be able to coarse-grain them to raise the credence above zero while leaving the same general structure. In any case, the normative quasi-conditionalization principle that after gaining total evidence E , one’s posterior credence $cr(T1|D)$ should be equivalent to one’s prior credence $cr(T1|D\&E)$ is plausible even independent of the ratio formula.

to doubt that at each stage, when a sentence is true for a speaker, it is scrutable.

Block and Stalnaker (1999) discuss the case of ‘jade’, in which we discovered two superficially similar substances and decided that there are two sorts of jade. Here, it is plausible that even before discovering the two substances, our conditional credence in ‘There are two kinds of jade’ given D , where D specifies relevant truths about the two substances, would have been high. If a diagnosis of this sort is right, there is no problem for scrutability in these cases. But even if the diagnosis is wrong—if one’s credence in ‘There are two kinds of jade’ given D at the earlier stage was low, for example—then we can say either that this credence was not ideally rational (one had not thought the matter all the way through) and that the utterance was true, or that the low credence was rational, and that the utterance was false. On the latter hypothesis, there will have been conceptual change along the way. But on any of these scenarios, there is no problem for scrutability.

The case of solidity provides another problem case. Let D be ‘Apparently solid objects contain mostly empty space’, and let S be ‘Tables are solid’. It might be held that before discovering the truth of D , then (i) it was reasonable to hold that if D , then $\neg S$, but (ii) S was true all the same. This suggests that even though S was true, $cr'(S|PQTI)$ was low and $PQTI \rightarrow S$ was not a priori, so that both A Priori and Conditional Scrutability fail. In response, one can note that in the actual world, after discovering D , we affirmed S . This observation combined with (i) yields a violation of conditionalization, so the reasoning in this chapter suggests that given rationality throughout, either (i) is false or that there was conceptual change here. I think the former is perhaps the more plausible option: it is arguable that our core concept of solidity all along was tied to functional role (resisting penetration and so on) rather than intrinsic structure. If there was conceptual change, on the other hand, then it is natural to hold that (ii) is false: although S is true of the actual world on its later meaning, it was not true on its earlier meaning. Either way there is no problem for scrutability.

An interesting sort of opposition here invokes epistemic conservatism: roughly, the idea that we ought to hold on to existing beliefs when we can. One might suggest that epistemic conservatism can yield violations of conditionalization without irrationality or conceptual change. On this view, for a proposition p (concerning solidity, for example) and evidence e (concerning physics, for example), when one thinks that e is very unlikely, one can rationally have a low conditional credence in p given e , but upon discovering to one’s surprise that e obtains, epistemic conservatism allows that one can or should rationally retain one’s belief in p . I stand with the orthodox Bayesian view that this form of conservatism is irrational: one should either have a high conditional credence at the first stage or a low credence at the second stage.¹⁷ But this sort of

opposition should at least be noted.

Another sort of opposition in the solidity case invokes semantic conservatism: roughly, charity principles that suggest that meanings should be assigned to expressions to maximize truth. Here it might be suggested that even if $cr'(S|PQTI)$ is low at the first stage, principles of charity might override conditional judgments to dictate that S is true. The diagnosis of this view depends on whether we accept S on discovering that physics is mostly empty space. If yes (as actually happened), then we have violations of conditionalization, and this view will reduce to a version of the previous view. If no (as might happen in some hypothetical analog case), on the other hand, we have a version of radical externalism on which even knowing the relevant underlying truths about the actual world does not put one in a position to know that tables are solid. Strong enough principles of charity might indeed yield a sort of radical externalism that violates scrutability. As discussed in the previous chapter, I think that this radical externalism is implausible, but in any case this objection has little to do with conceptual change.

Some tricky cases from the history of mathematics involve conceptual development that is not tied to new empirical evidence. Newton and Leibniz talked of the limit of a series 150 years before Bolzano and Cauchy gave the now-canonical definition of a limit.¹⁸ Take the sentence S : ‘The limit of a series S_n is that value a such that for all ϵ there exists k such that for all $n > k$, $|S_n - a| < \epsilon$ ’. This sentence is true, and today it seems a priori. But was it a priori for Newton or Leibniz? This is not obvious. To get at this matter, let us suppose that Newton and Leibniz had entertained the sentence and considered whether to accept it. Perhaps the most likely outcome is that on sufficient rational reflection, they would have accepted it. If so, then S was plausibly a priori in their context. But suppose that they would have rejected it. If so, then we must say either that S did not express a truth for them, so that there was conceptual change between them and Cauchy, or that the sentence did express a truth but that they were irrational in rejecting it. Any of these outcomes seems a possibility. It might even be that the actual world was somewhat indeterminate between these possibilities, so that it is indeterminate whether S in the mouth of

¹⁷For example, an epistemic conservative will be vulnerable to a diachronic Dutch book combination of bets [see 4.3], according to which they are guaranteed to lose money in this sort of case. For example, if $cr_1(p|e) = 0.1$ and $cr_2(p) = 0.9$, then the subject will bet at 1:2 odds on $\neg p$ conditional on e at stage 1 and at 1:2 odds on p at stage 2. These bets yield a guaranteed loss if e obtains and merely break even if e does not obtain (if e does not obtain, the first bet is refunded and the second is not offered). Furthermore, unlike some other diachronic Dutch book cases (but like the version in chapter 4), this behavior from an epistemic conservative will be predictable in advance, and so can be exploited by a Dutch bookmaker.

Newton or Leibniz would have been true.

Perhaps the most complex cases are tied to the “open texture” of language, so that later developments are somehow grounded in earlier concepts without being entirely determined by them. For example, Yablo (1999; 2002) suggests that there are cases where rational reflection on qualitative information underdetermines theoretical truth, which is settled only by pragmatic factors. There can certainly be such cases: perhaps at t_1 , our credence in M is unsettled, but later one comes to accept M , because of pragmatic factors that are not merely a product of rational reflection. (I leave aside D here on the assumption that the evidence is the same throughout.) But given that the pragmatic factors are rationally underdetermined, so that one’s initial *rational* credence in M is not high, then this is best seen as a case in which it is initially indeterminate whether M expresses a truth, and M later comes to express a truth because of terminological evolution. If so, there is no problem for scrutability.¹⁹ Here the pragmatic factors can be seen as producing mild conceptual change, or conceptual precisification: an intension that is initially indeterminate at a scenario becomes determinate at that scenario.

Likewise, in *Wandering Significance* (2006), Mark Wilson discusses an enormous number of interesting cases of conceptual evolution, mainly drawn from the history of science and mathematics, in which old concepts are extended in new and unpredictable ways by new developments. I cannot analyze all these cases here, but I think that all are open to diagnoses such as the above.²⁰ In some cases, the changes in response to new cases are prefigured, in other cases, they are post-figured, and in still others the matter is indeterminate. All of these diagnoses are compatible with a priori scrutability.

I think that if these cases of conceptual evolution pose a problem for the scrutability framework, it is a problem not of incorrectness but of incompleteness. The scrutability framework

¹⁸For related discussion of this case, see Peacocke 19xx and Burge 20xx.

¹⁹It might be suggested that the earlier meaning of M is fixed by its later meaning, so that there is no conceptual change, and M is true but inscrutable at the earlier stage. The mild violation of conditionalization here strongly suggests mild conceptual change, however. There may be cases in which past meaning is fixed by future meaning via a sort of semantic deference to future users, but these cases can be handled as we handle the cases of deference discussed in chapter 6.

²⁰Wilson does not talk much about either apriority or truth (although see his pp. 617-38), so it is hard to know how his arguments apply to a thesis such as A Priori Scrutability. Concerning key questions about whether old sentences are true when applied to a new and extended case, one natural view to extract from his discussion is that there is often no determinate answer: later extensions depend on idiosyncratic developments, and verdicts about such cases are not determinately prefigured in a user’s original use of an expression. I think that this view is plausible, especially where the hardest cases are concerned, but of course widespread indeterminacy is quite compatible with the scrutability thesis.

allows us to associate an expression with an intension at any given time, and indeed to chart the way that the intension changes with time. But it does not provide any explanation of why and how concepts evolve in the way that they do. Indeed, one might think that from the point of view of the scrutability framework, as with classical models of concepts, conceptual evolution should be expected to be rare and anomalous. I think that this is not quite right: there is no reason why the framework should not be combined with a principled account of the dynamics of conceptual change. Indeed, the discussion of conceptual pluralism in chapter 9 suggests a view on which conceptual evolution and diversity is constant and ongoing, driven by various practical purposes.

Ultimately we would like a positive theory of conceptual dynamics, answering questions about when, how, and why we can expect concepts associated with an expression to change. In modeling these dynamics, the intensional framework will at least provide a useful tool. For example, the Clyde and Neptune cases above suggests some generalizations about how change in intensions tends to accompany theory change. Other generalizations might concern the relation between intensions and purposes. Intensions need not do all the work. We can be pluralists about concepts and their content. For example, we might invoke coarser-grained features of concepts to analyze what stays constant in some cases where intensions change, and finer-grained features of concepts to analyze what changes in some cases where intensions stay constant. But a positive account of conceptual dynamics within the present framework remains an important open challenge.

Tenth Excursus: Constructing Epistemic Space

In *Meaning and Necessity* (1947), Carnap laid the foundations for much of the contemporary discussion of possible worlds and of intensional semantics. In particular, he developed a notion of “state-description” that serves in effect as a linguistic construction of possible worlds. He also argued that every expression can be associated with an intension. This extensive modal and semantic project serves as the background for “Meaning and Synonymy in Natural Language”, in which Carnap engages in the metasemantic project of determining what it is for a subject to use an expression with a given meaning. The modal and semantic projects deserve attention in their own right, however. In this excursus, I focus on analogs of the modal project within the current framework, while in the eleventh excursus, I focus on analogs of the semantic project.

Meaning and Necessity was published two decades after the *Aufbau*, and Carnap does not explicitly connect the projects. But as chapter 5 suggests, it is natural to draw a connection. The basic elements of the *Aufbau* provide a basic vocabulary for characterizing the world, and a set of atomic sentences in that vocabulary that characterize the actual world. In effect, these elements provide a state-description for the actual world. The definitional elements of the *Aufbau* provide a way to determine the truth of an arbitrary sentence, given the specification of truths in the basic vocabulary. This provides a way of evaluating the intension of a sentence given an arbitrary state-description. In the *Aufbau*, Carnap uses his basic vocabulary to characterize state-descriptions for non-actual states of the world, but one could certainly do so in principle. In this way, one could use the materials of the *Aufbau* to construct the state-descriptions and intensions that are needed for the project of *Meaning and Necessity*.

Of course there are some differences of detail. Where the *Aufbau* uses an austere basic vocabulary (logic plus a basic relation) and a rich system of semantic rules (arbitrary definitions), *Meaning and Necessity* allows a rich basic vocabulary (atomic sentences containing arbitrary predicates and individual constants) and an austere system of semantic rules (logical relations between atomic and complex sentences). And where the *Aufbau* requires only extensional adequacy of its definitions, *Meaning and Necessity* appeals to a notion of “L-truth” (truth in virtue of the semantic rules of a language L) that Carnap says is akin to a notion of analyticity or necessity. Still, as discussed in Chapter 1, there is a nearby *Aufbau* project that requires something like analyticity of its definitions, and this might be used to ground a nearby *Meaning and Necessity* project that uses austere bases to define state-descriptions and rich definitional connections to define intensions.²¹

²¹An austere base might help with certain internal tensions in the *Meaning and Necessity* construction. Carnap says

Something like this construction is at the heart of contemporary linguistic constructions of possible worlds. In these constructions, metaphysical necessity plays the role that L-truth plays for Carnap, and with expressions for fundamental objects and/or properties playing the role of the basic vocabulary. One can thereby define possible worlds and intensions for expressions, with the key property that a sentence S is metaphysically necessary iff its intension is true in all possible worlds.

The scrutability project allows us to execute a related but quite different construction: a construction not of metaphysically possible worlds, but of epistemically possible scenarios. On this construction, apriority plays the role that L-truth plays for Carnap, and expressions in a generalized scrutability base playing the role of the basic vocabulary. These scenarios can play a crucial role in the analysis of knowledge, belief, and meaning.²²

Intuitively, an epistemically possible scenario (or a *scenario* for short) is a maximally specific way the world might be, for all we know a priori. For example, we do not know a priori that gold is an element. For all we know a priori, we could be in a world in which gold is a compound. Correspondingly, there are many scenarios in which gold is a compound, as well as many scenarios in which gold is an element. This already suggests that epistemically possible scenarios are distinct from metaphysically possible worlds. On the usual understanding of metaphysical possibility, it is metaphysically necessary that gold is an element (given that it is actually an element), so there are no possible worlds in which gold is a compound. Instead, we have to understand epistemically possible scenarios in different terms.

To construct scenarios, we can start by stipulating that S is *deeply epistemically necessary* when S is conclusively a priori (see section 1 of chapter 4), and that S is *deeply epistemically possible* when $\neg S$ is not epistemically necessary. Note that this stipulative notion of deep epistemic

(p. 15) that where 'H' and 'RA' are predicates for 'human' and 'rational animal' respectively, $(x)(Hx \equiv RAx)$ is L-true, on the grounds that 'H' and 'RA' are synonymous. But on Carnap's official definitions (pp. 3-4 and pp. 9-10), L-truth requires truth in all state-descriptions, state-descriptions allow arbitrary recombinations of atomic sentences or their negations, and sentences of the form Hc and RAc (for any constant c) are atomic. On these definitions, there will be state-descriptions containing both Hc and $\neg RAc$, so that $(x)(Hx \equiv RAx)$ will not be L-true. More generally, the rich atomic language along with free recombination has the consequence that all sorts of apparently analytic sentences will be false in some state-description. An *Aufbau*-style version of the *Meaning and Necessity* project with an austere base vocabulary and a rich system of definitions corresponding to a rich notion of L-truth would help to avoid these problems. Alternatively, one could retain the rich base vocabulary but impose a constraint of L-consistency on recombinations of atomic sentences.

²²What follows is a highly abbreviated discussion of issues many of which are elaborated at much greater length in "The Nature of Epistemic Space" (Chalmers 2011).

possibility differs from the ordinary notion of what is epistemically possible for a subject, both because it is idealized (all mathematical theorems are epistemically necessary and their negations are epistemically impossible, even if no-one has proved them) and because it does not depend on what a given subject knows ('I am not conscious' is deeply epistemically possible, even though I am introspectively certain that I am conscious). For convenience I will abbreviate "deeply epistemically possible" as "e-possible" and sometimes as "epistemically possible" in what follows, but this should not be confused with the nonidealized subject-relative notion. For now we can restrict attention to context-independent (specifically, epistemically invariant) sentences S , although the definition can naturally be extended to define the e-possibility of a context-dependent sentence S in a context.

We can also stipulate as before that a sentence G is *epistemically complete* when G is e-possible and there is no H such that $G \& H$ and $G \& \neg H$ are e-possible. For example, if $PQTI'$ a priori entails all truths, then (as noted in section 4), $PQTI'$ will be epistemically complete.²³ We can also say that two epistemically complete sentences G_1 and G_2 are *equivalent* when $G_1 \rightarrow G_2$ and $G_2 \rightarrow G_1$ are both e-necessary.

Given Generalized Scrutability, there will be a compact vocabulary that can be used to specify epistemically complete sentences corresponding not just to the actual world but to arbitrary epistemic possibilities. The thesis says that there is a compact class C of sentences such that for all e-possible S , S is a priori scrutable from some e-possible subclass C' of C (where a class of sentences is e-possible iff its conjunction is e-possible). From here, one can argue that for all e-possible S , S is e-necessitated by some epistemically complete sentence in C . In effect, C provides an array of epistemically complete sentences akin to $PQTI$, each of which corresponds to a highly specific epistemic possibility.

We can then identify scenarios with equivalence classes of epistemically complete sentences in the vocabulary of a generalized scrutability base. Given a scenario w , any sentence D in the corresponding equivalence class is a *canonical specification* of w . For a context-independent sentence S , a scenario w *verifies* S when S is a priori scrutable from a canonical specification of

²³Indeterminacy raises a few complications. Its treatment will depend on the issue, discussed in the first excursus, of whether the vagueness of epistemic necessity goes along with the vagueness of truth or with the vagueness of determinate truth. If we take the latter route, one should say that S is e-possible when $\neg \text{det}(S)$ (rather than $\neg S$) is not e-necessary. Then when G is the conjunction of sentences in an a priori scrutability base and H is indeterminate, $G \rightarrow \text{indet}(H)$ will be e-necessary and neither $G \& H$ nor $G \& \neg H$ will be e-possible. If we take the former route, then under plausible assumptions, each of the two latter sentences will be indeterminately e-possible, but it will be determinately false but both are e-possible. So either way, G will be epistemically complete.

w : that is, when $D \rightarrow S$ is a priori, where D is an epistemically complete sentence corresponding to w . Given a Generalized Scrutability thesis for context-dependent sentences, one can likewise say that w verifies an arbitrary sentence S in a context when S is a priori scrutable from a canonical specification of w in that context.²⁴

Given the above, this construction ensures the crucial principle of Plenitude (along with a number of other principles discussed in “The Nature of Epistemic Space”): S is e-possible iff there exists a scenario that verifies S . Likewise, for context-dependent sentences S , S is e-possible in a context iff there exists a scenario that verifies S in that context. For example, the e-possible sentence ‘Gold is a compound’ will be verified by many scenarios: intuitively, these are scenarios in which a compound gives rise to the appearances that we associate with gold. Likewise, the e-possible sentence ‘Hesperus is not Phosphorus’ will be verified by many scenarios: intuitively, these are scenarios in which the relevant bright objects in the evening and morning skies are distinct.

For any subject s (at time t in world w), there will be one scenario that is *actualized* for s (at t in w). This will be a scenario corresponding to an epistemically complete sentence (such as $PQTI$) that is true of w for s at t : specifying the objective character of w (perhaps using P , Q , and T) and the position of s and t within it (using I). The scrutability thesis tells us that a sentence S will be true for s (at t in w) if and only if it is verified by the scenario that is actualized for s (at t in w). Different scenarios will be actualized for different subjects (even within the same world), as reflected in the fact that the I component of $PQTI$ will be different for different subjects.

Scenarios as defined have many applications. They can be used to help understand talk about skeptical scenarios in epistemology, and more generally to serve as “epistemically possible worlds” in the analysis of knowledge and belief. They can be used to help understand the objects of subjective probability: it is arguable that subjective probabilities are in effect distributed over the space of scenarios. And perhaps most importantly, they can play a key role in the analysis of meaning and content, helping to analyse Fregean notions of meaning and internalist notions of mental content.

For these purposes (which are discussed further in the eleventh excursus and chapter 8), a central role is played by intensions: functions from scenarios to truth-values. The intension of S (in a context) is true at a scenario w if w verifies S (in that context), false at w if w verifies $\neg S$

²⁴Here the relevant sort of context-dependence is epistemic context-dependence. Primitive indexicals such as ‘I’ and ‘now’ can be taken to be epistemically context-independent. The required Generalized Scrutability thesis then requires a base of epistemically context-independent sentences while allowing context-dependent sentences in the dependent class. See the tenth excursus for more on this issue.

(in that context), indeterminate at w if w verifies $indet(S)$, and so on. The intension of S will be true at all scenarios iff S is a priori. So the intension of ‘Gold is an element’ will be true at some scenarios and false at others.

The intension so defined is a version of the primary or epistemic intension familiar from two-dimensional semantics. A sentence’s primary intension is its epistemic profile (chapter 1), mapping epistemically possible scenarios to truth-values. A sentence’s secondary intension is its modal profile, mapping metaphysically possible worlds to truth-values. A sentence S is a priori (epistemically necessary) iff its primary intension is true in all scenarios, and metaphysically necessary iff its secondary intension is true in all worlds.

When S is an a posteriori necessity, such as ‘Hesperus is Phosphorus’, its secondary intension will be true at all worlds, but its primary intension will be false at some scenarios. There will be some specifications of scenarios— $PQTI^*$, say—describing a scenario in which the objects visible in the morning sky (around the individual designated by ‘I’) are entirely distinct from the objects visible in the evening sky. If we discovered that we were in such a scenario, we would accept ‘Hesperus is not Phosphorus’. Likewise, conditional on the hypothesis that we are in such a scenario, we should accept ‘Hesperus is not Phosphorus’. So ‘Hesperus is not Phosphorus’ is conditionally scrutable from $PQTI^*$. The arguments earlier in this chapter then suggest that it is a priori scrutable from $PQTI^*$. So the primary intension of ‘Hesperus is Phosphorus’ will be false at a scenario specified by $PQTI^*$.

More generally, to evaluate the primary intension of a sentence S at a scenario w , one considers w as actual: that is, one considers the hypothesis that w actually obtains, or equivalently, the hypothesis that D is actually the case, where D specifies w . Here we can use conditional scrutability at least as a heuristic guide to a priori scrutability, and ask: conditional on the hypothesis that w is actual, should one accept S ? For example, conditional on the hypothesis that a Twin Earth scenario in which the oceans and lakes are filled by XYZ is actual, one should accept ‘Water is not H_2O ’. So the primary intension of ‘Water is H_2O ’ is false at this scenario.

By contrast, to evaluate the secondary intension of a sentence S at a world w , one considers w as counterfactual: that is, one considers counterfactually what would have been the case if w had obtained, or equivalently, if D had been the case, where D specifies w . Here as a heuristic we can ask: if w had been obtained (that is, if D had been the case), would S have been the case? In the case of a Twin Earth world, we can ask: if the oceans and lakes had been filled by XYZ, would water have been H_2O ? Following Kripke and Putnam, the standard judgment about this subjunctive conditional is “yes”. (By contrast, the intuitive judgment about the indicative

conditional ‘If the oceans and lakes are filled by XYZ, is water H₂O’ is “no”.) If so, the secondary intension of ‘Water is H₂O’ is false at the Twin Earth world.

We can also associate subsentential expressions with primary and secondary intensions. In general, a primary intension maps an epistemically possible scenario to extensions—objects for singular terms, properties for predicates, and so on—while a secondary intension maps metaphysically possible worlds to extensions. For example, intuitively the primary intension of ‘water’ picks out H₂O in the actual scenario and picks out XYZ in a Twin Earth scenario. Making this precise requires us first to make sense of the notion of objects within scenarios, which takes some work (as scenarios have so far just been constructed from sentences), but we can work with an intuitive understanding for present purposes.

The secondary intension of ‘Hesperus is Phosphorus’ reflects the fact that ‘Hesperus’ and ‘Phosphorus’ are *metaphysically rigid*, picking out the same entity—the planet Venus, their referent in the actual world—in all metaphysically possible worlds. By contrast, the primary intension of ‘Hesperus is Phosphorus’ (in paradigmatic contexts) suggests that ‘Hesperus’ and ‘Phosphorus’ are not *epistemically rigid*: they do not pick out the same entity in all epistemically possible scenarios (see E13 for much more here). If they did pick out the same entity in all scenarios, then ‘Hesperus is Phosphorus’ would be true in all scenarios and therefore a priori, which it is not. Rather, to a rough first approximation, ‘Hesperus’ picks out a bright object in a certain location in the evening sky in a given scenario, while ‘Phosphorus’ picks out a bright object in a certain location in the morning sky. In many scenarios, these two objects will be distinct.

This point generalizes to arbitrary names of concrete objects, suggesting that no name for a concrete object is epistemically rigid. By contrast, it is arguable that some names for abstract objects are epistemically rigid: for example, ‘0’ arguably picks out 0 in all scenarios, and ‘identity’ arguably picks out the relation of identity in all scenarios. One can similarly hold that numerous predicates—perhaps ‘conscious’, ‘causes’, and ‘omniscient’, among many others—are epistemically rigid, having the same property as extension in all scenarios.

Making the notion of epistemic rigidity precise is tricky, in part because we have not yet formally populated scenarios with objects, and in part because it is unclear that there is a coherent general notion of trans-scenario identity: that is, of what it is for entities in two different scenarios to be the same entity. But a useful intuitive gloss on the notion is that an epistemically rigid expression is one that expresses an epistemically rigid concept, and that an epistemically rigid concept is one whose extension we can know a priori. For example, there is an intuitive sense in which we cannot know what water is a priori, and in which we cannot know what Hesperus is a

priori, but in which we might be able to know what zero is a priori or what identity is a priori. That is roughly the sense at play here. I discuss this and many other issues about epistemic rigidity at more length in the thirteenth excursus.

When an expression is epistemically rigid and also metaphysically rigid (metaphysically rigid *de jure* rather than *de facto*, in the terminology of Kripke 1980), it is *super-rigid*. In this case, the expression will pick out the same entity in all scenarios and all worlds. For example, it is plausible that ‘0’ picks out zero in all scenarios and all worlds. As with epistemically rigid expressions, there are plausibly no super-rigid expressions for concrete objects, but there are plausibly super-rigid expressions for some abstract objects and properties.

What is the relationship between epistemically possible scenarios and metaphysically possible worlds? As I have discussed them so far, these are independent sorts of entities. But it is common to see a close relationship between them, modeling epistemically possible scenarios as centered metaphysically possible worlds. The existence of super-rigid expressions, which function to pick out the same entities in scenarios and in worlds, helps us to explore the connections between these entities. They can also help us to construct scenarios nonlinguistically using worldly entities such as properties and propositions.

To analyze the correspondence between scenarios and worlds it is helpful to highlight two theses about super-rigid expressions that I discuss in the thirteenth excursus and chapter 8. The first is Super-Rigid Scrutability: all epistemically possible sentences S are scrutable from sentences including only super-rigid expressions and primitive indexicals (such as ‘I’ and ‘now’).²⁵ The second is an Apriority/Necessity thesis: when a sentence S contains only super-rigid expressions, S is a priori iff S is necessary. The first thesis is supported by the character of the scrutability bases we arrive at and by general considerations concerning the scrutability of sentences containing epistemically nonrigid expressions. The second thesis is supported by the observation that paradigmatic Kripkean a posteriori necessities all appear to involve epistemically nonrigid expressions.

Super-Rigid Scrutability (and the considerations that support it) suggests that a generalized scrutability base need contain only certain basic super-rigid sentences (that is, sentences containing only super-rigid expressions) and certain indexical sentences such as ‘I am F_1 ’ and ‘Now is F_2 ’, where F_1 and F_2 are predicates containing only super-rigid expressions. Given this, scenarios can be identified with epistemically complete sentences of the form $D\&I$, where D is a complex

²⁵This has roughly the strength of a generalized scrutability thesis, but I omit “Generalized” here and later for ease of discussion.

super-rigid sentence (conjoining basic sentences) and I is a conjunction of indexical sentences as above. D will say roughly that there exist objects bearing certain specific properties and relations, and I will attribute certain specific properties and relations of oneself and the current time.

D will express a complex Russellian proposition p , containing properties and relations (perhaps along with other abstract objects) as constituents, connected by logical structure. This proposition is quite reminiscent of a possible world. It is common to regard possible worlds as complex Russellian propositions, specifying the distribution of certain basic properties and relations over objects. If the Apriority/Necessity thesis is correct, then D will be metaphysically possible (as it is super-rigid and epistemically possible), so p will be metaphysically possible. In this case, p will correspond to a metaphysically possible world.²⁶ If the Apriority/Necessity thesis is incorrect, then D may be metaphysically impossible, in which case p will correspond to a metaphysically impossible world. But either way p will correspond to a world-like entity involving the distribution of properties and relations over objects.

If a scenario w were specified by D alone, we could then identify the scenario with the world specified by p .²⁷ Given the additional role of I , which says something like ‘I am F_1 and now is F'_2 , we can instead identify the scenario with a *centered world*. Centered worlds are usually taken to be ordered triples of worlds, individuals, and times. For present purposes, we can take them to be ordered triples of a Russellian proposition p and properties ϕ_1 and ϕ_2 (corresponding to maximally specific properties possessed by the individual and the time respectively). For the scenario in question, we can take ϕ_1 and ϕ_2 to be the properties expressed by the predicates F_1 and F_2 respectively.

²⁶Strictly speaking, p will correspond either to a complete or an incomplete metaphysically possible world. Which is correct depends on the thesis of Super-Rigid Necessitation (an analog of Super-Rigid Scrutability): for any metaphysically possible sentence S , S is metaphysically necessitated by some sentence T including only super-rigid expressions. If this thesis (along with Apriority/Necessity) is true, the Russellian propositions in the text will be metaphysically complete (by analogy with epistemic completeness) and will specify full metaphysically possible worlds. If this thesis is false, these Russellian propositions may be metaphysically incomplete and will specify incomplete worlds (worlds without all details filled in), which in effect correspond to equivalence classes of metaphysically possible worlds.

Super-Rigid Necessitation will be false on certain haecceitist views [E19], on which a super-rigid specification of a world may underdetermine which objects are present in a world. It will also be false on certain quidditist views [7.9, E19], on which a super-rigid specification of a world may underdetermine which intrinsic properties are present in that world. Still, given Apriority/Necessity and Super-Rigid Scrutability, these propositions will correspond at least to incomplete metaphysically possible worlds. Given Super-Rigid Necessitation in addition, the propositions will correspond precisely to metaphysically possible worlds.

²⁷This only works when D is super-rigid. If non-super-rigid expressions are involved in the base, identifying sce-

If Super-Rigid Scrutability is true, we can always model scenarios as centered worlds in this way. If Apriority/Necessity is also true, scenarios will correspond to centered worlds where the worlds in question are always metaphysically possible.²⁸

Take a paradigmatic a posteriori necessity, such as ‘Water is H₂O’ which is false in a Twin Earth scenario. Super-Rigid Scrutability entails that a Twin Earth scenario can be specified by an epistemically possible sentence *TE&I* conjoining a super-rigid sentence *TE* with indexical claims involving ‘I’ and ‘Now. Here *TE* might involve a super-rigid specification of the microphysical and phenomenological character of the world, without using epistemically nonrigid terms such as ‘water’. Because *TE* is super-rigid, it can be used to specify a world: that is, a Russellian complex of objects and properties. Apriority/Necessity entails that *TE* is metaphysically possible, so that this world is a metaphysically possible world. *TE* metaphysically necessitates ‘Water is H₂O’, while *TE&I* epistemically necessitates ‘Water is not H₂O’. So the primary intension of ‘Water is H₂O’ is false at the Twin Earth scenario while its secondary intension is true at the Twin Earth world. But there remains a close correspondence between the scenario and the world. The difference in intensions for ‘Water is H₂O’ arises not so much due to differences between them as due to the difference between epistemic and metaphysical necessitation.

I accept Apriority/Necessity and Super-Rigid Scrutability. (Relatives of these theses play a crucial role in the two-dimensional argument against materialism; see Chalmers 2010.) So I hold that scenarios correspond closely to centered metaphysically possible worlds. In practice, many philosophers already use centered metaphysically possible worlds to model epistemically possible scenarios (this is a standard practice in the literature on subjective probability, for example). The analysis above can be seen as providing a partial grounding for this practice.

At the same time, some philosophers will reject Apriority/Necessity or Super-Rigid Scrutability. If these theses are rejected, the relation between scenarios and possible worlds becomes more complex.

First, consider a view that rejects Apriority/Necessity while retaining Super-Rigid Scrutability:

narios with Russellian propositions will give the wrong results. For example, if ‘Hesperus’ and ‘Phosphorus’ are in the base, ‘Hesperus is such-and-such’ will always specify a Russellian proposition about Venus, as will ‘Phosphorus is such-and-such’. So ‘Hesperus is Phosphorus will come out true in all scenarios, even though it is not a priori.

²⁸As before, if Super-Rigid Necessitation is false, scenarios will correspond to centered incomplete metaphysically possible worlds, or to equivalence classes of centered metaphysically possible worlds. If Super-Rigid Necessitation (along with the other two theses) is true, scenarios will correspond near-perfectly with centered metaphysically possible worlds. The only exception involves certain symmetrical worlds, where more than one centered world (centered on symmetrical counterparts) may correspond to the same scenario.

a theist view on which ‘There is an omniscient being’ is necessary but not a priori. This sentence *S* plausibly contains only super-rigid expressions. So on the view in question, *S* will be a counterexample to the Apriority/Necessity thesis. Correspondingly, *S* will be false in some epistemically possible scenario. Given Super-Rigid Scrutability, there will be a sentence *D&I* specifying such a scenario, where *D* is super-rigid and *I* is indexical. *D* will then specify a world (a Russellian propositions constructed from properties): in effect, a world without an omniscient being. But on the view in question, this world will not be metaphysically possible. So the scenarios here will correspond to centered metaphysically impossible worlds.

Next, consider a view on which Super-Rigid Scrutability is false: a type-B materialist view (discussed in excursus 13) on which phenomenal concepts are both conceptually primitive and epistemically nonrigid. Then corresponding expressions such as ‘consciousness’ are needed in a scrutability base, but one cannot know what these expressions pick out a priori: perhaps ‘consciousness’ picks out a neurophysiological property *N* empirically. In this case, scenario descriptions (such as *PQTI*) will not be decomposable into a super-rigid part and an indexical part. Even once indexicals are removed, some remaining expressions (such as the phenomenal expressions in *Q*) will be epistemically nonrigid, so that they will not map a priori onto properties. While the sentence will still correspond empirically to a complex of properties (such as *N*), the resulting property-involving world will not adequately model the epistemic properties of the scenario. (For example, it becomes hard to see how ‘Consciousness is *N*’ can be false at a centered world.) In this case, it is better to stay with a linguistic model of scenarios where we do not invoke properties at all. Or perhaps best, we can see scenarios as structures of Fregean senses, where we associate the primitives here with primitive Fregean senses that determine the properties that serve as their referents only a posteriori.

To summarize: if both Apriority/Necessity and Super-Rigid Scrutability are true, as on my view, scenarios correspond closely to centered metaphysically possible worlds. If the former but not the latter is true, scenarios still correspond closely to centered worlds, but the worlds in questions may be metaphysically impossible in some cases. If the latter is false, one has to break the close connection between scenarios and centered (property-involving) worlds, instead constructing scenarios from sentences or from Fregean propositions.

What about two-dimensional evaluation? On the current model, this is tied to epistemic possibilities concerning what is metaphysically possible. For example, someone might hold that it is epistemically possible that there is exactly one metaphysically possible world, and epistemically possible that there are enormously many metaphysically possible worlds. We could model this

by a two-dimensional modal structure on which every epistemically possible scenario v is associated with a space of worlds that are metaphysically possible relative to v . If one holds (as I do) that Apriority/Necessity and Super-Rigid Scrutability are both a priori, then every scenario will be associated with a space of worlds such that there is a centered world for every scenario. Under slightly stronger assumptions (discussed in Chalmers 2006), one can use the same space of metaphysically possible worlds to play the role of the putative worlds for each of these scenarios. Under these assumptions one gets a simple “rectangular” two-dimensional structure, with every scenario corresponding to the same space of worlds. Under weaker assumptions one gets a more complex structure, on which some scenarios may be associated with smaller or larger spaces of worlds, and on which some worlds may be metaphysically impossible (relative to the actual scenario).

Either way, one can use this two-dimensional structure to define two-dimensional intensions for sentences: given a scenario v and a world w that is possible relative to v , the two-dimensional intension of S will map (v, w) to the truth-value of S at w , conditional on the assumption that v is actual. These two-dimensional intensions are a version of the two-dimensional matrices familiar from two-dimensional semantics.

An important residual issue concerns the idealized notion of epistemic possibility that we started with. This works well in modeling the epistemic states of idealized agents, but not as well in modeling the epistemic states of nonideal agents. For example, a nonideal agent might disbelieve certain moral or mathematical truths, even though these truths are a priori, or the agent might simply fail to believe these. The current framework does not easily model these states, as a priori truths are true in all scenarios. To model these states, it would be useful to invoke a less idealized notion of epistemic necessity, and a correspondingly more fine-grained space of nonideal scenarios.

One idea here is that here one might understand nonideal epistemic necessity in terms of a notion of analyticity rather than apriority (such as the notion discussed later in this book, especially chapter 9 and the eighteenth excursus), and construct fine-grained scenarios and fine-grained intensions from there. Then insofar as the relevant moral truths (for example) are a priori but not analytic, there will be scenarios where they are false. If so, we may have groups of fine-grained scenarios with the same natural truths and different moral truths. However, it is not clear that this sort of model will help with ignorance of logical truths: these truths are often taken to be analytic, and it is not obvious how best to model scenarios in which they are false. The understanding of nonideal epistemic space remains open as a challenge for future work.²⁹

²⁹Bjerring (2010) explores a number of models of nonideal epistemic space, and demonstrates that there are serious

difficulties in developing a model in which complex logical truths are false but simple logical truths are not. As a result, one may be left with a choice between logical omniscience (logical truths are true in all worlds) and triviality (any set of sentences determines a scenario).