

# Coda: Explaining Scrutability

David J. Chalmers

I have argued that various scrutability theses are true, but I have not yet explained why they are true. I have argued for them mainly by attending to truths about the world and arguing that they are all scrutable from various limited bases. This provides an argument, but it does not yet provide an explanation.

How might we explain the truth of the scrutability theses? This question can take two subtly different forms. First, *given* that the theses are true, we can look for explanations of why they are true, assuming simpler premises than scrutability theses themselves. Second, without assuming that the theses are true, we can look for arguments for why they *have* to be true, assuming a priori premises of which we are antecedently more confident than we are of scrutability itself.

The first project is arguably more important, but the second project can contribute to it, as a strong a priori argument for scrutability might also provide the elements of an explanation of scrutability. The second project is also useful for dialectical purposes: making an a priori argument for scrutability that does not depend on empirically surveying truths about the actual world may help to block off certain sorts of opposition.

I do not have conclusive answers to either of these questions, but I have some relevant thoughts. I will start with the first project.

Why are scrutability theses true? Of course these could be brute facts that are not to be further explained, but they seem too complicated for this to be a plausible diagnosis. If they are true, there are surely simpler principles in terms of which they can be explained. Given this much, explanatory pluralism suggests that there are probably multiple explanations to be found.

In section 8 of chapter 8, I have done some preliminary charting of the explanatory relations *between* scrutability theses. There I suggested that Acquaintance Scrutability might explain Primitive Scrutability (via the thesis that acquaintance concepts are primitive concepts) and Narrow Scrutability (via the thesis that in the actual world, acquaintance concepts are narrow concepts). We have also seen that Acquaintance Scrutability along with the Apriority/Necessity Thesis entail

Fundamental Scrutability. Furthermore, Fundamental Scrutability, combined with the thesis that the class of expressions for fundamental properties (plus indexicals) is compact, entails Compact Scrutability. If all this is right, we might see Acquaintance Scrutability, along with the ancillary theses mentioned here, as explaining, the other four theses. That reduces the question of explaining those theses to that of explaining Acquaintance Scrutability, and perhaps explaining the ancillary theses.<sup>1</sup>

We still need to explain Acquaintance Scrutability, however. And if one rejects the Apriority/Necessity thesis and Fundamental Scrutability (as many will), one will also need to explain Compact Scrutability in some other way.

What about the modality of a scrutability thesis? A Priori Scrutability can arguably explain Conditional Scrutability and Inferential Scrutability. A Priori Scrutability theses might themselves be explained in terms of Analytic Scrutability theses, or the Translucent Scrutability theses of chapter 9. Assuming that analyticity entails apriority, then analytic scrutability entails a priori scrutability. Analytic scrutability theses seem more basic in some respects: in particular, analytic scrutability bases seem to better reflect the class of primitive concepts. So it is arguable that Analytic Acquaintance Scrutability can explain A Priori Acquaintance Scrutability, and that Analytic Compact Scrutability can explain A Priori Compact Scrutability. But the explaining theses themselves need explanation.

At this point, it is natural to try explaining these theses in terms of the character of our concepts. For example, one might invoke *conceptual descriptivism*: the thesis that every concept is composed from a compact set of primitive concepts. (Here I construe concepts as types of mental representations, though something similar will apply if one construes concepts as abstract objects.) Given that thoughts are composed from concepts, and given that sentences express thoughts and subsentential expressions express concepts, it follows that every expression is equivalent to one composed of expressions that express primitive concepts. If this is right, then all truths will be analytically entailed by truths involving expressions that express primitive concepts alone.

So conceptual descriptivism might be taken to ground a Primitive Analytic Scrutability Thesis. Given the further theses that there is a compact class of primitive concepts, or that primitive concepts are all acquaintance concepts, this thesis might also explain Compact Analytic Scrutability and Acquaintance Analytic Scrutability.

I think that conceptual descriptivism is probably false, for reasons discussed in chapter 1. Just

---

<sup>1</sup>My best attempt at both explaining the Apriority/Necessity thesis and providing an a priori argument for it is in Chalmers 2010, chapter 6.

as most expressions are not definable from a limited class of primitive expressions, most concepts are not equivalent to complexes composed from a limited class of primitive concepts. But we saw in chapter 9 that anchored inferentialism, a weaker cousin of conceptual descriptivism, avoids these problems. This raises the prospect that anchored inferentialism might explain the scrutability theses above.

According to inferentialism, concepts are individuated by their inferential roles. According to anchored inferentialism, there is a limited set of primitive concepts, and all other concepts are grounded in their inferential role with respect to these concepts. If we construe concepts as a certain sort of mental state (or mental representation), this thesis can be seen as a thesis about the existence and content of concepts: a state counts as a concept in virtue of its inferential role, and a concept has its content in virtue of its inferential role. (If concepts were construed as abstract objects, we could construe the thesis as a thesis about the possession of concepts instead.) The view is neutral on how primitive concepts are individuated, but one natural view holds that they are individuated by acquaintance with their referents.

The inferential role of a concept can be construed as a normative role, constituted by *good* inferences that the concept might be involved in. On one construal, the inferential role will be a *a priori* role, involving the *a priori* justified inferences that the concept is involved in. On another, it will be a sort of analytic role, involving the trivial or cognitively insignificant inferences that the concept is involved in. Either way, inferentialism will give special weight to *entry inferences*: good inferences from thoughts constituted by primitive concepts alone to thoughts involving the concept in question. There may also be a role for *exit inferences*: involving good inferences from thoughts involving the concept to thoughts involving primitive concepts alone. Typically, however, a pattern of entry inferences will itself fix a pattern of exit inferences. If so, one can hold that concepts are individuated by entry inferences alone.

This model fits the scrutability framework well. Entry inferences can be represented as inferences from sets of sentences composed of expressions for primitive concepts to sentences that also involve non-primitive concepts. Under certain reasonable assumptions, the pattern of entry inferences for a nonprimitive concept *C* will itself fix a pattern of scrutability inferences: inferences from certain complete sets of sentences involving expressions for primitive concepts (where these sets correspond to full scenarios) to sentences also involving *C*. And in the reverse direction, these scrutability inferences will fix the pattern of entry inferences. So on this model, the content of any concept corresponds to a pattern of scrutability inferences.

If this is right, then the inferential role associated with any concept will fix the scrutability

inferences it is involved in. If we grant that inferences fixed by constitutive inferential roles are a priori or analytic, then this will explain a priori and analytic scrutability from the relevant bases. I think it is most natural to appeal to a sort of analytic scrutability here, but this requires first unpacking a relevant notion of analyticity.

I discussed a few notions of analyticity in chapter 9, but there is an important way of understanding analyticity that I have not discussed so far. This understanding invokes the framework of warrants discussed in fourth excursus. We can say that  $S$  is *warrant-analytic* for a subject if there is a *conceptual warrant* the subject to accept  $S$ . Warrant-analyticity can be seen as a variety of positional analyticity, discussed in chapter 9 ( $S$  is positionally analytic when

We can say that a sentence  $S$  is *warrant-analytic* when there exists a conceptual warrant for any subject who uses the expressions in  $S$  (with full competence) to accept  $S$ . Warrant-analyticity is a relative of the notion of positional analyticity discussed in chapter 9:  $S$  is positionally analytic if any subject who uses the expressions in  $S$  (with full competence) is in a position to know  $S$ . Warrant-analyticity cashes out “in a position to know” in terms of warrants, as in the fourth excursus, and more importantly, it constrains the sort of warrants that are relevant to a special sort: conceptual warrants.

Here, intuitively, a conceptual warrant for accepting a sentence is one that derives wholly from the concepts expressed by the expressions in that sentence. For example, it is natural to hold that there is a warrant for accepting ‘Vixens are female foxes’ that derives from the concept expressed by ‘vixens’, ‘foxes’, and so on. The existence of a warrant does not entail that any given subject will use the warrant, so the warrant-analyticity of a sentence  $S$  does not entail its epistemological analyticity, and warrant-analyticity is not subject to Williamson’s critique of epistemological analyticity. But it is still an epistemological notion, and is one that may be able to play some roles of the traditional notion of analyticity.

Can we characterize conceptual warrants more precisely? Here the inferentialist analysis of concepts provides some potential tools. On this view, most concepts can be characterized by certain inferential roles with respect to other concepts. One might suggest that there is a conceptual warrant to accept a proposition  $p$ , constituted by various concepts, when  $p$  is warranted by the constitutive inferential roles of the concepts that constitute it.

For example, suppose that the inferential role of *vixen* is constituted by the obvious inferential relations to *female* and to *fox*. Then it is natural to hold that an inference from  $x$  is a *vixen* to  $x$  is a *female fox* is warranted by these constitutive inferential relations. It is not a large step from there to hold that *vixens are female foxes* is warranted by these inferential relations (perhaps along

with inferential relations deriving from *are* and from various logical concepts). If this is right, then there is a conceptual warrant to accept *vixens are female foxes*. As before, none of this entails that a subject possessing the concepts must accept the proposition: constitutive inferential roles are normative rather than descriptive. But the roles nevertheless provide a warrant.

Much more would need to be said to make this picture fully precise. As well as spelling out the rules out what can be a constitutive inferential role, we would need a precise account of just how these inferential roles have to be related to a proposition to warrant it. There are also questions about how precisely to represent conceptual warrants in the framework of support structures: they might be seen as a sort of basic warrant, providing basic a priori evidence, or alternatively they might be seen as grounded in concepts somehow. But we have enough on the table here to provide the basic picture.

Given that all nonprimitive concepts have constitutive inferential roles connecting them to primitive concepts, one can then see the outlines of an explanation of why all truths are analytically scrutable from primitive truths. Take a sentence *S* expressing (or apt to express) a thought *T*. Every nonprimitive concept involved in *T* will be associated with have entry rules endorsing inferences from thoughts involving primitive concepts to thoughts involving it. We If we put together the inference rules for all nonprimitive concepts in *T*, we can expect these entry rules to determine inferences from certain sets of thoughts *T'* involving only primitive concepts to *T*. If *S'* expressing a conjunction of the thoughts in *T'*, then these constitutive inferences will also determine a conceptual warrant for inferences from a sentence *S'* to *S*. Likewise, they will also provide a conceptual warrant for accepting a conditional 'If *S'*, then *S*'. So that conditional will be warrant-analytic. So *S* is analytically scrutable from the conjuncts of *S'*, if analytic scrutability is understood in terms of warrant-analyticity. This reasoning applies to any sentence *S*, so it follows that any sentence is analytically scrutable from some set of sentences expressing primitive concepts. This explains a generalized version of Analytic Scrutability.

This explanation is highly congenial with the basic picture of conceptual application set out in section 3 of chapter 1. On that picture, grasp of a concept goes along with a conditional ability to identify an expression's extension, given sufficient information about how the world turns out and sufficient reasoning. We can now see how such a conditional ability might fall out of constitutive inferential connections to primitive concepts, along with the claim that subjects are always in a position (ideally) to make the constitutive inferences. When they do, the inferences will have a conceptual warrant. Once a picture of this sort is granted, analytic scrutability is only to be expected.

What about the scrutability base? If we grant that all primitive concepts involve acquaintance, then this sort of inferentialism will then yield versions of Acquaintance Scrutability. If we grant that there is only a compact class of primitive concepts, this sort of inferentialism will then yield versions of Compact Scrutability. The first thesis is a reasonably natural one for an anchored inferentialist. The second thesis is not obvious, though. Anchored inferentialism might suggest that the class of primitive concepts will be *relatively* compact compared to the class of all concepts: for any given class of primitives, many more nonprimitives will be definable by inferential roles with respect to them. But it is not clear why this class should be absolutely compact, involving just a handful of families.

At this point, we might appeal to contingent facts about the human mind. One could argue that as a matter of fact humans have only a handful of families of primitive concepts. Alternatively, we might appeal to necessary constraints on possible thinkers. One could argue that any primitive concept must meet certain constraints satisfied only by a limited number of concepts. For example, a structuralist might hold that any primitive concept must be structural, so that only a few logical and relational concepts can be primitive. An acquaintance theorist might hold that any primitive concept must be grounded in acquaintance, which imposes strong constraints: for example, the constraint that one is only acquainted with phenomenal properties and perhaps with properties that are phenomenally represented. These constraints might limit the primitive concepts enough to yield a compact class.

An intermediate view holds that while there are many *possible* primitive concepts, in the actual world these concepts are limited for contingent reasons. For example, the actual world has a limited fundamental structure, involving a small number of fundamental properties and laws. Only a few primitive concepts are required to represent this structure. Given Fundamental Scrutability, all truths are scrutable from truths involving these concepts. This suggests that a limited anchored inferentialism with these concepts as anchors may suffice for the purposes of representing the empirical structure of our world. The intermediate view leaves open the possibility that other worlds may be more complex, involving all sorts of alien properties and alien structure that may require entirely different concepts to represent them. In such worlds, there is no guarantee that a compact class of concepts will suffice, although (given the necessity of inferentialism) there is still reason to think that a relatively compact class will suffice.

One might object that an explanation in terms of concepts puts too much weight on the semantic and the psychological. After all, scrutability theses are partly about the world, holding that all truths about the world are scrutable from this limited base. One could respond that given

that truths here are true sentences, we should expect these to be constrained by what we can think and express. Scrutability itself does not tell us how many facts (true propositions, say) about the world are expressible in language. It is arguable that there are some inexpressible facts: facts about ungraspable quiddities, for example. If so, then the best we can hope for is scrutability of the expressible facts. If there are no inexpressible facts, on the other hand, then the explanation of scrutability above might then explain why all facts about reality are scrutable, although at cost of requiring an explanation of why there are no inexpressible facts. I do not have a settled view on that question, although an explanation might naturally go through the theses that all fundamental facts are expressible.

Overall, I think there is a reasonable case that anchored inferentialism can explain Acquaintance Scrutability (in analytic or a priori varieties). Acquaintance Scrutability along with the Apriority/Necessity thesis can explain Fundamental Scrutability. These theses along with contingent theses about our world (e.g., that it has a compact fundamental nature) can explain Compact Scrutability.

Anchored inferentialism about concepts is probably not an explanatorily basic thesis. We can ask how the truth of inferentialism is to be explained, and we can expect there to be an answer. The same goes for the ancillary assumptions. Still, one cannot explain everything at once. Even if inferentialism and the ancillary assumption are not themselves fully explained, they can still explain the scrutability thesis.

Inferentialism about concepts is controversial. But given the apparent truth of the scrutability thesis and the power of an inferentialist model to explain it, the inferentialist model is arguably supported by a sort of inference to the best explanation.

[Answer to second problem still to come!]

## Nineteenth Excursus: Scrutability, Supervenience, and Grounding

Our central scrutability theses have been epistemological theses, not metaphysical theses. But it is natural to ask what sort of metaphysical upshot these scrutability theses might have. For example: does a minimal scrutability base serve as a guide to the fundamental metaphysical structure of the world?

To ask this is to ask about the reach of conceptual metaphysics. In the introduction, I said that conceptual metaphysics investigates the structure of our conception of reality, with one eye on how well this structure corresponds to reality itself. At a finer grain, conceptual metaphysics divides into three parts. The first focuses on the *structure of concepts*: relations among the concepts involved in our conception of the world, unconstrained by external reality. The second focuses on the *conceptual structure of reality*: conceptual relations among truths about reality. The third focuses on the *metaphysical structure of reality*: using conceptual relations as a guide to metaphysical relations among truths about reality.

I have occasionally engaged in the first project in this book. For example, discussions of generalized scrutability and the class of primitive concepts are largely unconstrained by external reality. Some of our primitive concepts may have no application to the actual world: for example, it may be that our basic conception of the world is of an Edenic world that is very different from reality. Still, the structure of concepts at least serves as a constraint on the conceptual structure of reality.<sup>2</sup>

I have mainly engaged in the second project in this book. Ordinary scrutability theses reflect conceptual and epistemological relations among truths about the world. The focus on truths means that the project is constrained by empirical reality. For example, the base truths include truths from physics, and an important constraint is that all truths about reality be scrutable. In effect, we isolate conceptually and epistemologically fundamental truths about the world, helping to understand the structure of reality as reflected in our concepts.

I have only rarely engaged in the third project in this book. That is, I have largely been unconcerned with how well the conceptual and epistemological relations reflect metaphysical relations. The main exception has been the discussion of whether all truths are scrutable from metaphysically fundamental truths [8.6]. I think that there is a great deal of promise in the third project, however.

---

<sup>2</sup>The first project is closely related to Strawson's descriptive metaphysics, though it focuses on our concepts rather than on our beliefs. It is also not out of the question for conceptual structure to be revisionary in that primitive concepts are unfamiliar concepts.

Metaphysical relations among truths about reality do not float free of conceptual relations, but are heavily constrained by them. So we should expect conclusions about the conceptual structure of reality to have at least some consequences for the metaphysical structure of reality.

We might think of the third project as *conceptually-guided global metaphysics*: using concepts as a guide to the global metaphysical structure of reality.<sup>3</sup> Of course this project, like the second, involves a heavy interplay of the conceptual and the empirical. Empirical methods such as those of physics play an enormous role in delivering fundamental truths and in delivering nonfundamental truths. But the relation between the fundamental and the nonfundamental requires careful philosophical analysis. The analysis of this relation can play a significant role in constraining which truths are fundamental in turn. In this excursus, I concentrate on the role that scrutability and related notions can play in this project.

We can approach the question by comparing scrutability to two related notions often thought to do metaphysical work in connecting the fundamental and the nonfundamental. One notion is *supervenience* (Kim 19xx): B-properties supervene on A-properties when any two possible worlds that are indiscernible with respect to their A-properties are indiscernible with respect to their B-properties. Another notion is *grounding* (Schaffer 2009): B-properties are grounded in A-properties when B-properties are instantiated in virtue of A-properties being instantiated.<sup>4</sup>

I will start with supervenience. How does the claim that B-truths are scrutable from A-truths relate to the claim that B-properties supervene on A-properties? The most obvious difference here is that scrutability theses concern truths (sentences or perhaps propositions) while supervenience theses concern properties. Another obvious difference is that the first claim is cast in terms of a priori entailment (an epistemological notion) whereas the second is cast in terms of possibility (a modal notion). Apart from these differences, there is also a structural difference in the way the notions are cast.

---

<sup>3</sup>For closely related projects in conceptually-guided global metaphysics, see Jackson 1998 and Thomasson 20xx. For recent projects in global metaphysics that are not especially conceptually guided, see Schaffer 2009 and Sider 20xx.

<sup>4</sup>A third notion is that of *metaphysical definition* (Fine 1994; Sider forthcoming): B-properties are metaphysically definable in terms of A-properties when for each B-property, there is a metaphysical definition of it that appeals only to A-properties. Unlike the definitions on which I have focused in this book, metaphysical definitions are usually not constrained to be conceptual or a priori truths. Still, I think the counterexample problems outlined in chapter 1 also pose problems for metaphysical programs grounded in metaphysical definition. In any case, the three metaphysical projects (metaphysical definition, supervenience, metaphysical grounding, metaphysical definition) can be seen as analogous to the three epistemological/conceptual projects (definitional scrutability, a priori scrutability, analytic scrutability or conceptual grounding).

On the first difference, we can line up truths and properties by saying that some sentences (e.g. the A-sentences) *characterize* some properties (e.g. the A-properties) when the sentences fully specify the instantiation of the properties. More precisely, A-sentences characterize A-properties when two worlds are indiscernible with respect to the A-properties iff the same A-sentences hold in them. The following discussion will focus on pairs of scrutability theses and supervenience theses satisfying a characterization assumption: the scrutability thesis holds that B-truths are scrutable from A-truths and the supervenience thesis holds that B-properties supervene on A-properties, where the A-truths are just the true A-sentences, A-sentences characterize A-properties, and B-sentences characterize B-properties.

The characterization assumption serves in part a formal rather than substantive constraint on the scrutability thesis at issue: roughly, it has to concern truths about the instantiation of properties. Most of the scrutability theses we are concerned with can be put into this form straightforwardly. But the assumption also builds in a substantive claim about the expressibility of the properties involved in the supervenience thesis. If there are A-properties that cannot be referred to by any expression, then there may be two A-discernible worlds in which the same sentences are true. If so, there will be no A-sentences that characterize the A-properties. This sort of inexpressibility provides one way in which supervenience (a thesis about properties) and scrutability (a thesis about sentences) can come apart. I will set worries about inexpressibility aside for now, and return to them later.

The second difference concerning necessity and a priority will play an important role shortly, but to allow us to address the structural issues, it is useful to first abstract away from the modal/epistemological difference by comparing supervenience not to A Priori Scrutability but to Necessary Scrutability: the thesis that all B-truths are necessitated by A-truths. We might put this thesis by saying that in every possible world where all the (actual) A-truths are true, all the (actual) B-truths are true. Unlike other scrutability theses, this thesis is not cast in epistemological terms (so it is really a scrutability thesis only in a weak sense), but it nevertheless shares a common structure with them.

Given these assumptions, scrutability (that is, Necessary Scrutability) is very nearly a consequence of supervenience. One might reason: a world  $w$  where all the actual A-truths are true will be A-indiscernible from our world, so (by supervenience)  $w$  will be B-indiscernible from our world, so all the actual B-truths will be true at  $w$ . The only questionable step here is the first: perhaps all actual A-truths are true at  $w$  but some other A-sentences are true there as well. This cannot happen if we assume that the A-sentences are closed under negation (setting aside indeterminacy), so that A-truths include both positive and negative A-truths, that is, truths about both the instan-

tiation and non-instantiation of A-properties. If they include only the former, it will also suffice to assume that the A-sentences include a “that’s-all” truth saying that these are the instantiations of A-properties. Call the assumption that the A-sentences are either closed under negation or include a that’s-all truth the completeness assumption. Given the completeness assumption (along with the structural assumption), supervenience entails Necessary Scrutability. Without the completeness assumption, the entailment will not quite go through: for example, the number of apples plausibly supervenes on applehood, but it is not necessitated by positive truths about applehood alone.

In the reverse direction, supervenience is not a consequence of scrutability, even given the characterization and completeness assumptions. Necessary Scrutability says that actual A-truths necessitate actual B-truths, but it makes no such claim about A-sentences and B-sentences in other worlds. While the connections between the actual A-truths and the actual B-truths must be necessary, the thesis itself may be contingent. For example, a version of Necessary Scrutability holding that all truths are necessitated by physical truths may be true in some physicalist worlds and false in other nonphysicalist worlds. By contrast, supervenience theses as defined so far are not tied to the actual world, and (at least given S5) will be necessary is true at all.

Scrutability theses are more closely analogous to a weaker sort of supervenience thesis (Lewis 19xx, Chalmers 1996) tied to a specific world. These are sometimes called a contingent supervenience theses (Lewis 19xx, Chalmers 1996), although they might more accurately be called worldwise supervenience theses, as the relation may or may not hold contingently. We can say that B-properties c-supervene on A-properties in world  $w$  if any world that is A-indiscernible from  $w$  is B-indiscernible from  $w$ . Given the structural and the completeness assumptions, Necessary Scrutability is equivalent to c-supervenience in the actual world. Without the completeness assumption, Necessary Scrutability will be slightly stronger than c-supervenience.

Another way to draw the notions more closely into alignment is to move from Necessary Scrutability to the stronger Generalized Necessary Scrutability: here understood as the thesis that in every world, the A-truths in that world necessitate the B-truths in that world. Supervenience is certainly a consequence of Generalized Necessary Scrutability. Under the structural and completeness assumption, supervenience will be equivalent to Generalized Necessary Scrutability. Without the completeness assumption, Generalized Necessary Scrutability will be slightly stronger than supervenience.

The moral is that when the modalities are aligned, ordinary scrutability theses have approximately the same strength as contingent supervenience theses (at least given the relevant assump-

tions), and ordinary supervenience theses have approximately the same strength as generalized scrutability theses. The same goes if we compare a priori scrutability theses to epistemic supervenience theses: for example, B-properties epistemically supervene on A-properties when all epistemically possible scenarios that are A-indiscernible are B-indiscernible (or to avoid worries about reidentifying properties across scenarios, one might cast such a thesis in terms of concepts or expressions). These epistemic supervenience theses have roughly the same force as generalized a priori scrutability theses, while ordinary a priori scrutability theses have roughly the same strength as (epistemically) contingent supervenience theses.

Why not cast scrutability theses as epistemic supervenience theses from the start? This is partly because I have been most concerned with ordinary rather than generalized scrutability theses, and these align less well with the most familiar supervenience theses. More importantly, it is because I have not wanted to presuppose the relatively unfamiliar apparatus of epistemically possible scenarios. Casting supervenience theses in terms of possible worlds rather than necessitation is useful because worlds are so familiar and vivid, but it is more straightforward to cast scrutability theses in terms of the a priori.

We can now abstract away from the structural differences, comparing the epistemological thesis that B-truths are scrutable from A-truths to the modal thesis that B-properties c-supervene on A-properties. Contingent supervenience is arguably the most important sort of supervenience for metaphysical purposes. For example, it is plausible that the metaphysical thesis of physicalism does not require that mental properties supervene on physical properties, but it requires (at least) that mental properties c-supervene on physical properties. Physicalism is a thesis about the actual world, and is consistent with various supervenience-falsifying claims: for example, it is consistent with the claim that there are non-actual worlds that are physically indiscernible and differ in that one also has additional nonphysical minds. So it is most relevant to compare a priori scrutability to c-supervenience, or equivalently (given the structural and completeness conditions) to necessary scrutability.

If apriority and necessity were equivalent, then a priori scrutability (a-scrutability) and necessary scrutability (n-scrutability) would be equivalent. But given that there are truths that are necessary but not a priori, or vice versa, the theses come apart. For example, 'There is water' is n-scrutable but not a-scrutable from 'There is H<sub>2</sub>O'.

Still, a weaker link between a-scrutability and n-scrutability remains tenable. It is arguable that the gap between apriority and necessity in 'water' cases and the like arise because 'water' is not super-rigid. The Apriority/Necessity thesis discussed earlier [8.5] says that sentences composed of

super-rigid expressions are necessary iff they are a priori. If one accepts this thesis, it follows that if A- and B-truths involve only super-rigid expressions, B-truths will be n-scrutable from A-truths iff they are a-scrutable from A-truths.

A complication arises because the scrutability bases we have considered have not been restricted to super-rigid expressions: they also involve primitive indexicals such as 'I' and 'now'. These indexicals can generate a gap between n-scrutability and a-scrutability. Still, one might suggest a weaker link:

Linking Thesis: For any class of super-rigid A-truths, all truths are n-scrutable from the A-truths iff all truths are a-scrutable from the A-truths plus indexical truths.

The Linking Thesis articulates a strong link between supervenience and scrutability theses. Roughly, any (contingent metaphysical) supervenience base yields an (a priori) scrutability base and vice versa, as long as the relevant base expressions are super-rigid, and the scrutability base is augmented by indexical truths. So I will spend some time assessing the prospects for this thesis, and for theses in the vicinity.

I have already in effect given an argument for the left-to-right direction of the Linking Thesis in Chapter 8, in the argument for Fundamental Scrutability there. The key premises there were the Apriority/Necessity thesis (super-rigid truths are necessary iff they are a priori) and the Acquaintance Scrutability thesis (all truths are a-scrutable from super-rigid truths plus indexical truths). Suppose all truths are n-scrutable from the A-truths, which are super-rigid. Then by Apriority/Necessity, all super-rigid truths are a-scrutable from the A-truths. By Acquaintance Scrutability, all truths are a-scrutable from these super-rigid truths plus indexical truths, so all truths are scrutable from A-truths plus indexical truths.

As in Chapter 8, one can deny this link from n-scrutability to a-scrutability by denying one of the key premises. Some theists, some ontologists, and some type-B materialists may deny the Apriority/Necessity thesis, while other type-B materialists may deny the Acquaintance Scrutability thesis. Still, these theses have significant support. I have argued for a variant of the Apriority/Necessity thesis at length in Chalmers 2009, and for the Acquaintance Scrutability thesis more briefly in chapter 8.

What about the right-to-left direction of the Linking Thesis? If all truths are a-scrutable from super-rigid A-truths plus indexical truths, then are all truths n-scrutable from the A-truths? We could derive this claim by assuming the Apriority/Necessity thesis along with the auxiliary claims

that (i) if a super-rigid truth is a-scrutable from the super-rigid A-truths plus indexical truths, it is a-scrutable from the A-truths alone, and (ii) all truths are n-scrutable from super-rigid truths.

The first auxiliary claim is a consequence of the rules for indexical truths in scrutability bases (in the fifth excursus). If a super-rigid truth  $S$  is a-scrutable from ‘I am  $\phi$ ’ and super-rigid A-truths, it will be a-scrutable from ‘Something is  $\phi$ ’ and A-truths. The rules require that ‘Something is  $\phi$ ’ is scrutable from the non-indexical truths in the base in any case, so it will be scrutable from A-truths. So  $S$  will be a-scrutable from A-truths.

The second auxiliary claim, Super-Rigid Necessitation (discussed briefly in E9) is not obvious, however. Potential counterexamples will arise on haecceistic views (Adams 19xx), on which certain truths about concrete objects are not necessitated by underlying “qualitative” truths. On such a view, there can be a world that is qualitatively identical (microphysically and phenomenally identical, for example) to our world but in which different objects exist: where our world contains Obama, perhaps, the other world contains Twin Obama. Given the plausible claim that there are no super-rigid expressions that refer to concrete objects, so that super-rigid truths are all qualitative, the actual truths about Obama will not be necessitated by super-rigid truths.

One could reply by simply denying the relevant haecceitistic view, which is controversial. But if one accepts haecceitism, one can weaken the Linking Thesis by retreating to the claim that if all truths are a-scrutable from super-rigid A-truths plus indexical truths, then all *qualitative* truths are n-scrutable from A-truths. Here qualitative truths are understood to exclude object-dependent truths (this might involve a ban on singular terms, along with certain restrictions on predicates and the like). This weaker claim can then be supported by replacing auxiliary thesis (ii) with a weaker thesis (iii), which we might call Super-Rigid/Qualitative Necessitation: all qualitative truths are n-scrutable from super-rigid truths. Haecceitistic views will not pose an objection to these weaker theses.

Another potential counterexample arises if truths about quiddities are not necessitated by truths about non-quiddities and if there are no super-rigid expressions for quiddities. If one accepts a no-quiddity or a graspable-thick-quiddity view [7.9], one will reject these views. If one accepts a thin-quiddity view or an ungraspable-thick-quiddity view, on the other hand, one may well accept these views. If so, one could always retreat to the claim that if all truths are a-scrutable from super-rigid A-truths plus indexical truths, then all *super-rigid* truths are n-scrutable from A-truths. Alternatively, we can expand the class of qualitative truths above to exclude quiddity-involving truths. In what follows, I will assume the Super-Rigid/Qualitative Necessitation thesis, and readers can adjust the notion of qualitiveness, perhaps to exclude object-involving and/or

quiddity-involving truths, according to their own views of whether this is needed.

Where does this adjustment leave the connection between scrutability bases and supervenience bases? In chapter 8, we saw that there is plausibly an a-scrutability base involving just super-rigid expressions and indexicals. If we assume Apriority/Necessity along with Super-Rigid/Qualitative Necessitation, it follows that these super-rigid truths form a necessitation base for qualitative truths: all qualitative truths will be n-scrutable from the super-rigid truths in such a base.<sup>5</sup> Furthermore, given that the super-rigid truths and indexicals form a minimal scrutability base, the super-rigid truths in question will form a minimal qualitative necessitation base: a minimal class of truths such that all qualitative truths are necessitated by (or n-scrutable from) those truths. So the scrutability base yields a sort of supervenience base: the properties involved in the super-rigid truths will in effect be a supervenience base at least for qualitative properties.

Given this connection between scrutability and supervenience, we can then ask about the connection between scrutability and metaphysical fundamentality. For example: can we conclude that the super-rigid truths in such a scrutability base are the metaphysically fundamental truths? There are a few obvious obstacles to this thesis: one involving nonqualitative truths, one involving inexpressible properties, and one involving metaphysical priority. But addressing these obstacles can help us to better understanding the connection between scrutability and metaphysical fundamentality.

The first obstacle is posed by nonqualitative truths. We know that the super-rigid truths in question necessitate all *qualitative* truths, but one might think that the metaphysically fundamental truths should necessitate *all* truths. Matters are not entirely clear here, however. In practice, many philosophers at least implicitly take it that necessitation of object-involving truths is not required. For example, physicalists often allow that microphysical truths do not necessitate object-involving truths (that this, they allow that there are microphysically identical possible worlds involving different objects) without taking this to threaten physicalism. The issue is subtle. If the stronger thesis is required, then the move from n-scrutability to fundamentality will require either ruling out haecceitism or else fleshing out the n-scrutability base with certain object-involving truths (object-involving truths about certain microphysical objects, for example) so that the base becomes a full necessitation base. For present purposes, however, I will take it that at least one interesting sort of metaphysical fundamentality is compatible with failure to necessitate non-object-involving

---

<sup>5</sup>This claim requires only the less controversial direction of the Apriority/Necessity thesis: if a super-rigid sentence is a priori, it is necessary. The claim about minimality in the next sentence requires both directions, however.

truths.<sup>6</sup>

An analogous worry arises if there are no super-rigid expressions for quiddities, as on views with ungraspable thick quiddities and with thin quiddities. On these views, quiddistic truths will not be necessitated by the super-rigid truths in a scrutability base. Most believers in quiddities take at least some of them to be metaphysically fundamental, so this problem cannot be dismissed as with haecceities above. Rather, the super-rigid truths involved in a scrutability base will have to be augmented by non-super-rigid truths concerning quiddities in order to yield an n-scrutability base for all truths. Then the truths involved in this n-scrutability base (or perhaps a minimal subset of it) may well be metaphysically fundamental, at least as far as quiddities are concerned.

The second obstacle is posed by the possibility that certain metaphysically fundamental truths are inexpressible. We have already dealt with views on which certain truths are not expressible in super-rigid terms, but now the issue is true propositions that are not expressible by sentences at all. Quiddities and haecceities are plausibly expressible at least non-super-rigidly, but someone might hold that there are fundamental properties in other realms that we cannot refer to at all. If there are such properties, then there will also be inexpressible propositions concerning them. Then our n-scrutability base for sentences will not yield an n-scrutability base for propositions, and will not yield a base of metaphysically fundamental properties. Still, if we make the fairly weak assumption that we can refer to all fundamental properties, then (given (iii)) truths about these properties will either be in our n-scrutability base or will be n-scrutable from our base, and this obstacle will be removed. If we make the stronger assumption that we can refer super-rigidly to all fundamental properties, then we do not need (iii). Given the Apriority/Necessity thesis and (i), a-scrutability of all truths from super-rigid A-truths plus indexical truths yields a-scrutability of all super-rigid truths by A-truths (by (i)), which yields n-scrutability of all super-rigid truths by A-truths. Given the assumption, all fundamental true propositions will be expressible by super-

---

<sup>6</sup>See Hofweber 2005 and Almatohari and Rochford forthcoming for differing perspectives on this matter. My view is that even if object-involving truths are not necessitated by underlying qualitative truths, they may nevertheless be grounded in underlying qualitative truths. For example, suppose there are just two particles. Then the fundamental truth about the world might take the form  $\exists x\exists y : x \neq y \& Fx \& Gy$ . There may also be object-involving truths about this world of the form  $Fa$  and  $Gb$ , but I do not think it is compulsory to see  $Fa$  and  $Gb$  as the fundamental truths here. Instead, they may themselves be grounded in the existential truths. This requires rejecting the standard view that existential truths are always grounded by object-involving truths. This view is consistent with a haecceitistic view on which there is a distinct world in which  $Fb \& Ga$ . In effect, once there are objects in our world, we can use them to characterize various counterfactual possibilities involving them, but the original objects are nevertheless grounded in qualitative matters.

rigid truths, which will themselves be n-scrutable from A-truths. Given that fundamental true propositions necessitate all true propositions (perhaps setting aside object-involving propositions), this yields necessitation of all true propositions by A-propositions.

The third and most important obstacle arises from metaphysical priority. Being a member of a minimal necessitation base (even a minimal necessitation base for all true propositions) does not suffice for fundamentality. To see this, we can note that given a nonfundamental truth such as ‘Paris is in France’, there will be many necessitation bases including that truth. Some of these bases will be minimal among this class, in that they do not include any other bases in that class. Some of these bases will have the further property that if one subtracts ‘Paris is in France’, one would no longer have a necessitation base. These bases will be minimal necessitation bases and will include ‘Paris is in France’. But ‘Paris is in France’ is not plausibly fundamental.

This point is sometimes put by saying that where matters of fundamentality are concerned, we need to appeal to a relation more fine-grained than necessitation: the relation of *grounding* (Fine 200x; Schaffer 2009). Here the thought is that B-truths are grounded in A-truths when B-truths hold *in virtue of* A-truths holding. A metaphysically fundamental truth will then be a truth that is not grounded in any other truths. Under certain assumptions, the metaphysically fundamental truths will form a minimal grounding base: a minimal set of truths that ground all truths. Likewise, any minimal grounding base will be the set of metaphysically fundamental truths.

Grounding can be understood as a relation among propositions, facts, properties, or objects. I will use grounding relations among sentences as a stand-in for all of these. If grounding is understood as a relation among true propositions (perhaps Russellian propositions), we can translate by saying that sentence  $S_1$  grounds sentence  $S_2$  iff the proposition expressed by  $S_1$  grounds the proposition expressed by  $S_2$ . One can do the same if grounding is construed as a relation among facts. Grounding relations among properties will correspond to grounding relations among sentences that characterize those properties. Something similar may apply to grounding relations among objects, depending on how those relations are understood. So my talk of grounding relations among sentences can be translated to apply to these other sorts of grounding, though as before we need to keep worries about inexpressibility in mind.

It is tempting to hold that if A-truths ground B-truths, A-truths necessitate B-truths, but this is not entirely obvious (Leuenberger 20xx?). For example, some hold that the collection of fundamental positive truths grounds all truths, both positive and negative, even though it does not necessitate all negative truths. On the view in question, a that’s-all truth needs to be added for necessitation, but this truth is itself grounded in the collection of positive truths. On some haec-

ceitistic views, one might also hold that object-involving truths are grounded in qualitative truths even though they are not necessitated by those truths. I will not take a stand on these matters here. I am more sympathetic with the second point than the first (I am inclined to think that a that's-all truth is itself metaphysically fundamental), but these points will make only a minor difference for present purposes.

More importantly, it is not the case that if A-truths necessitate B-truths, A-truths ground B-truths. For example, if  $A$ ,  $B$ , and  $C$  are microphysical truths, then  $A$  is necessitated by  $A \& (B \vee C)$ , but it is not plausible that  $A$  is grounded by  $A \& (B \vee C)$ . More plausibly, the latter truth is grounded in some combination of  $A$ ,  $B$ , and  $C$ . Likewise, the minimal necessitation base including 'Paris is in France' necessitates all truths but it does not ground all truths. It may even be that some necessary truths, such as mathematical truths, are not grounded by any other propositions, even though they are necessitated by all other propositions. If so, they will be in a minimal grounding base, although they are not in any minimal necessitation base. These phenomena arise because grounding requires a much stronger connection between truths than necessitation.

Because of this, even if one can argue from super-rigid A-truths plus indexicals forming a minimal a-scrutability base to their forming a minimal qualitative necessitation base, one cannot argue directly from here to their forming a minimal grounding base. There will certainly be non-fundamental super-rigid truths. Some of these will be in minimal scrutability and necessitation bases without being in minimal grounding bases. At best, we might be able to move in reverse and hold that a minimal grounding base (perhaps with the addition of a that's-all truth) will itself be a minimal qualitative necessitation base and will therefore, if it involves only super-rigid truths, be a minimal scrutability base (with the addition of some indexical truths). That is in effect a version of the argument for Fundamental Scrutability offered earlier.

We have seen how to move from premises about scrutability to conclusions about supervenience and vice versa, at least given certain assumptions. We have also seen how to move from premises about fundamentality to conclusions about supervenience and scrutability. But this leaves open the question raised above: can we move from premises about scrutability to conclusions about fundamentality?

To do this, I think we have to appeal to a more fine-grained relation that stands to scrutability roughly as grounding stands to necessitation. We might call the more fine-grained relation *conceptual grounding*. We have investigated relations in this vicinity when discussing the thesis that all truths are analytically scrutable from truths involving primitive concepts. One might hold that one truth is conceptually grounded in other truths when it is analytically entailed by those truths

and those truths are conceptually prior to it. Or perhaps better, one might understand it in terms of the notion of translucent settling in chapter 9. The discussion in chapter 8 and 9 gives at least some plausibility to the claim that there is a notion of conceptual grounding in this vicinity.

In what follows, I will assume that we have pinned down a conceptual grounding relation, although of course the matter requires a more sustained analysis than I have given. To get a rough grip on it, we can work with the approximate definition picture, so that when  $E$  is approximately definable as  $D$  (under criteria of adequacy that include conceptual priority), truths involving  $E$  are conceptually grounded in truths without  $E$  involving the terms in  $D$ . So about bachelors will be conceptually grounded in truths about gender and about marriage, while truths about electrons will be conceptually grounded in truths about playing the electron role. I will also take it that standard logical grounding relations yield conceptual grounding: so  $A$  and  $B$  jointly ground  $A \& B$ ,  $A$  or  $B$  separately ground  $A \vee B$ , and so on.

The discussion in chapters 8 and 9 makes a case that there is a minimal conceptual grounding base such that all truths are conceptually grounded in those truths. These truths will involve primitive concepts: perhaps some or all of nomic, phenomenal, spatiotemporal, and quiddistic concepts, as well as normative and mathematical concepts, perhaps among others. The inclusion of normative and mathematical truths in a minimal conceptual grounding base makes clear that such a base can go well beyond a minimal scrutability base.

Can one make inferences from claims about conceptual grounding to claims about metaphysical grounding? Certainly, the claim that  $A$  conceptually grounds  $B$  does not seem to be equivalent to the claim that  $A$  metaphysically grounds  $B$ . For example, a claim about a table might be metaphysically grounded by microphysical truths about charge, spin, and the like, but it is not plausibly conceptually grounded in those truths. The truth that an entity has a certain charge may be conceptually grounded in the claim that it has a property that plays a certain role, but (at least on some views) it will not be metaphysically grounded in that truth.

Correspondingly, charge and spin may be metaphysically fundamental, but the concepts *charge* and *spin* are certainly not conceptually fundamental. In the reverse direction, some may hold that *conscious* and *I* are conceptually fundamental, while denying that consciousness and *I* are metaphysically fundamental. So fundamentality of a concept need not go along with fundamentality of its referent.

Still, all these problems also arose when considering the relation between apriority and necessity, and there is a familiar solution. ‘Charge’, ‘spin’, and ‘I’ are not epistemically rigid, so we can restrict the thesis to super-rigid truths, as follows. The case of consciousness is still a potential

exception, to be sure, but this case is controversial, and as before one might use the restricted thesis to argue for the metaphysical fundamentality of consciousness.

Conceptual/Metaphysical (C/M) Thesis: When *A* and *B* are super-rigid, *A* conceptually grounds *B* iff *A* metaphysically grounds *B*.

On the left-to-right direction: it is very plausible that when *A* conceptually grounds *B* for super-rigid *A* and *B*, *A* metaphysically grounds *B*. The obvious candidates for conceptual grounding without metaphysical grounding all involve non-super-rigid expressions: for example, truths involving natural kind terms ('charge', 'electron') or names ('Jack the Ripper'). If we take a conceptually grounded super-rigid expression, such as 'friendly' perhaps, it is highly plausible that those expressions involved in its conceptual grounds (for example, expressions involving certain mental states and dispositions to behave) are equally involved in its metaphysical grounds. Certainly, when *A* is definable super-rigidly as *D* (where super-rigidity excludes devices of a posteriori rigidification and the like within *D*), we can expect that *D*-truths to metaphysically ground *A*-truths: so truths about unmarried males plausibly ground truths about bachelors. Something similar goes for approximate definitions, for translucent grounding, and for logical grounding. So there is a strong prima facie case for the left-to-right direction here.

On the right-to-left direction: this direction is clearly more controversial, as the case of consciousness illustrates. But setting aside that and related controversial cases for now, are there any clear exceptions? One might worry that super-rigid microphysical truths will metaphysically ground super-rigid high-level truths without conceptually grounding them. After all, microphysical truths seem far from being conceptually primitive.

To assess this matter, we should first consider what super-rigid microphysical truths will involve. The main class will be broadly structural truths, cast in terms of logical, mathematical, nomic, and perhaps spatiotemporal vocabulary. An example is the truth that there exists an entity with a property that plays a certain specified nomic role with respect to other properties. On some views there will also be a distinct class of super-rigid quiddistic truths, characterizing intrinsic quiddities of microphysical entities. The broadly structural truths are cast in conceptually primitive vocabulary, and they are plausible candidates to be conceptually primitive truths, not grounded in any further truths: the main exception is that on a quiddistic view, certain existential claims within them (there exists a property that plays a role) may be grounded in a corresponding quiddistic truth (quiddity *Q* plays that role). As for quiddistic truths, quiddistic concepts are

certainly unfamiliar, but it is natural to hold that basic quiddities can serve as conceptual grounds higher-level quiddities.

Next, we should consider what super-rigid high-level truths will involve. Most high-level expressions are not super-rigid, and the super-rigid expressions derive from a limited number of categories: causal, spatiotemporal, mathematical, quiddistic, phenomenal, normative, and a few others. In the case of causal and spatiotemporal high-level truths (truths involving ‘computer’ or ‘square’ perhaps), it is plausible that these truths will be both conceptually and metaphysically grounded in structural microphysical truths. In effect, fine-grained nomic and spatiotemporal microphysical structure will serve as conceptual and metaphysical grounds for coarse-grained macrophysical nomic and spatiotemporal structure. In the case of high-level quiddistic truths (if any), these are again unfamiliar, but there is no obvious reason to doubt that if they exist, they are both conceptually and metaphysically grounded in microphysical quiddistic truths.

In the case of pure mathematical truths, it is arguable that these are neither conceptually nor metaphysically grounded by microphysical truths. On the face of it, the microphysical truths are simply irrelevant to pure mathematical truths, and play no role in grounding them. One could hold that mathematical truths are conceptual truths and that conceptual truths need no grounds: they are not fundamental, but they are conceptually grounded in an empty base. If they are not conceptual truths, however, then one could hold either that they are metaphysically grounded in an empty base, or that they are not metaphysically grounded in microphysical truths at all. I think that the last view is perhaps the most plausible of these options.

As for phenomenal truths: of course type-B materialist views will reject the Conceptual/Metaphysical Grounding thesis, just as it rejects the Apriority/Necessity thesis, but we can set that view aside for now. On type-A materialist views such as analytic functionalism, phenomenal truths (if super-rigid at all) will be both conceptually and metaphysically grounded in structural microphysical truths. On dualist views, phenomenal truths will be grounded in neither way in microphysical truths. On a Russellian monist view, phenomenal truths are metaphysically grounded in certain quiddistic truths: either phenomenal or protophenomenal truths. It might seem odd to suggest that familiar phenomenal truths are conceptually grounded in much less familiar protophenomenal truths; but this may be not much more odd than the plausible claim that truths about phenomenal color are conceptually grounded in truths about phenomenal hue, saturation, and brightness. It is certainly possible for unfamiliar primitive concepts that play a role in conceptual grounding truths involving familiar concepts. Here it is worth keeping in mind that primitive concepts may be quite different from the concepts that we first acquire.

An especially tricky case is that of normative truths. Basic normative expressions are arguably super-rigid. If one is a naturalist normative realist, one will hold that normative truths are metaphysically grounded in non-normative truths. But we have seen that (setting aside normative descriptivism and the like) it is arguable that normative truths are not conceptually grounded in non-normative truths. One could respond by embracing normative irrealism, normative non-naturalism, or normative descriptivism. I am inclined to think that the moral is that one should be either a normative irrealist or a weak sort of normative non-naturalist who holds that normative truths are partly grounded (conceptually grounded and metaphysically grounded) in fundamental normative moral principles, which are not themselves conceptually grounded or metaphysically grounded in non-normative truths (although they are necessary and therefore necessitated by those truths). In this way, fundamental normative truths are akin to fundamental mathematical truths. But if one rejects non-naturalism, irrealism, and descriptivism, one may need to allow that there are cases of super-rigid metaphysical grounding without conceptual grounding.

Something very similar goes for the case of ontological truths. We saw earlier (chapter 6) that some ontologists hold that the existence of a mereological sum (say) is necessitated but not a priori entailed by the existence of its parts. Likewise, some will hold that the existence of the sum is metaphysically grounded but not conceptually grounded in the existence of its parts. I am inclined to reject these views, holding that one should be either an ontological irrealist (so there is no truth to ground) or an ontological deflationist (so the truth is conceptually grounded in truths about the parts). An ontological realist could also preserve the C/M Thesis by endorsing ontological nonreductionism and holding that the existence of the sum is not entirely metaphysically grounded in truths about its parts. But if one rejects these three views, one might allow that there are cases of super-rigid metaphysical grounding without conceptual grounding.

Overall, the moral of this discussion is that there are no clear exceptions to the C/M Thesis. Certain philosophical views entail the existence of exceptions: type-B materialism, some strong forms of naturalist normative realism, and some strong forms of ontological realism. But these views are all controversial and far from obviously correct (although to be fair, their negations are also controversial, and they are far from obviously wrong). Good reasons to accept the C/M Thesis will also be good reasons to reject these views.

I will not try to argue for the C/M Thesis at any length here. I think that one can argue for it in ways parallel to arguments for the Apriority/Necessity thesis. In the latter case, one can argue that any a posteriori necessities involving super-rigid expressions (such as putative necessities connecting consciousness and physical properties) will be brute necessities (Chalmers 1996,

2009). One can likewise argue that any a posteriori, and perhaps any nonconceptual, grounding claims involving super-rigid expressions (such as grounding claims connecting consciousness and physical properties) will be brute grounding claims. And one can argue that there can be no brute necessities and no brute grounding claims. More strongly, one can argue that our modal concepts are grounded in epistemic concepts, so that we do not have a grip on a notion of metaphysical necessity that is not tied to epistemic necessity as the Apriority/Necessity thesis suggests. In the same way, one can argue that we do not have a grip on a notion of metaphysical grounding that is not tied to conceptual grounding as the C/M thesis suggests. For now, however, I simply note that the C/M Thesis remains on the table as a highly attractive view about grounding.

If the C/M thesis is true, then a minimal conceptual grounding base for super-rigid truths is also a minimal metaphysical grounding base for super-rigid truths, and vice versa. If the right-to-left half of the C/M thesis is false but the left-to-right half is true, then a conceptual grounding base for super-rigid truths is also a metaphysical grounding base for super-rigid truths (although a minimal conceptual grounding base need not be a minimal metaphysical grounding base).

What about the stronger claim that any conceptual grounding base for *all* truths is also a metaphysical grounding base for all truths and vice versa? This does not follow immediately from the C/M Thesis, as we now have worries about non-super-rigid expressions to contend with. For the left-to-right direction, the biggest worry concerns object-involving truths. (Related issues arise for other non-super-rigid truths such as kind-involving truths, but the issues are parallel.) For example, perhaps existential truths such as '*ExFx*' collectively serve as conceptual grounds for singular truths such as '*Fa*' [7.10], while the latter collectively serve as metaphysical grounds for the former. If so, conceptual and metaphysical grounding bases will look quite different. To respond, one could take the line discussed earlier according to which even a metaphysical grounding base involves the existential truths here. Alternatively one could weaken the thesis to the claim that any conceptual grounding base *corresponds* to a metaphysical grounding base, where correspondence requires replacing existential truths by singular truths of an otherwise similar form.

As for the right-to-left direction, there is an obvious worry about indexical truths. Given the C/M thesis, a super-rigid metaphysical grounding base for all super-rigid truths will also conceptually ground all super-rigid truths, but it will not conceptually ground indexical truths. One needs to add indexical truths to obtain a full conceptual grounding base. If the right-to-left direction of the C/M thesis is false, one may need to add further truths (perhaps phenomenal truths, normative truths, and so on) to obtain a full conceptual grounding base.

So if the C/M thesis is correct, a minimal conceptual grounding base will not be a minimal

metaphysical grounding base: one will have to subtract indexical truths for that purpose. One may also have to convert existential truths to singular truths, depending on one's view of the role of these truths in grounding. In the reverse direction, to go from a minimal metaphysical grounding base to a minimal conceptual grounding base, one will need to add indexical truths, and perhaps convert singular truths to existential truths. If the C/M thesis is false, one will need to add or subtract further truths (such as phenomenal truths and normative truths) along with the indexical truths. On my own view, the only difference between the two bases will be the inclusion or exclusion of indexical truths. These aside, metaphysically fundamental truths will be conceptually fundamental truths and vice versa.

It might seem surprising to say that metaphysically fundamental truths, such as those in physics, are conceptually primitive truths. But once one reflects on the fact that metaphysically fundamental truths in physics will themselves either involve quiddities (perhaps with nomic and spatioemporal links) or else nomic profiles, powers, and the like, this no longer seems so surprising. If there are concepts of these quiddities at all, they will be novel concepts and we should not be surprised that they are primitive. Concepts of nomic profiles and powers, expressed in an appropriately structural way, themselves appear to be good candidates to be conceptually primitive truths. And of course if phenomenal truths are metaphysically fundamental, it is no surprise that they should also be conceptually fundamental.

Given this close a connection between metaphysical and conceptual grounding, it is natural to ask about the relation between the two: are the two identical, is one grounded in the other, or are they more independent than that? An unrestricted identity thesis seems unlikely, because of the way the two relations come apart from non-super-rigid truths. But grounding theses are still on the table. It is not out of the question that the most basic principles of metaphysical grounding are themselves conceptual truths. For example, one could hold that it is a conceptual truth that all true propositions  $p$  metaphysically ground propositions  $p \vee q$ , thereby explaining the metaphysical grounding claim above via a conceptual truth. Then one could argue that less basic truths about metaphysical grounding themselves follow from conceptual truths and fundamental truths.

This line of thinking suggests the intriguing idea that conceptual truths along with fundamental truths conceptually ground all truths about metaphysical grounding, and thereby metaphysically ground those truths. If we see conceptual truths as corresponding to conceptual grounding claims, we might put this pithily as: conceptual grounding grounds metaphysical grounding. One could then suggest that conceptual truths do not themselves require explanation or grounding (perhaps they are grounded in the empty set). If so, this provides perhaps as good an explanation of meta-

physical grounding as one might get.

On another intriguing view, metaphysical grounding grounds conceptual grounding. For example, one could hold that at least for non-indexical acquaintance concepts, to grasp the concept depends on being acquainted with its referent. Then one could suggest that grounding relations about the concepts reflect metaphysical relations among the referents. For example, a hue concept may ground a color concept in virtue of a hue property grounding a color property. If so, then conceptual grounding relations among the concepts are grounded in metaphysical grounding relations among the properties.

Both of these views are attractive, and I do not know which of these two views is correct. It would not surprise me if elements of both of them are correct. Either way, there will be a close and even constitutive connection between conceptual grounding and metaphysical grounding.

An opponent might object that metaphysical grounding is mind-independent while conceptual grounding is mind-dependent, so the two cannot be as closely connected as this. One response here would be to adopt the broadly Kantian idea that metaphysical grounding is itself mind-dependent and depends on our contingent cognitive scheme. I am inclined to the opposite response, however: conceptual grounding is mind-independent. That is, conceptual grounding relations among truths do not depend on our cognitive apparatus at all. On this view, the primitiveness of a concept is not a fact about humans. Of course our grasping of these concepts is mind-dependent, as are the beliefs we form with them. To the extent that we are well-functioning, the relations among concepts may be reflected in various contingent cognitive relations in us. But there are mind-independent truths about conceptual relations, just as there are mind-independent truths about numerical relations.<sup>7</sup> If this is right, conceptually-guided metaphysics can lead us to mind-independent metaphysical truths.

Overall, we have seen that the relationships between scrutability, supervenience, and grounding are complex, but they can be drawn. The most important principles in drawing these connections are modal/epistemological bridging principles for super-rigid truths. The Apriority/Necessity Thesis connects supervenience and scrutability, while the C/M Thesis connects conceptual and metaphysical grounding. Smaller obstacles along the way include structural differences, worries about non-super-rigidly expressible propositions, and the status of indexical and that's-all truths. But these smaller obstacles can be handled in reasonably straightforward ways, leaving a fairly strong connection between the theses in place. If all this is right, we may truly say that scrutability and

---

<sup>7</sup>This not to endorse platonism about concepts, any more than to hold that it is a mind-independent truth that  $2 < 3$  is to endorse platonism about numbers.

conceptual grounding are guides to the structure of the world.

Of course I have not argued at any length for the Apriority/Necessity thesis or the C/M thesis. I have just tried to make the case that they have some plausibility, and that standard worries about the connection between apriority and necessity (and so on) are not worries for these theses. I have argued for the former thesis elsewhere, and I am more confident of it than of the latter thesis. But I think that both deserve further investigation. I leave that investigation to future work.

## Twentieth Excursus: Reference Magnets and the Grounds of Intentionality

A recently popular idea is that especially natural properties and entities serve as “reference magnets”. Expressions refer to these properties at least partly in virtue of their naturalness. This idea sometimes serves as part of a radically externalist approach to reference and content, on which content is fixed by worldly factors that are largely independent of our cognitive apparatus. I think that there is considerable interest in the idea of reference magnetism, but there is little reason to believe in the radically externalist versions of the idea. Still, an analysis of the core issues here can help us to shed light on the grounds of intentionality more generally.

The locus classicus for the idea of a reference magnet is David Lewis’s article “Putnam’s Paradox”. There, Lewis proposes the idea as a response to Hilary Putnam’s “model-theoretic” argument against realism. Putnam’s argument is itself closely related to Newman’s problem, discussed in Chapter 1 and again in the appendix to chapter 7.

As Lewis puts things, Putnam’s argument is an argument against a certain theory of reference: *global descriptivism*.<sup>8</sup> Let us say that an expression *E* refers wholly by description when there is an associated description ‘the *D*’ such that *E* refers to some entity wholly in virtue of that entity’s being the referent of a description ‘the *D*’ that is associated with *E*. A description ‘the *D*’ is associated with *E* when ‘*E* is the *D*’ is the core of a theory held by users of *E*. For example, ‘electron’ refers by description if its referent is constrained to be the entity that plays such-and-such a role (in physics and chemistry, say), where ‘electrons play such-and-such a role’ is the core of users’ theory of electrons.

Here a “theory” need not be especially theoretical: a theory can be any set of sentences that speakers (perhaps “expert” speakers in some cases) are disposed to accept. The notion of a core part of a theory can be understood in various ways. One obvious suggestion is that the core theory associated with *E* involves only those sentences involving *E* that are a priori (or perhaps analytic), reflecting those principles that hold over any epistemically possible scenario that users of *E* might consider. When an expression refers wholly by description, we might say that the only

---

<sup>8</sup>Putnam clearly intends his argument to be an argument against an enormous number of theories of reference, not just global descriptivism. But the argument proceeds by arguing in effect that all theory-external constraints on reference can be turned into theory-internal constraints (this is Putnam’s “just more theory” point). So one might see the argument as proceeding by arguing that any theory of reference can be turned into a version of global descriptivism, and then invoking Newman’s argument against the latter.

constraints on its referent are *theory-internal*, deriving wholly from speaker's associated beliefs and dispositions to judge.

According to *local descriptivism*, discussed earlier, many or most expressions refer wholly by description, while some other nonlogical expressions, the primitive expressions, do not refer by description but refer in some other way. Global descriptivism says that *all* expressions, or perhaps all nonlogical expressions, refer by description: in effect, the referent of every (nonlogical) expression is determined by some associated theory. Where local descriptivism has both theory-internal constraints on reference (for descriptive expressions) and theory-external constraints (for primitive expressions), global descriptivism says that the only constraints on reference are theory-internal.

Global descriptivism suffers from a by-now familiar problem: Newman's problem. If every nonlogical expression refers by description, then the referent of every expression is determined by an overall theory that can be put using logical expressions alone, and such a theory is almost vacuous. As long as the world has an appropriate cardinality, the theory will be satisfied. Furthermore, it will be satisfied in innumerable ways: any expression can be mapped on to any entity of the appropriate category (so a name can be mapped to any object, a predicate to any property) in a way that satisfies all the constraints of global descriptivism: that is, such that all relevant theoretical claims of the form '*E* is the *D*' come out true.

Lewis draws the conclusion that global descriptivism is false: not all expressions refer wholly by description. It follows that for at least some expressions, one needs nondescriptive constraints on reference. Put differently, one needs theory-external constraints: constraints on an expression's referent that do not derive from users' associated beliefs and dispositions to judge.

At this point, Putnam's argument objects that any further constraints here are "just more theory". In effect, the point is that these constraints are then theory-internal, so that the Newman-style argument applies equally to them. But the mere fact that such constraints can be turned into a theory does not make them theory-internal. The relevant point is that when constraints are theory-external, their reference-determining role in no way derives from speakers' acceptance of the theory, or from their inclination to make judgments in accordance with the theory. Rather, it derives simply from the truth of the theory.

Lewis considers and rejects one candidate for a theory-external constraint: the causal constraint, holding that an expression's referent should stand in an appropriate causal relation to the expression's uses. Lewis takes the role of causation to be theory-internal. On this view, when the causal constraint plays a role, it does so in virtue of being part of speakers' associated theory. For

example, it is part of speakers' tacit theory that Gödel is responsible for our use of 'Gödel': Lewis notes elsewhere that the causal constraint is revealed precisely through speakers' judgments about cases. For other expressions, the constraint is not present at all. Lewis takes all this to recommend causal descriptivism: the thesis that relevant expressions *E* are associated with a description 'the *D*' that involves causation, such as (for example) 'the entity causally responsible for my use of 'Gödel''. Given that causation plays a theory-internal role, we have to look elsewhere for a theory-external constraint.

Instead, Lewis endorses another theory-external constraint: the *naturalness* constraint, holding that an expression referent should be relatively natural. Here an entity is perfectly natural when it is fundamental: the perfectly natural entities are the fundamental entities of physics. An entity is more natural than another when it is closer to fundamental: when it can be defined more easily in terms of fundamental entities, perhaps. Then the naturalness constraint says roughly that where there are multiple candidates to be an expressions' referent that are equally well-qualified on other grounds (such as fitting theory), the most natural candidate is the referent. Or considering all expressions at once: the referents of all expressions are determined to be that assignment of entities to expressions that optimizes the combination of naturalness and theoretical fit.

This view has come to be known as *reference magnetism*.<sup>9</sup> The idea is that relatively natural entities serve as "magnets" for expressions to refer to: they are intrinsically more eligible than less natural entities to be referents. When any expression refers to an entity, it does so partly in virtue of relative naturalness: perhaps the relative naturalness of the entity itself, or at least in virtue of the relative naturalness of other entities that serve as referents of related expressions. According to Lewis's *strong reference magnetism*, this naturalness constraint is theory-external and is the only theory-external constraint on reference.

Even if one accepts that naturalness plays a role in reference, there is an obvious alternative to strong reference magnetism: namely *weak reference magnetism*, according to which naturalness plays only a theory-internal role. On this view, it is part of our core theory of electrons that electrons are relatively natural, and when naturalness plays a role in determining a theory's reference, it does so in virtue of being part of core theory like this. This view leads naturally to a counterpart of Lewis's causal descriptivism: *naturalness descriptivism*, where naturalness is part of the descriptions associated with relevant expressions. For example, the description associated with 'electron' might be 'The relatively natural entity that plays such-and-such role'.

---

<sup>9</sup>See Hawthorne, Sider, etc.

Interestingly, each of Lewis's reasons for rejecting a theory-external role for causation applies equally to naturalness. On the face of it, the naturalness constraint applies much more in some cases than others. For example, while perhaps 'mass' is constrained to be natural, 'weight' is less constrained (if it were constrained further, it might presumably pick out mass!), while expressions such as 'Canberra' seems less constrained again. This reflects the fact that we are more disposed to regard mass *as* natural. The variation in the role of naturalness and way this variation reflects variation in our attitudes and dispositions already strongly suggests a theory-internal role.

Furthermore, just as the causal constraint is revealed in speakers' judgments about scenarios, so is the naturalness constraint. At least, they are reflected in speakers' informed and reflective judgments: judgments in which speakers are informed about relevant empirical facts (although not semantic facts) and in which they are appropriately rationally reflective. Take a case in which the naturalness constraint seems operative: for example, 'gold' refers to the element *Au* rather than the disjunctive kind involving both *Au* and fool's gold, in part because the element is more natural. Users of the expression were in effect presented with the empirical facts about this scenario and made the judgment for themselves that gold is the simple kind *Au*. Likewise, the fact that 'plus' refers to the simple operation of addition rather than quaddition is reflected in our reflective judgments about specific pairs of numbers. All this tends to suggest that naturalness is part of speakers' theory associated with these expressions, and that to the extent that it plays a role in determining their referent, it plays a theory-internal role.

So insofar as Lewis's arguments recommend causal descriptivism is over a theory-external causal theory of reference, arguments of the same sort recommend naturalness descriptivism over a theory-external reference magnetism. This might seem to be merely a *tu quoque* against Lewis—plenty reject causal descriptivism, after all—but I think considerations like this yield a powerful case against strong reference magnetism.

One can put the case as follows. Proponents of reference magnetism (e.g. Sider 20xx) often say that considerations of naturalness can "trump" considerations of theoretical fit: given an entity *E* with an associated theory *T*, it can happen that entity *a* fits *T* better than *b*, but that in virtue of *b*'s greater naturalness, *b* is the referent of *E*. If the theory *T* is constrained to be one in which naturalness plays no role, even weak reference magnetism supports this conclusion. Let us say that *thin theoretical fit* is theoretical fit when a role for naturalness is included, whereas *thick theoretical fit* is theoretical fit when any theory-internal role for naturalness is included. Then naturalness can certainly trump thin theoretical fit. But the more relevant question is: can naturalness trump thick theoretical fit?

One can approach this question by asking: in cases where naturalness trumps thin theoretical fit, is the trumping reflected in speakers' (informed and reflective) judgments about the case? Or are speakers' reflective judgments themselves trumped by naturalness? The first option is apparent in the case of 'gold': here speakers' reflective judgments are that the expression refers to the element rather than to the disjunctive kind, so insofar as naturalness plays a trumping role, it is reflected in judgments about cases. The alternative second option requires that trumping is not reflected in judgments about cases, so that speakers reflectively judge that *E* refers to a less simple entity *a*, when in fact it refers to *b* due to *b*'s greater naturalness.

We might call the second option *ultra-strong reference magnetism*: there is not just a theory-external role for naturalism, but one that can trump speakers' judgments about cases. This view is theoretically coherent, but I think that there is little reason to think that it is correct. The main linguistic evidence for reference magnetism comes precisely from judgments about cases: speakers judge that expressions such as 'gold', 'mass', and so on refer to relatively simple entities. If ultra-strong reference magnetism is correct, we should expect that informed and reflective speakers will frequently be simply wrong about the referents of their expressions, and correspondingly about the truth-values of their sentences. To pick an extreme case, due to John Hawthorne (?): it could be that although we take 'Europe' to refer to a highly convoluted area, the expression in fact refers to a rectangular area, so that our judgments about the borders of Europe are largely wrong. I think that this is clearly an unattractive view. If there are reasonable alternatives to this view, we should embrace them.<sup>10</sup>

A nearby view might allow that the truth of reference magnetism will be available to sufficiently informed and reflective speakers, who will then be in a position to make correct judgments about cases: once one knows the correct semantic theory, one will know that Europe is rectangular.

---

<sup>10</sup>Some further questions for the ultra-strong view: can simple apparent definitions such as 'vixens are female foxes' can be trumped by naturalness? Can stipulations such as 'glubs are octagonal tables' be trumped by naturalness? If yes, the view is even more bizarre. If no, this suggests that the role of naturalness is not global: this behavior suggests the sort of role one gets on a local descriptivism, where the reference of some expressions is fixed by naturalness and that of others is fixed by descriptions using those expressions.

A related question: are apparently a priori claims such as '*E* is the *D*', where *D* articulates the theory associated with *E* that corresponds to the speakers' informed and reflective judgments about cases, truly a priori? Likewise, are scrutability conditionals such as '*PQTI* → Europe is such-and-such' a priori? If reference magnetism is itself a priori, then presumably some other nearby claims (such as '*E* is the natural *D*') may be a priori instead. Or perhaps reference magnetism is not a priori (just true), and the original putatively a priori claims are only nonconclusively a priori. If so, they are defeasible and may turn out to be false if trumped by naturalness.

If so, naturalness will not trump total theory, but it will at least trump ordinary nonsemantic theory, which is arguably good enough for a theory-external role in a reasonably strong sense. Still, I think a view on which everyone but a few philosophers is badly wrong about the shape of Europe is almost as unattractive as the view in the last paragraph. And once again, the linguistic support for a role for naturalness, just like the support for a role for causation, comes from judgments about cases that are available to ordinary speakers who have never thought about semantic theory. In fact, most of these judgments were originally made by speakers of this sort. So I think this view remains unattractive and unsupported.

Moving to the first option: on this option, cases where naturalness trumps thin theoretical fit are always reflected in speakers' (informed and reflective) judgments about cases, and naturalness never trumps those judgments themselves. On this option, naturalness does not trump thick theoretical fit. Instead, the role of naturalness in determining reference is always reflected in a theory-internal role.

The problem for this option is clear: if naturalness never trumps thick theoretical fit, the obvious conclusion is that naturalness plays only a theory-internal role. The role of naturalness can be encapsulated in naturalness descriptivism: speakers' theories involve claims of the form '*E* is the natural *D*'. Given this much, then at least once the referent of 'natural' is fixed so that it picks out naturalness, any theory-external role for naturalness is rendered redundant.

A strong reference magnet theorist has a couple of options remaining. They might say that although the naturalness constraints are part of speaker's theories associated with expressions, their being part of associated theories plays no role in determining reference. If one has any sympathy for a role for theories or descriptions in determining reference, however, this view seems unattractive. The naturalness constraint meets the standard requirements for being part of a reference-determining theory: in particular, it appears to be reflected in speaker's a priori judgments about cases. Alternatively, the theorist might say that naturalness plays both a theory-internal role and a theory-external role, where the theory-external role is redundant so that reference is overdetermined. This view leaves it unclear why we should posit a theory-external role in the first place, though.

Perhaps best, the reference magnet theorist might say that although naturalness never trumps theory, theory leaves certain indeterminacies (highlighted by Newman's problem), yielding ties between candidates that naturalness can settle. Thick theories may determine descriptions such as 'the *D*' or 'the natural *D*' associated with every expressions, but naturalness is still needed to settle the overall assignment of referents for the language as a whole.

Still, it is at least suspicious that the external role for naturalness is reflected so neatly in its theory-internal role. The “no trumping, just tie-breaking” view tends to suggest a picture on which naturalness is needed only to settle the reference of a few basic expressions, and then theory can do the rest. In fact, given a view on which naturalness is the only theory-external constraint and on which this constraint is always reflected in theory-internal descriptions, it suffices for the external role for naturalness to settle the referent of precisely one expression: namely, ‘natural’. This leads to a picture where naturalness is itself an ultra-natural property, one that is picked out theory-externally by the term ‘natural’ precisely because it is so natural. Once this is done, the theory-internal role of naturalness can do the rest. This picture is aesthetically pleasing, but it is certainly odd. I think that around this point, it is natural to look elsewhere.

The dilemma here for strong reference magnetism is paralleled by a similar dilemma for a strong causal theory of reference, on which causal constraints play a theory-external role. We can ask once again whether these constraints can trump reflective and informed judgments about cases. If they do, we arrive at an ultra-strong causal theory of reference, which is as unsupported by linguistic evidence and has as bizarre consequences as ultra-strong reference magnetism. If they do not, then the natural conclusion is that causation plays only a theory-internal role, as causal descriptivism holds. The option remains open that it plays parallel theory-internal and theory-external roles, but this option is unattractive as above: it seems that causation need only fix the referent of ‘causation’, and then theory can do the rest. So I think that just as weak reference magnetism is more plausible than strong reference magnetism, a weak causal theory of reference is more plausible than a strong theory.

Of course Newman’s problem tells us that some theory-external constraints on reference are needed. But it makes sense to seek theory-external constraints that are not paralleled so directly by theory-internal constraints.

One clue about such constraints comes from the role of theory itself. Even reference magnetism holds that reference is determined by theoretical fit plus naturalness. So the question arises: in virtue of what are expressions associated with descriptions or theories? Here further external constraints seem to be needed. It is most natural to say that an *E* is associated with a description *D* in virtue of its inferential role, or in virtue of speaker’s dispositions. But ‘inferential role’ and ‘speaker’s dispositions’ play very little role in our ordinary theories: “electrons refer to whatever satisfies the inferential role associated with ‘electron’” is not folk theory revealed in judgments about electrons in cases. And even if these were part of our theories, the question of how the theories get to be associated with expressions still arise. At the very least, we would need some

mechanism to associate ‘inferential role’ with inferential role, and it is not plausible that naturalness or causation could do this alone: some role for theory would still be needed. I take the moral to be that external constraints are needed to associate expressions with theories, and that inferential role is one candidate for such an external constraint. (This provides another reason to reject Lewis’s strong reference magnetism on which naturalness is the only theory-external constraint, and likewise for an analogous strong causal theory of reference.)

Another clue comes from considering a view that Lewis sets aside: local descriptivism. For all Newman’s problem suggests, it remains plausible that the reference of very many expressions is fixed by description, and the reference of a few primitive expressions is fixed in some other way. We have seen that models with primitive concepts are attractive in numerous respects, so one might want to look for plausible theory-external constraints that work specifically for these concepts. One may need give some theory-external role to the expression relation, in order to get from primitive expressions to associated primitive concepts (that is, to ground linguistic content in associated mental content). But it is natural to hope that there might be some especially simple story about fixing the reference of primitive concepts.

Even if one rejects definitional models, as I do, the scrutability model suggests a similar moral. On this view, most expressions can be associated with intensions across scenarios, where the scenarios can be specified using a few primitive expressions. If one accepts that these intensions fix reference for the complex expressions, then in effect reference is grounded in (i) whatever fixes the intensions of complex expressions and (ii) whatever fixes the reference of simple expressions. Or if we aim to ground linguistic content in mental content, reference may be grounded in (i) whatever associates expressions with concepts (construed as mental representations), (ii) whatever fixes the intensions of nonprimitive concepts, and (iii) whatever fixes the reference of primitive concepts. There are three corresponding roles for theory-external constraints here. The first role is presumably played by some sort of expression relation (perhaps an appropriate causal/intentional relation), and the second role is plausibly played by something like inferential role. But this leaves open what plays the third role.

What fixes the reference of primitive expressions? We have seen that these are plausibly narrow, and mainly super-rigid, so causal connections to the environment do not seem crucial here. Given the previous discussion, there are two natural initial candidates: (i) acquaintance, and (ii) inferential role. On the face of it, acquaintance seems especially suited for the reference of phenomenal concepts, Edenic concepts of color and spacetime, and indexical expressions such as ‘I’ and ‘now’. On the face of it, inferential role seems especially suited for logical and mathematical

expressions, perhaps along with nomic and fundamentality expressions.

Of course 'by acquaintance' is not an especially informative answer to the question of how primitive expressions get their referents. Acquaintance is itself a relation between subjects and referents, and it raises the question: in virtue of what is someone acquainted with something? We might take acquaintance as a primitive relation, in which case this view is not far from primitivism about reference, at least for primitive concepts. Or we might try to explain acquaintance in terms of something more basic, in which case the something more basic will be the more fundamental external constraint.

What might acquaintance be grounded in? Perhaps there might be account in terms of narrow causal role in a cognitive system or inferential role, but I have my doubts, especially given the conceivability of spectrum inversion and the like. I suspect that if a grounding story is to be told, it will give a central role to consciousness. If we are acquainted with qualities such as (Edenic) redness, for example, this is plausibly in virtue of our having certain conscious experiences, including but not limited to perceptual experiences.

This raises the question of the structure of conscious states and what they might be grounded in. If conscious states are grounded in physical or functional states these might thereby also play a role in grounding reference. On my own view, though, conscious states are not grounded this way, and may well be primitive. The question then arises as to whether the primitive structure of conscious states is itself relational. On one intentionalist view, perceptual experience fundamentally involves a relation of awareness to certain qualities. If we take such a relation to be fundamental, then again we are not far from primitivism about the most basic referential relations. On another view, intentional structure is itself grounded in nonintentional phenomenal qualities of consciousness. On that view, these qualities may be serving as the ultimate external constraints. I am inclined to prefer the first view to the second, but I view all of these questions as open.

As for inferential role: a detailed story needs to be told about how logical, mathematical, nomic, and fundamentality expressions might acquire their content through inferential role. These inferential roles are likely to be structural inferential roles, so that content is not grounded in inferential links to other specific concepts, but rather in quite general structural patterns of inference that these concepts are involved in. I do not have the details of such an account, but I do not think it is out of the question that such an account can be developed.

Perhaps the biggest challenge to inferential-role accounts is Kripke's plus/quus argument that non-normative dispositions and inferential roles may underdetermine reference. Kripke argues that the same dispositions and inferences might be associated with a badly functioning plus-user

and a well-functioning quus-user. Even though the former is disposed to accept '65+78=5', their expression '+' still refers to addition. In response, one might give a grounding role to normative facts about inference, or to what speakers *should* infer: perhaps the difference between the plus and quus user in this case is that the 'plus'-user is being irrational and not judging as they should. Of course normative facts about inference may well be grounded in something else in turn, in which case these grounding facts will help to ground the facts about reference, in a theory-external way.

If we do not want to give normativity a role in explaining intentionality, we might instead give a role to naturalness. One might hold that at least in cases involving nonideal subjects whose dispositions are limited or irrational, reference may be fixed to certain entities (such as addition rather than quaddition) in part because they are so natural. This is compatible with the claim above that naturalness never trumps informed and reflective judgments, at least where reflectiveness involves absence of irrationality, but it gives a role to naturalness in trumping irrational judgments, or in fixing reference in cases where speaker's dispositions are silent. If so, we may have a sort of intermediate reference magnetism, on which naturalness plays a large theory-internal role but also a limited theory-external role, as one of a number of theory-external constraints.

All this suggests a picture on which intentionality is grounded partly in acquaintance and/or consciousness (or whatever grounds them), partly in inferential role, and partly in norms or naturalness. In effect, contents for primitive concepts are fixed by consciousness or acquaintance; contents for nonprimitive concepts are fixed by inferential role plus naturalness or norms (with inferential role fixing dispositions to judge and naturalness or norms extending these dispositions into full intensions). Linguistic content is grounded in this intentional content along with the expression relation.

These factors might be incorporated into theories, but in many cases they are not part of speakers' associated theories (unlike the roles for causation and naturalness considered earlier), and in any case they do not play their roles in virtue of being part of these theories. All of these factors will play a theory-external role, but none of them are radically external, in that their role will never trump informed and (ideally) reflective judgments about reference.

An enormous amount of work is required to turn a brief sketch such as this into a full picture of the grounds of intentionality and of linguistic content. But the framework developed here might at least illuminate what needs to be done.