

Chapter 9: Verbal Disputes and Philosophical Progress*

David J. Chalmers

1 Introduction

Carnap held that many key disputes, in philosophy and elsewhere, are verbal disputes. In the latter sections of the *Aufbau* (part V), he tries to dissolve a number of philosophical problems, including the mind–body problem and the problem of the external world, as verbal disputes. This theme runs through Carnap’s later works, from “Pseudoproblems in Philosophy”, which applies this method to epistemological problems, to “Empiricism, Semantics, and Ontology”, which applies it to ontological problems.

I think that more philosophical disputes are substantive than Carnap did, but the dissolution of verbal disputes is nevertheless part of my metaphilosophical religion. The diagnosis of verbal disputes has the potential to serve as a sort of universal acid in philosophical discussion, either dissolving disagreements or boiling them down to the fundamental disagreements on which they turn. If we can move beyond verbal disagreement to either substantive agreement or to clarified substantive disagreement, then we have made philosophical progress.

To make this progress, though, we need some general tools to help us determine when a dispute is verbal and when it is not. The scrutability framework helps to provide these tools. To see the connection, note first that a definitional framework such as Carnap’s, if correct, would provide a natural way of approaching potential verbal disputes. The proponents in a dispute can define their key expressions in terms of a fixed class of primitive expressions. If the definitions differ, then the

⁰This chapter is a revised version of an article published in *Philosophical Review* 120:4, 2011. Thanks to audiences at ANU, Beijing, Bristol, Buffalo, Copenhagen, Georgia State, NYU, Rabindra Bharati, Reading, Rutgers, St. Andrews, Toronto, UC Riverside, Yale, and at conferences of the AAP, GAP, SIFA, and APA between 2005 and 2007. Thanks also to Berit Brogaard, Jim Pryor, Gideon Rosen, Wolfgang Schwarz, Daniel Stoljar, Stephen Yablo, and reviewers for the *Philosophical Review* (as well as others acknowledged in footnotes) for comments, and to Rachael Briggs, Peli Grietzer, Mike Jacovides, Derk Pereboom, Geoff Pynn, and David Wall, among many others, for suggesting examples.

proponents may be differing only verbally. Substantive disputes can in principle be carried on at the level of primitive concepts. Verbal issues about the correctness of definitions for nonprimitive expressions are left to one side.

On the scrutability framework, the absence of definitions makes things harder but not impossible. I will set out a way to diagnose verbal disputes that is inspired by the scrutability picture, along with a way of boiling down high-level substantive disputes to the most basic substantive disputes that underlie them.

Verbal disputes are also interesting as a subject matter for the philosophy of language and mind. Reflection on the existence and nature of verbal disputes can reveal something about the nature of concepts, language, and meaning. Later in this chapter, I will use ideas drawn from the analysis of verbal disputes to help characterize a class of primitive concepts, and to help play the role of something like an analytic-synthetic distinction. Even philosophers who reject an analytic-synthetic distinction often appeal to the notion of a verbal dispute, so if the latter notion can be used to play some of the roles of the distinction and perhaps to vindicate a version of it, that is a conclusion worth noting.

This approach allows us shed light on a key thesis discussed in the previous chapter: that all truths are analytically scrutable from truths involving primitive concepts. In effect, the analysis of verbal disputes helps to characterize a version of this thesis and helps to make it plausible. It is also highly suggestive in the analysis of meaning and content, giving support to a broadly inferential view of these phenomena. In this way the analysis of verbal disputes can simultaneously help us to investigate the roots of the scrutability framework and help us to bring out its philosophical and metaphilosophical applications.

For much of this chapter, I aim to work largely from first principles without explicit use of the scrutability apparatus, although the general picture on which high-level truths are scrutable from more basic truths will be in the background throughout. Attributions of verbal disputes and the use of methods for resolving them are already common in contemporary philosophy and elsewhere, and I will start by trying to make these familiar phenomena explicit. In sections 2 through 6, I characterize verbal disputes, spell out a method for isolating and resolving them, and draw out conclusions for philosophical methodology. In later sections (7 through 9), I use the earlier framework to draw out conclusions in the philosophy of language and mind, and to shed light on the analytic scrutability thesis.

2 What is a Verbal Dispute?

Intuitively, a dispute between two parties is verbal when the two parties agree on the relevant facts about a domain of concern, and just disagree about the language used to describe that domain. In such a case, one has the sense that the two parties are “not really disagreeing”: that is, they are not really disagreeing about the domain of concern, and are disagreeing only over linguistic matters.

Perhaps the most famous case of a verbal dispute in philosophy comes from William James’s *Pragmatism* (1907). A man walks rapidly around a tree, while a squirrel moves on the tree trunk. Both face the tree at all times, but the tree trunk stays between them. A group of people are arguing over the question: Does the man go round the squirrel or not? James presents himself as resolving the dispute with the following magisterial speech:

Which party is right depends on what you practically mean by ‘going round’ the squirrel. If you mean passing from the north of him to the east, then to the south, then to the west, and then to the north of him again, obviously the man does go round him, for he occupies these successive positions. But if on the contrary you mean being first in front of him, then on the right of him then behind him, then on his left, and finally in front again, it is quite as obvious that the man fails to go round him... Make the distinction, and there is no occasion for any farther dispute. (James 1907, 25)

One might question the plausibility of either of James’s proffered analyses of ‘going round’, but in any case the key is the final sentence. Once we resolve an issue about language, the dispute over the nonlinguistic domain evaporates, or at least should evaporate. This potential evaporation is one of the central marks of a verbal dispute.

Sometimes words matter. Disputes over words are sometimes important disputes, when something important rests on matters of linguistic usage. In linguistics and the philosophy of language, for example, words are the primary domain of concern, so that a disagreement over the meaning of ‘round’ cannot simply be set aside. The same applies in some cases where we are studying the users of words. In literary criticism or in history, for example, it might be crucial to know how a given word is used by a given individual or community. In psychology and the philosophy of mind, the use of a word may give us evidence about the concepts that thinkers use and about patterns of thought.

In other cases, the answer to questions about the usage of words has serious practical consequences. For example, if we are arguing over whether a law has been violated, one often needs to

settle the meaning of relevant words. Questions about what falls into the extension of ‘marriage’ and ‘murder’ may in some sense be verbal, but the answer to these questions may make a serious difference to people’s lives. The same goes for contracts and promises. In Monty Python’s argument sketch, when Michael Palin protests “This isn’t argument, it’s just contradiction”, he and John Cleese may be engaged in a verbal dispute about what counts as ‘argument’, but the dispute cannot just be set aside. Resolving the verbal issue is crucial to knowing whether a contract has been fulfilled: “Five pounds for an argument”. Something similar applies to promises and to conventions.

In cases where words have fixed connotations and associations, too, verbal issues often have serious practical import. This applies especially when those connotations are normative. What counts as ‘torture’ or as ‘terrorism’ might be, at one level, a verbal issue that a philosopher can resolve by distinguishing senses. But in a rhetorical or political context, words have power that transcends these distinctions. If the community counts an act as falling into the extension of ‘torture’ or ‘terrorism’, this may make a grave difference to our attitudes toward that act. As such, there may be a serious practical question about what we *ought* to count as falling into the extension of these terms.

Often, however, words do not matter. It often happens that we are concerned with a first-order domain, not with the usage of words, and in such a way that nothing crucial to the domain turns on the usage of words. In this case, a verbal dispute is a *mere* verbal dispute. Mere verbal disputes are usually impediments to understanding. In effect, they are obstacles that we do better to move beyond, in order that we can focus on the substantive issues regarding a domain. I take it that this phenomenon is familiar and ubiquitous in science, in philosophy, and in everyday life.

To make progress here, we need to characterize verbal disputes more precisely. Let us take a dispute where one party utters a sentence S , and another party rejects S , perhaps by uttering a sentence that appears to be S ’s negation. What is it for such a dispute to be verbal?

First, it is worth noting that whether a dispute over S is verbal does not depend on S alone. It also depends on the parties to the dispute, and in particular on their background of agreement and disagreement. The same sentence S can typically be the focus of both verbal and substantive disputes, depending on this background. For example, against the background of agreement on the motion of a squirrel, a dispute over ‘The man goes around the squirrel’ may be verbal, but against a different background (perhaps we agree that the man goes around some animal but disagree about whether it is a squirrel or a mouse), the dispute may be substantive. So we cannot just speak of a “verbal question”, independent of context. For our purposes disputes are individuated in part by

the disputing parties, and background facts about agreement should play a role in characterizing whether a dispute is verbal.

One common proposal for characterizing verbal disputes is the following:¹

A dispute over the truth of S is verbal if S is definitionally equivalent to S_1 for one party and definitionally equivalent to S_2 for the other, and the parties agree on the truth-values of S_1 and S_2 .

This is a reasonable heuristic to use in detecting the presence of a verbal dispute, and it plausibly yields a sufficient condition, at least if we set aside cases in which the parties' agreement over S_1 or S_2 is merely verbal. But this criterion does not plausibly yield a necessary condition. Many expressions do not have clear definitional analyses, so that there will be no S_1 and S_2 available, but a dispute involving these expressions may be verbal all the same.

Consider the following passage from chapter 32 of *Moby Dick*:

I take the good old fashioned ground that the whale is a fish, and call upon holy Jonah to back me. This fundamental thing settled, the next point is, in what internal respect does the whale differ from other fish. Above, Linnaeus has given you those items. But in brief they are these: lungs and warm blood; whereas all other fish are lungless and cold blooded.

Here, there is a clear sense in which Ishmael and Linnaeus do not or at least need not differ on the nonlinguistic issues. For example, Ishmael and Linnaeus might agree that whales are like typical fish in their superficial qualities, while they are like mammals in various biological respects that are relevant for scientific purposes. They simply differ on the broadly verbal matter of whether to use 'fish' for the superficial kind or the scientific kind. But there need not be any clear definitional gloss of 'fish' either for Linnaeus or for Ishmael.

Another suggestion is the following common proposal:

A dispute over S is verbal iff S expresses distinct propositions p and q for the two parties, so that one party asserts p and the other denies q , and the parties agree on the truth of p and $\neg q$.

¹See for example Hirsch 1993, 181, although Hirsch appeals to believed equivalence rather than definitional equivalence and goes on to require the further condition that the parties will not retract their view on reflection. I think that "equivalence" has to be read as something like definitional equivalence here (rather than material or even necessary or a priori equivalence) to avoid counting substantive disputes (in material, necessary, or a priori domains) as verbal.

This is a useful conception for some purposes, but I think it is still an overly narrow conception. Even in James's squirrel case, it is far from obvious that the two parties assert and deny distinct propositions.

Consider the *deferential squirrel case*. This combines the original squirrel case with the stipulation that the quarreling parties both use the expression 'go round' with deference to their linguistic community, so that what they mean by the expression is determined by the expression's meaning in the wider community.² We can also stipulate that the linguistic community uses it with a single relevant meaning (perhaps one of James's two meanings above). If so, then the meaning of 'go round' for both parties will be the community meaning, although at least one of them will be under a misapprehension about that meaning. Then the two parties will use 'The squirrel goes round the tree' to express the same proposition, and one will deny the same proposition that the other asserts. Indeed, one party's assertion may be correct where the other's is incorrect. Nevertheless, the debate still seems clearly to be a verbal dispute, in that the first-order disagreement is grounded in a metalinguistic disagreement about the meaning of 'go round'.

Someone might insist that this debate is not a verbal dispute, as it does not fit the distinct-propositions template of the definition above. Instead it is a nonverbal dispute about what it is for something to go around something else. Now, there is no point getting into a verbal dispute over the meaning of 'verbal dispute'. We can simply note that the dispute between the two parties is a pointless dispute in the same way that paradigmatic verbal disputes are pointless and that resolving the verbal issue about the meaning of 'going round' will resolve their disagreement. We might say that this is a *broadly* verbal dispute, while the definition in terms of propositions defines a *narrowly* verbal dispute.³

Broadly verbal disputes are my main focus in this article, and when I speak of 'verbal disputes' simpliciter I will be referring to broadly verbal disputes. For almost any narrowly verbal dispute, there is a corresponding broadly verbal dispute that is not narrowly verbal: we need only tweak the dispute as above so that the two parties use the key terms with deference to a linguistic community in which those terms are used univocally. On the face of it, the tweaked dispute shares many properties with the original dispute: the sense of pointlessness remains, and the linguistic resolution is as apt as ever.

²See Burge 1979 on incomplete mastery and Putnam 1975 on the division of linguistic labor.

³If an opponent insists that these disputes are not in any sense verbal, because they are disputes over non-metalinguistic matters of fact, we could in principle (adopting a variant of the methodology discussed in section 4) call them "schmerbal disputes". Nothing substantive rests on the term "verbal".

Disputes that are broadly but not narrowly verbal typically arise in cases illustrative of semantic externalism, broadly construed to include cases in which a subject has false beliefs about the meaning of a term. Semantic deference is one source of these false beliefs, and inadequate reflection is another. For example, suppose that *A* and *B* agree that Sue made a false statement that she did not believe to be false and also agree on the moral status of Sue's assertion and other relevant properties. *A* says 'Sue did not lie'. *B* initially says 'Sue lied', believing falsely that 'lie' refers to any false statement, but on reflection comes to accept 'Sue did not lie', through reflection on the concept of lying. Then initially, *A* and *B* need not be having a narrowly verbal dispute: both may use 'lie' to express the same concept. But they are having a broadly verbal dispute all the same: intuitively, they agree on the important facts of the case and are merely disagreeing on whether the word 'lie' should be used to describe it.

A third example: suppose that the epistemic theory of vagueness (Williamson 1994) is correct, so that there is a sharp but unknowable boundary between cases in which a person is tall and cases in which the person is not. And suppose two speakers who accept the epistemic theory are engaged in a dispute about whether someone is tall, against a background of agreement on her height and on other relevant underlying truths. Their dispute might not be a narrowly verbal dispute (assuming that 'tall' in fact has the same meaning for both of them), but it is nevertheless a broadly verbal dispute.

Eli Hirsch (2005, 2009) observes that semantic deference causes problems for some conceptions of a verbal dispute, and proposes an alternative conception along the following lines (paraphrased from Hirsch 2009, 238-40):

Two parties *A* and *B* are having a verbal dispute iff, were *A* and *B* to inhabit an *A*-community and a *B*-community (respectively) in which everyone exhibits the same sort of linguistic behavior that *A* and *B* actually exhibit (respectively), then on the correct view of linguistic interpretation, *A* and *B* would agree that both speak the truth in their own language.

The stipulation about correct linguistic interpretation here might be read as requiring that *A* and *B* know which propositions each of them are expressing. Under this reading, Hirsch's criterion is a counterfactual version of the distinct-proposition criterion: it requires that if *A* and *B* were to inhabit the relevantly uniform *A*- and *B*-communities (and were to engage in the right sort of interpretation), they would agree on the truth of the distinct propositions *p* and $\neg q$ that they would express.

This criterion provides a useful heuristic for cases of semantic deference (if a somewhat imprecise one, as Hirsch acknowledges), but it does not cover the problem cases stemming from other sorts of semantic externalism. For example, it appears to predict that the disputes above over ‘lie’ and ‘tall’ are nonverbal: these cases need not involve deference to a community, so the move to uniform communities will make no difference to the propositions expressed. But these cases are naturally classed with the deferential cases as broadly verbal. They are certainly pointless in the same sort of way, and resolving the verbal issue about what counts as ‘lying’ or ‘tall’ will resolve the disagreement, just as in the original squirrel case and the deferential squirrel case.

There is a simpler characterization of a broadly verbal dispute that has the potential to cover all of these cases. At least to a first approximation, we might say that a broadly verbal dispute is one in which an apparent first-order dispute arises *in virtue* of a metalinguistic disagreement. That is:

A dispute over *S* is (broadly) verbal when for some expression *T* in *S*, the parties disagree about the meaning of *T*, and the dispute over *S* arises wholly in virtue of this disagreement regarding *T*.

For example, in the first case above, the parties disagree (at least tacitly) about what ‘go round’ means, and it is plausible that their first-order disagreement arises in virtue of this disagreement about meaning. In the second and third cases, the parties disagree about what ‘lie’ and ‘tall’ mean, and it is plausible that their first-order disagreement arises in virtue of this disagreement about meaning. This model can handle narrowly verbal disputes, on the plausible assumption (see below) that in the relevant cases subjects disagree about the meaning of *S*, but it can also handle disputes that are broadly but not narrowly verbal in the same fashion.

This gloss on a broadly verbal dispute works best if we assume a version of the common view that whenever speakers use an expression, they do so with beliefs about the expression’s meaning, where these beliefs may be tacit beliefs rather than explicit beliefs. In cases of full competence, it is arguable that speakers use expressions with tacit knowledge of an expression’s meaning. Then any case along the lines above, whether a narrowly or a broadly verbal dispute, will plausibly involve a disagreement about meaning, in that it will involve differing beliefs (perhaps tacit beliefs) about the meaning of a key term: one party believes that the meaning of ‘go round’ is such-and-such, while the other party believes that the meaning of ‘go round’ is something else.⁴

⁴The relevant beliefs for each speaker may concern the community meaning, or the content of the term in the current

The relevant notion of meaning can be left intuitive for present purposes. There are some constraints: it should be a substantive notion of meaning, as two verbally disagreeing parties might nevertheless agree on deflationary meaning-sentences such as “‘round’ means round”. It need not be verbally articulable: in believing that ‘go round’ means such-and-such, a speaker need not be able to articulate an expression corresponding to ‘such-and-such’. It should be a notion of meaning that allows variation in meaning between two users of the same context-dependent term. For example, if I say ‘Michael Jordan is tall’ and you say ‘No, Michael Jordan is not tall’, where my use of ‘tall’ expresses tall-for-an-academic and yours expresses tall-for-a-basketball-player, then we are having a broadly verbal dispute, one that arises in virtue of different beliefs about what ‘tall’ means in the current context. It is also worth noting that the meaning of a general term or a predicate should not simply be understood as its extension: two parties who disagree substantively over whether O.J. Simpson is a murderer will also disagree about the extension of ‘murderer’, but intuitively this does not disqualify them from agreeing on the term’s meaning.

What about “in virtue of”? Here, as elsewhere in philosophy, one may have to take something in this vicinity as primitive, but one can at least offer some clarification. I think this should be understood as an *explanatory* “in virtue of”: the idea is that the metalinguistic disagreement *explains* the apparent first-order disagreement. And the relevant sort of explanation should require something stronger than an arbitrary causal or evidential explanation: it would not suffice, for example, for a metalinguistic disagreement to cause personal enmity that then causes first-order disagreement.

An opponent might reject the in-virtue-of characterization of putative broadly verbal disputes. If one denies that language use requires substantive beliefs about meaning, for example, one might well hold that although the parties disagree over ‘The man goes round the squirrel’, they do not disagree over the meaning of ‘go round’ because they do not have any beliefs about its meaning. And even if one accepts that such metalinguistic beliefs are present, one might deny that they have explanatory priority with respect to the apparent first-order disagreement. The opponent might hold that the disagreement over whether certain cases are cases of going round itself explains the

context (especially in cases of context-dependence). It is arguable that for two speakers to have a dispute at all, they must share a context. Perhaps there is a broad sense in which two speakers can have a disagreement without sharing a context. If we want to account for verbal disagreements in this broad sense, and if an appeal to community meaning is unavailable (perhaps because of context-dependence), we will need to appeal to mere differences in beliefs about meaning—one speaker believes her word means such-and-such (in her current context) while the other speaker believes his word means something different (in his current context)—rather than appealing to a disagreement about meaning per se. Something similar goes if we want to allow that there can be verbal disagreement in a crosslinguistic case.

metalinguistic disagreement, or they might hold that neither dispute is explanatorily prior to the other.⁵

I do not know whether this objection is correct: an analysis of the explanatory priority between metalinguistic and first-order disputes here would require a detailed analysis and would turn on numerous disputed theses. It seems clear that it is at least possible for a metalinguistic disagreement to generate an apparent first-order dispute in the deferential squirrel case. But it is arguable that there could be other versions of the deferential squirrel cases and related cases that are not generated in this way. Even if this is right, I think that in all these cases there is at least an intuitive sense in which it is apt to say that the apparent first-order dispute arises in virtue of a metalinguistic disagreement. For example, in all these cases a certain counterfactual variant of the in-virtue-of thesis appears to be satisfied: if the parties were to agree on the meaning of the key term, their disagreement over *S* would be removed. One might use this gloss to help isolate a sense in which the in-virtue-of thesis is approximately correct, although as with most counterfactual analyses, the gloss would not provide a perfect definition.⁶

Still, for reasons like this, it is best not to take the in-virtue-of gloss as providing a stipulative definition of a broadly verbal dispute. We can instead see the characterization as pointing us toward a salient and familiar phenomenon, rather than delineating its contours precisely. The phenomenon is pinned down partly through cases and partly through resemblance to narrowly verbal disputes. The key feature is the distinctive sort of pointlessness that broadly verbal disputes

⁵Another potential objection holds that apparently substantive disputes might satisfy the definition. One potential example, suggested by Tim Williamson: some substantive disagreements over the law of the excluded middle arguably arise in virtue of disagreement about whether ‘not’ expresses classical or intuitionistic negation. Now, if this were merely a disagreement about actual usage, with both parties allowing that both meanings are possible, the disagreement would plausibly be verbal. But if it is a disagreement about possible usage—that is, about what candidate meanings are even available in principle for use—than it is less obviously verbal. One might hold that in this case the disagreement does not really arise in virtue of a disagreement about the meaning of ‘not’, but instead arises in virtue of a nonmetalinguistic issue about logical truths or about what meanings are available. Or one might refine the gloss to restrict it to disagreements about actual usage that do not correlate with disagreements about possible usage.

⁶The counterfactual variant might say that a dispute over *S* is broadly verbal when for some term *T* in *S*, if the parties were to agree over the meaning of *T*, then they would (if reasonable) agree over the truth of *S*. Like other counterfactual versions of in-virtue-of theses, there would be various counterexamples: for example, substantive disputes in which agreeing over meaning would lead to ancillary changes (resulting in a co-operative personality, say) that would lead to agreeing over *S*, and substantive disputes between unreasonable participants that would be dissolved if the participants were reasonable. Still, the heuristic can be used as at least a rough approximation of the in-virtue-of gloss, and has the advantage of not requiring a view on which every use of an expression involves beliefs about its meaning.

share with narrowly verbal disputes.

To bring out the role of pointlessness, it is worth noting that when faced with a potentially verbal dispute we often ask: what turns on this? For example, what turns on our verdict about whether the squirrel goes round the tree? In the cases discussed above in which words matter, something may turn on the verdict: in these cases we have a verbal dispute but not a merely verbal dispute. And the very existence of a merely verbal dispute may have some uses: in helping to make distinctions and in clarifying the important issues at hand, for example. But setting aside these indirect ways in which a verbal dispute might matter, it typically seems to an outside observer that nothing turns on the verdict: the dispute is pointless. Of course there are pointless disputes that are not verbal, as with a dispute about whether the number of odd digits in the first 10^{100} digits of pi is odd or even. But verbal disputes have a familiar and distinctive sort of pointlessness. I will not try to precisely define this sort of pointlessness, but I will use it as a heuristic guide to the presence of a broadly verbal dispute.⁷ In the next section I develop further heuristics for the detection of broadly verbal disputes.

The current definition characterizes *wholly* verbal disputes, in which apparent first-order disagreements arise wholly in virtue of metalinguistic disagreements.⁸ There are also *partly* verbal disputes, in which an apparent first-order dispute arises partly in virtue of a metalinguistic disagreement and partly in virtue of a substantive non-metalinguistic disagreement. In the dual-proposition framework, partly verbal disputes will arise when one party takes S to mean p , the other party takes S to mean q , and when the parties disagree substantively over p , q , or both. There are also cases of (mere) *verbal agreement*, in which apparent agreement over a sentence is explained partly or wholly by disagreement over the meaning of a key term in the sentence.

⁷An opponent might go on to argue that the disputes over roundness, tallness, and so on are not pointless. I consider opposition of this sort in section 5.

⁸On one understanding of a *merely* verbal dispute, a dispute is merely verbal when it is wholly verbal. One might also require, following a line of thought in the last section, that a merely verbal dispute is one in which usage of the key term is not itself a central object of concern. (Thanks to Carrie Jenkins for discussion here.) We can also distinguish explicitly verbal disputes, in which parties disagree explicitly about metalinguistic claims, from implicitly verbal disputes, in which parties disagree explicitly about first-order claims in virtue of disagreeing about metalinguistic claims without disagreeing explicitly about the latter.

3 Resolving Verbal Disputes

The foregoing gloss on a broadly verbal dispute immediately suggests a method for resolving verbal disputes. Typically, a broadly verbal dispute is one that can be resolved by attending to language and resolving metalinguistic differences over meaning. For example, these disputes can sometimes be resolved by settling the facts about the meaning of key terms in our community: someone might resolve the squirrel dispute by establishing that in their community, ‘go round’ has their favored meaning and not another. They can also sometimes be resolved simply by distinguishing senses of a key term, as James did in the case of ‘round’. By contrast, substantive disputes cannot usually be resolved in this way. For example, a dispute over whether a meteor caused the extinction of the dinosaurs cannot plausibly be resolved by attending to language, whether by settling facts about the meaning of the terms involved or by distinguishing senses for of the term.

Still, resolving disagreements over meaning in this way is often difficult. Settling the facts about the meaning of key terms in our community may take substantive empirical investigation. Distinguishing senses of a key term is particularly difficult when these senses do not correspond to clear explicit definitions. More generally, we are not always able to give a good articulation of what our terms mean, and it is often far from obvious whether or not two speakers disagree about meaning. So it is useful to have a method that does not directly depend on the analysis of meaning in this way.

An alternative heuristic for detecting and dealing with verbal disputes is inspired by the scrutability framework, and in particular by the idea that most truths are scrutable from more basic truths. We might call this heuristic the *method of elimination*. Here, the key idea is that one eliminates use of the key term, and one attempts to determine whether any substantive dispute involving more basic expressions remains.

To apply this method to a dispute over a sentence S that is potentially verbal with respect to term T , one proceeds as follows. First: one bars the use of term T . Second: one tries to find a sentence S' in the newly restricted vocabulary such that the parties disagree nonverbally over S' , and such that the disagreement over S' is part of the dispute over S . Third: If there is such an S' , the dispute over S is not wholly verbal, or at least there is a substantive dispute in the vicinity. If there is no such S' , then the dispute over S is wholly verbal (except in the special case of vocabulary exhaustion, discussed below).

For example: to adjudicate whether a dispute over whether Lee Harvey Oswald acted alone is verbal with respect to ‘alone’, one may bar the use of the term. In this case, it is likely that there

will be various sentences S' for which nonverbal dispute over S' is part of the original dispute: for example, did others help Oswald in his plan to kill Kennedy? If so, the original dispute is not wholly verbal. To adjudicate whether a dispute over 'Pluto is a planet' is verbal, one may bar the use of 'planet'. Here it may be hard to find any nonverbal dispute not involving the term 'planet' that is part of the original dispute (although see below). If there is no such dispute, then the dispute is verbal.⁹

What is it for one dispute to be part of another? It is certainly not required that the dispute over S' be an exact analysis or paraphrase of the dispute over S . We have seen that such analyses may not be available. The dispute over S' need only capture some aspect of the dispute over S . (If one wants to get at further aspects, one can always repeat the method once the status of S' is clarified.) If we are not helping ourselves to definitions or an analytic-synthetic distinction, it is best to understand the relation broadly.

A natural suggestion is to say that a dispute over S' is part of a dispute over S when the parties disagree over S partly *in virtue of* disagreeing over S' . In the case above, for example, it is arguable that parties disagree over whether Oswald acted alone partly in virtue of disagreeing over whether others helped him in his plan to kill Kennedy, whereas it is not clear that parties disagree over Pluto in virtue of any separate substantive disagreement. Here one can adopt the earlier explanatory reading of "in virtue of", requiring that a disagreement over helping explains the disagreement over acting alone, in a stronger than causal way. In this case we might also say that that a disagreement over S' *underlies* the disagreement over S .

If we are to take the talk of "in virtue of" and "underlying" seriously, then this definition requires assuming an asymmetrical structure on the space of disputes, according to which some disputes are more fundamental than others. I think that we have an intuitive sense of such an asymmetrical structure, one that is closely tied to our sense of the relative primitiveness of concepts. But this matter is somewhat controversial, and I do not need to presuppose it at this stage. So for now, it is useful to explore a somewhat broader understanding of "part of". I will return to the narrower understanding in terms of asymmetrical structure in section 7.

To give a broader and easier-to-use characterization of "part of", one can give the notion a counterfactual gloss. We can say that a dispute over S' is part of a dispute over S when: (i) if the parties were to agree that S' is true, they would (if reasonable) agree that S is true, and (ii) if they were to agree that S' is false, they would (if reasonable) agree that S is false. This characterization

⁹Readers may wish to proceed to the next section for further illustrations drawn from philosophy, before returning to this section for analysis of the method.

is subject to some of the subtleties about counterfactual glosses mentioned earlier, but can still play a useful role.

A consequence of the broad definition is that we will sometimes be able to find substantive disputes that are part of verbal disputes. In the Pluto case, some such disputes might include disputes over ‘Astronomical terms should be used in the way that is most useful for science’, ‘Traditions should be respected’, or ‘*X* is president of the astronomical society’ (where the parties agree that the president is the arbiter of meaning, and agree about what *X* says). In these cases, resolving the residual dispute resolves the original dispute over Pluto precisely in virtue of resolving a metalinguistic dispute over the key term ‘planet’. We might say that in these cases, the relation between *S* and *S'* is metalinguistically mediated: the parties disagree over *S* in virtue of disagreeing over a metalinguistic sentence *M*, and disagree over *M* in virtue of disagreeing over *S'*. We can then stipulate that for a dispute over *S'* to be part of the dispute over *S*, in the sense relevant for our purposes, the relation must not be metalinguistically mediated.¹⁰

Even with this stipulation in place, the counterfactual version of the heuristic will sometimes find a substantive dispute over *S'* associated with a verbal dispute over *S*. This will happen when agreeing over *S'* would change one party’s views sufficiently to change their mind about *S*, and likewise for agreeing over $\neg S'$, leading either way to apparent agreement over *S* despite the difference in usage.¹¹ Many such cases will be partially verbal disputes over *S*, but some are wholly verbal disputes as defined earlier. In practice, this does not matter much. Even in the cases in question, the method serves usefully to refocus parties on a substantive dispute in the vicinity. And these cases do not affect the key thesis for our purposes below: the thesis that when there is no such *S'*, then (except in certain special circumstances) the dispute over *S* is wholly verbal.

This heuristic is not intended to provide a reductive definition of a verbal dispute or a wholly mechanical procedure for determining when a dispute is verbal. These ambitions are out of the question, both because of exceptions to the heuristic and because the method itself makes appeal

¹⁰This stipulation will have to be qualified to handle some substantive disputes in linguistics in which the same term is used and mentioned: for example, debates about the meaning of the word ‘meaning’, or the length of the word ‘length’.

¹¹To illustrate using propositions: Suppose that *S* expresses *p* for Jo and a related proposition *q* for Mo, while *S'* expresses *r* for both. Jo and Mo agree on the truth of *p* and $\neg q$, so they have a verbal dispute over *S*. Jo accepts *r*, but if he rejected *r* would he reject *p* (as well as *q*). Mo rejects *r*, but if he accepted *r* he would accept *q* (as well as *p*). Then the two of them disagree substantively over *S'*, but if they agreed on the truth-value of *S'* (either way), they would agree over *S*. This requires disagreeing not just about *r* but also about its evidential relevance to *p* and to *q*, so these cases will not be especially common.

to the notion of a verbal dispute. Rather, the method is a heuristic device that allows us to use clear cases of substantive or verbal disputes to help adjudicate the status of unclear cases.

The method of elimination is extremely useful in practice, both inside and outside philosophy. Indeed, it is often used already, although not as often as it might be. When faced with a dispute that is potentially verbal with respect to a term *T*, one can simply ask the parties: can you state what you are disagreeing about without using (or mentioning) *T*? Or can you at least state some part of the debate this way? If the parties can do this, and if the resulting debate is nonverbal, this is strong evidence that the original dispute had substantive elements. If they cannot, then it is evidence that the original debate was wholly verbal (with the exception discussed in section 7).

The method of elimination can have various outcomes. If it reveals a clearly non-verbal dispute over some relevant *S'*, then the original dispute is not wholly verbal, but in any case the parties might focus on their dispute over *S'* in order to yield a clarified debate. If there is a dispute over some relevant *S'* but it is not clear whether this dispute is verbal, one can iterate the procedure, barring not just *T* but relevant further expressions *T'* used in *S'*. One can make this method more efficient by stipulating that when one bars the original term *T*, one also bars the use of all terms that are undisputed cognates or synonyms of *T* (though it is worth noting that sometimes, a verbal dispute is mirrored in a dispute over which terms are cognates or synonyms). If the method reveals agreement over all relevant *S'* in the restricted vocabulary, this provides reason to think that the dispute over *S* was verbal (or else that it falls into the special case below). If so, then the parties can either set aside the dispute, if they are interested in the first-order domain, or they can focus on the metalinguistic issue, if interested in the words in their own right.

The exception that I have alluded to involves cases of vocabulary exhaustion. If a language has a limited vocabulary, then it might be that once one bars a key expression, one can no longer even formulate any issue that might potentially resolve the original issue. For example, in a language that has just one predicate ('tall', say), barring this predicate will leave one unable to state any dispute relevant to 'X is tall', but that dispute might be substantive all the same. Likewise, if a language contains just one moral term, such as 'good', then in some cases, barring this term may leave one unable to state disputes relevant to 'X is good'. But this phenomenon does not imply that moral disputes are not substantive. So we should acknowledge the possibility of exceptions of this sort.

Later, I will argue that such exceptions are rare. In a language such as English, these exceptions will arise only for terms expressing especially primitive concepts. For other terms, if one cannot find a correlative substantive dispute that does not use the term, the original dispute is itself verbal.

For now, however, I will simply note that if we cannot formulate any such correlative dispute, then one must determine independently whether this is a case of vocabulary exhaustion or a verbal dispute.

4 The Method of Elimination within Philosophy

The method of elimination can be applied to many disputes in philosophy. To illustrate a possible use, I will start with an issue that has often been accused of giving rise to verbal disputes, and in which proponents are relatively sophisticated about these issues: the question of free will and determinism.

Suppose that a compatibilist says ‘Free will is compatible with determinism’, and an incompatibilist says ‘No, free will is not compatible with determinism’. A challenger may suggest that the dispute is verbal, and that the dispute arises only because the parties mean different things by ‘free will’.

We can then apply the method of elimination: bar the term ‘free will’, and see whether there are residual disputes. There are various possible outcomes, depending on the compatibilist and the incompatibilist in question. One possible outcome is that the parties will disagree over a sentence such as ‘Moral responsibility is incompatible with determinism’ as part of the original dispute. If so, this is a *prima facie* indication that the dispute is non-verbal— though one may want to reapply the method to ‘moral responsibility’ to be sure. Another possible outcome is that there will be no such residual disagreement. For example, the parties might agree on ‘Determinism is compatible with degree D of moral responsibility’, ‘Determinism is not compatible with a higher degree D’ of moral responsibility’ (for example, a degree involving desert that warrants retributive punishment), and other relevant sentences. This outcome is a *prima facie* indication that the dispute is verbal, resting on a disagreement about whether the meaning of ‘free will’ requires more than degree D of moral responsibility.

Of course, different diagnoses may apply to different compatibilist/incompatibilist pairs.¹² And further diagnoses may be possible: for example, perhaps the dispute might correlate with a disagreement about whether the possibility of doing otherwise is compatible with determinism, rather than a dispute over moral responsibility. In this case, one might reapply the method in an attempt to see whether this dispute is substantive or verbal. As always, there is no question of the issue of compatibilism being a verbal issue *simpliciter*: rather, disputes over compatibilism may

be verbal for some pairs of parties, and not for others.

In the Socratic tradition, the paradigmatic philosophical questions take the form “What is *X*?”. These questions are the focus of many philosophical debates today: What is free will? What is knowledge? What is justification? What is justice? What is law? What is confirmation? What is causation? What is color? What is a concept? What is meaning? What is action? What is life? What is logic? What is self-deception? What is group selection? What is science? What is art? What is consciousness? And indeed: What is a verbal dispute?

Despite their traditional centrality, disputes over questions like this are particularly liable to involve verbal disputes.¹³ So these disputes are particularly good candidates for the method of elimination. For disputes of this form, we can apply a special case of the method, which we can call the *subscript gambit*.

Suppose that two parties are arguing over the answer to “What is *X*?”. One says ‘*X* is such-and-such’, while the other says ‘*X* is so-and-so’. To apply the subscript gambit, we bar the term *X*, and introduce two new terms *X*₁ and *X*₂ that are stipulated to be equivalent to the two right-hand-sides. We can then ask: do the parties have non-verbal disagreements involving *X*₁ and *X*₂, of a sort such that resolving these disagreements will at least partly resolve the original dispute? If yes, then the original dispute is non-verbal, and the residual disagreement may serve as the focus of a clarified dispute. If no, then this suggests that the original dispute was verbal (unless we have reached the point of vocabulary exhaustion, as discussed shortly). Either way, the method helps to clarify the dispute.

For example, in the dispute over free will, one party might say ‘Freedom is the ability to do what one wants’, while the other says ‘Freedom is the ability to ultimately originate one’s choices’. We can then introduce ‘freedom₁’ and ‘freedom₂’ for the two right-hand-sides here, and

¹²The first diagnosis (involving substantive disagreements about moral responsibility) might apply to P.F. Strawson (1962) and Galen Strawson (1986), for example, while the second diagnosis (involving substantive agreement about moral responsibility underlying verbal disagreement about free will) might apply to J.J.C. Smart (1961) and Derk Pereboom (2001). For these philosophers, in effect, disputes about free will are anchored in the bedrock concept (in the sense discussed later) of moral responsibility. For other philosophers, such as John Fischer (1995) and Peter van Inwagen (1983), an underlying dispute over the possibility of doing otherwise is more central: for these philosophers, disputes about free will are taken to be at least partly anchored in a bedrock concept of what an agent could have done. Still other philosophers (Kantians, perhaps?) might take ‘free’ itself to express a bedrock concept.

¹³I think that the philosophical literature over almost all of the questions in the last paragraph is beset by verbal disputes, in a fashion that is occasionally but too rarely recognized. Of course each debate involves an admixture of substantive elements too.

ask: do the parties differ over freedom₁ and freedom₂? Perhaps they will disagree over ‘Freedom₂ is required for moral responsibility’, or over ‘Freedom₁ is what we truly value’. If so, this clarifies the debate. On the other hand, perhaps they will agree that freedom₁ conveys a certain watered-down moral responsibility, that freedom₂ would be really valuable but that freedom₁ is somewhat valuable, and so on. If so, this is a sign that the apparent disagreement over the nature of free will is merely verbal.

Something similar applies to the debate over the semantics/pragmatics distinction in the philosophy of language. Suppose that one party says that semantic (as opposed to pragmatic) properties are those associated with an expression type by linguistic convention, while another says that semantic properties are those that contribute to the truth-conditions of an utterance. Here the former but not the latter will classify conventional implicature (the difference between ‘and’ and ‘but’, say) as semantic, while the latter but not the former will classify context-dependent referents (the referent of an utterance of ‘I’, say) as semantic.

We can then bar ‘semantic’ (and ‘pragmatic’) and introduce ‘semantic₁’ and ‘semantic₂’ for the two right-hand-sides. Is there a residual disagreement? Perhaps one may find a disagreement over the claim that semantic₁ phenomena rather than semantic₂ phenomena can play such-and-such a role in a linguistic theory, or a theory of understanding. If so, this is a sign that the dispute is substantive. But perhaps one will find no relevant residual disagreement between the parties. In this case, the dispute was verbal all along.

The same goes for some disputes over the nature of justification. Suppose that an internalist foundationalist holds that a belief is justified iff it is rationally grounded in evidence available to the subject, while an externalist reliabilist holds that a belief is justified iff it is produced by a truth-conducive method. If we apply the subscript gambit, is there a residual disagreement over justification₁ and justification₂? For some pairs of internalists and externalists, this is far from obvious. It may be that the parties can agree that justified₁ beliefs go with having reasons and certain subjective norms (perhaps corresponding to a subjective ‘ought’), while justified₂ beliefs go with getting things right and with certain objective norms (perhaps corresponding to an objective ‘ought’). For two parties like this, the original dispute was very likely verbal.¹⁴ For other pairs of internalists and externalists, there may be a residual disagreement, but the method of elimination will at least have clarified the issue between them.

One frequently finds verbal elements in disputes over the formulation of physicalism. For

¹⁴Something like this line is taken by William Alston in his book *Beyond ‘Justification’*, which can be seen as an in-depth application of the method of elimination to disputes in epistemology.

example, some physicalists hold that physicalism is the thesis that everything supervenes on the properties invoked by a completed physics (whether or not they are mental), while others hold that physicalism is the thesis that everything supervenes on the properties invoked by a completed physics and that these properties are non-mental.¹⁵ To adjudicate whether this dispute is verbal, one can bar ‘physicalism’, introduce ‘physicalism₁’ and ‘physicalism₂’, and see whether there is a residual disagreement.

Here, residual substantive disagreements are not easy to find. The most likely candidates are sociological or normative claims: for example, ‘Physicalism₁ is what people in a certain debate are concerned with’, ‘Physicalism₁ is the more important issue’, or ‘Physicalism₁ is what matters for purpose *X*’. Even in these cases, however, it is not clear that any of these sentences will be such that the subjects disagree about it, and such that resolving that disagreement would resolve the original disagreement. But in any case, if these are the residual disagreements, then one can focus on these issues, putting the debate in the sociological or normative realm where it belongs.

The list of cases can be multiplied. I think that almost all the cases on the “What is *X*?” list above can be usefully subjected to the subscript gambit, or to other versions of the method of elimination. For other philosophical disputes that do not take the “What is *X*?” form, the subscript gambit may not directly apply, as there may not be any obvious *X*₁ and *X*₂ to substitute for a key term *X*. But in these cases, we can apply the method of elimination all the same, stepping back to a neutral language and looking for a residual disagreement. This method almost always yields clarification of the original dispute (with a few exceptions discussed below), and sometimes yields significant philosophical progress.

The method of elimination can be useful even when a debate is not verbal. If two philosophers have conceptual mastery of exactly the same concept of physicalism, but one asserts ‘Physicalism is true’ and the other rejects it, then asking them to state relevant disagreements without using the term ‘physicalism’ is likely to clarify what is at issue. Likewise, the method of elimination can usefully be applied even to philosophical assertions made by a single party outside the context of a dispute. If a compatibilist is asked to state her thesis, or relevant aspects of her thesis, without using the term ‘free will’, this may well clarify that thesis for an audience and may help boil the thesis down to the key underlying issues.

¹⁵A version of the first line is taken by Janice Dowell (2006), while a version of the second line is taken by Jessica Wilson (2006). An ancestor of this article, “65,536 Definitions of Physicalism” (delivered at the Bowling Green conference on Formulating Physicalism in 2005) applied the method of elimination to the debate over physicalism in some detail, with the dispute between Dowell and Wilson at that conference serving as one focus.

5 A Possible Reaction

For each of these philosophical disputes, a diagnosis of verbal dispute might be met with the following reaction. “Even though we agree on all the properties of X_1 and X_2 in non- X -involving language, we still disagree. We disagree about whether X is X_1 or X_2 !”. For example, the parties might insist that they have a substantive disagreement about whether physicalism₁ or physicalism₂ is really physicalism, or about whether semantics₁ or semantics₂ is really semantics.

In these cases, however, it is difficult to see what this alleged residual dispute comes to and what could rest on the dispute. Of course the parties might disagree on whether physicalism₁ or physicalism₂ best fits the use of ‘physicalism’ in a certain community, or over whether semantics₁ or semantics₂ best fits the use of ‘semantics’ in a given community. To resolve these issues of usage, one can do sociology, anthropology, linguistics, or experimental philosophy. Once one has agreed on the first-order properties of physicalism₁ and physicalism₂, it is hard to see that anything else in the first-order domain will rest on these questions of usage.

Alternatively, the parties might disagree over whether freedom₁ or freedom₂ best fits our ordinary *concept* of freedom—the one that we use in thought on relevant occasions. To resolve this issue, one might engage in psychology or psychoanalysis, as well as experimental philosophy and the other fields mentioned above. Again, once we have agreed on the first-order properties (including the value) of freedom₁ and freedom₂, it is hard to see that anything else in the first-order domain will rest on this conceptual issue.

Once we have agreed on the relevant first-order properties of physicalism₁ and physicalism₂, and we have agreed on the sociological, linguistic, and psychological questions above, and we reflect adequately, there is no important dispute left to resolve. And even before agreeing on the sociological questions and the like, there is no important *philosophical* dispute left to resolve, at least if we are not directly concerned with language. It may be that as in the deferential squirrel case, there is a fact of the matter about the usage of ‘physicalism’ in the community. If so, and if the parties both defer to their community, then before resolving questions of usage there will be an unresolved fact of the matter about whether the party who says ‘Physicalism is physicalism₁’ or the party who says ‘Physicalism is physicalism₂’ is correct. If we are concerned with language, for example to respect tradition or to facilitate communication, or if we are linguists or philosophers of language, then it may be important to get the answer to this question right. But if we are not, then there is no importance in this residual question.¹⁶

¹⁶At least, there is no importance to the question qua residual question once the relevant first-order questions are

If we do not defer to community usage, then there will be even less reason to resolve questions of usage, and it is arguable that there are no unresolved facts at all. That is, it is arguable that if one uses ‘physicalism’ nondeferentially, and if one knows all the relevant truths about physicalism₁ and physicalism₂ and other truths in non-‘physicalism’-involving vocabulary, and one reflects adequately, then there will be no hidden facts about the truth of ‘Physicalism is physicalism₁’. But even if there are such hidden facts, as on some radically externalist theories of meaning (recall the epistemic theory of vagueness, on which there may be hidden facts about the meaning of ‘tall’), nothing of importance will rest on knowledge of these facts, just as nothing of importance rests on knowledge of such facts in the case of social externalism above. Just as in that case, a dispute over ‘Physicalism is physicalism₁’ among parties who agree on the non-‘physicalism’-involving truths will be a broadly verbal dispute.

In both the deferential and nondeferential cases, a semantic externalist might insist that independent of matters of usage, the question is important: after all, the question of whether physicalism is physicalism₁ or physicalism₂ is *prima facie* a first-order question, and like other first-order questions, value attaches to getting it right.¹⁷ The same goes for getting right the question of whether something is going round something else, even once we know that it is going round₁ and is not going round₂, and so on. But the significance of the residual question (once the other questions are settled) seems to reduce to the significance of the linguistic question of what ‘physicalism’ and ‘going round’ refer to (or should refer to), and has no interest of its own. In knowing that the man is going round₁ the squirrel, that he is not going round₂ the squirrel, and so on, we know the important first-order matters of fact about them. At worst, we are missing knowledge of one of those facts under a certain linguistic mode of presentation. If we are not concerned with language, then we should not be concerned with the missing sort of knowledge.

Of course there are possible cases in which true beliefs here have derivative value. For example, someone might announce a prize that pays only if physicalism is proved (or only if the man goes round the squirrel). But in this case, given that we agree on the relevant prior first-order issues, we can resolve the residual question simply by resolving matters of usage. So in effect, the residual dispute depends on linguistic matters. This behavior contrasts with that of a substantive dispute that is normally pointless. For example, a substantive dispute over the exact location of a specific fundamental particle would normally be pointless, but one can imagine a case in which settled. If the first-order questions are not yet settled, the linguistic facts may have an evidential bearing on them, as discussed in the next section.

¹⁷Thanks to Jonathan Schaffer and Tim Williamson for pushing this line of resistance.

the trajectory of an asteroid heading for the earth depends on this location. In this case, we cannot resolve the dispute over the particle's location simply by resolving matters of usage. This connection to linguistic mediation helps to bring out the way in which the (normal) pointlessness of broadly verbal disputes is distinctive.

Someone might say that true beliefs are intrinsically valuable, and that a true first-order belief about whether Pluto is a planet (say) will carry intrinsic value. But I have suggested above that in the relevant cases, the value of this belief will reduce to the value of true beliefs about whether Pluto is a planet₁, a planet₂, and so on (specified under a non-deferential mode of presentation without hidden facts about reference) along with the value of true beliefs about the meaning of 'planet'. With this much fixed, there is no significant value in the further true belief. This contrasts greatly with the significant value of true beliefs about planethood when they are not held against such a background of existing underlying beliefs.

Perhaps there could be an intellectual value system that gave true beliefs about roundness or physicalism or planethood significant nonderivative value, over and above the value of the other true beliefs in question. But such a value system would seem fetishistic.¹⁸ Once we have the other true beliefs, the extra true belief does not correspond to any significant increase in our understanding of the world. And in practice, sensible philosophers and scientists who are not concerned with metalinguistic matters are almost always willing to set aside these residual questions as insignificant. As always, these residual questions carry with them a distinctive sort of pointlessness. If we are interested in understanding, it is best to simply move on.

6 Bedrock Disputes

The disputes we have discussed are relatively easy cases. Some cases are much harder.

One hard case involves moral disputes. Say that a Kantian and a consequentialist disagree about when an act is right. The Kantian holds that an act is right iff it has a certain deontological property, while the consequentialist holds that an act is right iff it has a certain consequence property. We can apply the usual method to determine whether this is a verbal dispute, barring 'right'

¹⁸An opponent might suggest that our own intellectual value system here is just a function of our psychology, and is not objectively preferable to the "fetishistic" system. I cannot hope to settle questions about the objectivity of epistemic values here, but I am strongly inclined to reject the relativistic view. If one did accept such relativism, then insofar as the notion of a broadly verbal dispute is grounded in the notion of pointlessness, one might also have to be relativistic about whether a dispute is broadly verbal.

and introducing ‘right₁’ and ‘right₂’. Is there a residual disagreement? We might find a dispute over ‘People always ought to perform right₁ acts’ and the like. But ‘ought’ (in the relevant sense) is more or less cognate with ‘right’, so this is not much progress. In any case, we can bar ‘ought’ and introduce ‘ought₁’ and ‘ought₂’, which are stipulated to apply to acts with the deontological and the consequence property respectively. Is there a residual disagreement now?

As we proceed, the disagreement gets harder and harder to state. It is plausible that once all moral terms are gone, no disagreement can be stated. We might agree on all the nonmoral properties of the relevant actions, but still disagree on whether it is right.

In the case of ‘semantics’, ‘physicalism’, and so on, this situation suggested a verbal dispute. Should we likewise diagnose a verbal dispute here? Intuitively, the answer is no. For all we have said, moral disputes are substantive disputes. Instead, we have simply exhausted the relevant vocabulary. It appears that at a certain point (perhaps once we have fixed on the appropriate moral “ought”), we have reached bedrock: a substantive dispute involving a concept so basic that there is no hope of clarifying the dispute in more basic terms.

Another hard case involves disputes over consciousness. Suppose that we disagree over ‘Mice are conscious’. If we bar the term ‘conscious’ are there residual disagreements? We might disagree over ‘Mice are phenomenally conscious’, but this is at best clarification. If we bar ‘phenomenally conscious’, it appears that we are left with cognate disagreements over sentences such as: ‘Mice have experiences’ and ‘There is something it is like to be a mouse’. Again, this is not huge progress.

Once enough phenomenal terms are barred, it may be that no disagreement can be stated. We might agree on all the nonphenomenal properties of a mouse, but disagree on whether it is phenomenally conscious. Again, however, it would be hasty to conclude that the original dispute is verbal. Instead, we have simply exhausted the relevant vocabulary. Intuitively, once we reach a certain point (‘phenomenally conscious’, say), we have reached bedrock.

A third hard case involves disputes over existence. Suppose that a mereological nihilist (one who denies that there are any composite objects) and a non-nihilist disagree over ‘Only particles exist’. We might bar ‘exist’, and introduce terms such as ‘exists₁’ as a quantifier that ranges only over simple objects and ‘exists₂’ so that ‘There exist₂ an *F*’ comes to ‘There exist₁ simples arranged *F*-wise’. The parties might then agree that only particles exist₁, and that not only particles exist₂. Is there a residual disagreement?

At least according to the parties, there is a residual disagreement: it is over whether only particles *exist*, and over whether existence₁ coincides with *existence*. Once certain basic quantifiers

are barred, this dispute may become impossible to state. But according to the parties, we should not conclude that the dispute is verbal. Again, this is a case of vocabulary exhaustion. Arguably, once we have reached a basic quantifier such as the absolute, unrestricted existential quantifiers, we have reached bedrock.¹⁹

These results are compatible with the method of elimination. Insofar as the debates are substantive, then the residual disputes in question are all substantive, until one runs out of vocabulary. Still, these cases bring out an important point of methodology. At a certain point, when the method of elimination is applied, a proponent may say “That’s bedrock”. Here, the thought is that we have reached a point where only cognate expressions can be used to state the issue and where there is no hope of finding a relevant disagreement at an “underlying” level. Further applications of the method will just lead to wheel-spinning, and ultimately to vocabulary exhaustion.

Let us say that disputes of this sort are *bedrock disputes*. To a first approximation, a bedrock dispute is a substantive dispute for which no underlying dispute can be found by the method of elimination. Strictly speaking, one should say that a dispute is bedrock relative to an expression: so the dispute over ‘Mice are conscious’ might be bedrock with respect to ‘conscious’ but not with respect to ‘mice’. A substantive dispute is bedrock relative to an expression *E* when no underlying dispute can be found by applying the method of elimination to *E*: roughly, when there is no underlying dispute that does not involve *E* or cognates.

The “bedrock” card can be played only rarely. Most disputes are not plausibly bedrock disputes. For example, if two people agree on all the underlying non-‘car’-involving statements about an entity (that is, not involving ‘car’ or cognates), but disagree about whether it is a car, it seems clear that the disagreement will be broadly verbal. Correspondingly, substantive ‘car’-disputes will always correlate with disputes about various underlying non-cognate truths, and can thereby be clarified by the method of elimination. So ‘car’-disputes are not bedrock disputes.

Of course a proponent could try claiming that disputes over ‘physicalism’ or ‘semantic’ are bedrock disputes. After all, the move I have recommended concerning ‘phenomenally conscious’ and ‘ought’ is at least structurally analogous to the move discussed in the last section regarding ‘physicalism’. So a proponent could try suggesting that the dispute over physicalism is bedrock, and that no further progress can be made by barring the term. But for the reasons given in section

¹⁹Hirsch (2005; 2009) diagnoses many ontological disputes as verbal disputes, in effect denying that there is a single bedrock absolute quantifier. Sider (2009) replies by arguing that these disputes turn on an especially fundamental quantifier. I discuss the application of the current method to ontological disputes briefly at the end of section 8 of the current article, and in more depth in sections 3 and 4 of Chalmers 2009.

5, this move is implausible. It seems that if two parties agree on all the underlying truths in non-‘physicalism’-involving language, then any further dispute over ‘physicalism’ will be broadly verbal. This indicates that disputes over physicalism are not bedrock disputes.

This suggests that there is a crucial difference between expressions such as ‘ought’, ‘conscious’, and ‘exist’, on one hand, and ‘physicalism’ and ‘semantic’ on the other. We might put this putative difference by saying that the expressions in the first class but not the second are *bedrock expressions*: expressions such that some disputes are bedrock with respect to those expressions. Of course some ambiguous expressions (such as ‘conscious’) may generate bedrock disputes with some uses but not with others. So we can also speak of bedrock *uses* of expressions: uses of an expression in a dispute that is bedrock with respect to that expression.

It remains to clarify the notion of a bedrock dispute. I have said that a substantive dispute is bedrock with respect to *E* when no underlying dispute can be found by applying the method of elimination to *E*. In a little more detail, we might say that a dispute is bedrock with respect to *E* when the dispute is substantive and there is no dispute not involving *E* (and analogous in other respects) that underlies the original dispute.

What is it for one dispute to underlie another? The relation of underlying may need to be taken as something of a primitive here, but there are a few things we can say to clarify it. As discussed in analyzing “part of” earlier, one might say that one dispute underlies another when the second dispute arises in virtue of the first. But the “in virtue of” here must be understood tightly rather than broadly. For present purposes it is crucial that the relation be asymmetrical: if one dispute underlies another, the second does not underlie the first. (For the methodological purposes of the method of elimination, asymmetry is useful but less crucial.) This rules out counterfactual glosses such as “if the first dispute were resolved, the second dispute would be resolved”, because these often apply symmetrically.

As before, one can invoke explanatory notions here: one dispute underlies another when the first dispute at least partly explains the second, by a certain sort of stronger-than-causal explanation. For example, some disputes over whether free will is compatible with determinism might be explained in terms of disputes over whether moral responsibility is compatible with determinism, and so are not bedrock with respect to ‘free will’. But for some disputes over whether mice are phenomenally conscious, there might be no dispute not involving ‘phenomenally conscious’ that explains the original dispute: at best there will be disputes over cognate expressions (‘Mice have experience’) or less basic expressions (‘Mice can feel pain’) that cannot be said to explain the original dispute.

The relevant explanatory relation needs to be something of an idealized dispositional relation. We might say that a dispute over ‘Mice have visual experience’ is explained by a dispute over ‘Mice are conscious’ even though the two parties never utter ‘Mice are conscious’ or even entertain a corresponding thought. Here (and in general) the second dispute might be understood as a potential dispute, or better, as a tacit disagreement. Even in the absence of an explicit dispute, it is reasonable to hold that some disputes over ‘*X* can see’ are explained by disagreement over the functioning of *X*’s eyes, while in other cases, disagreement over functioning is irrelevant and the dispute is explained by disagreement over whether *X* has visual experience.

The explanatory relation is not a mere causal or evidential relation. I might hold that mice are phenomenally conscious because they behave a certain way, and you might deny the former because you deny the latter, but (at least if we are phenomenal realists), the dispute over their behavior does not underlie or explain the dispute over consciousness in the relevant sense. Correspondingly, bedrock disputes should not be identified with disputes over epistemologically bedrock theses: theses that one cannot argue for and that constitute brute intuitions or premises. Rather, what is relevant is a sort of conceptual bedrock.

Instead, the relevant explanatory relation can be seen as a sort of conceptual relation: an underlying dispute must help explain what is at issue in the original dispute. For example, what is at issue in certain substantive disputes over free will can be explained in terms of disagreements about moral responsibility. What is at issue in certain substantive disputes over justification can be explained in terms of disagreements about what people ought to believe. What is at issue in certain substantive disputes about color can be explained in terms of disagreements over what properties we are most directly aware of in color experience. And so on. This is a relationship that philosophers and others are adept at deploying in practice: when asked to explain what is at issue in a given dispute, we can usually give an answer. For present purposes, I will not try to theoretically explicate the notion further (the remarks about inferentialism in the next section suggest one possible theoretical basis for the notion).

Can one dispense with asymmetrical notions such as “underlying” or “in virtue of”? One might suggest that a dispute is bedrock with respect to an expression *E* when it is substantive and unaccompanied by any dispute not over *E*. But this does not work: a dispute over *E* will also be accompanied by disputes over synonyms of *E*, and may be accompanied by entirely unrelated disputes. One might require instead that there be no correlative dispute not over *E* or a synonym of *E*, where ‘correlative’ is stronger than ‘accompanying’ but not as strong as ‘underlying’ (one might gloss it in counterfactual terms as above, for example). But one still faces the problem that

a dispute over ‘Mice are conscious’ may have a correlative dispute over ‘Mice can see’. Likewise, a dispute over ‘X is right’ may have a correlative dispute over ‘X is murder’. Eliminating these disputes requires further structure.

Still, ‘murder’ is a broadly moral term, and can be seen as part of the same family as ‘ought’: the family of broadly moral expressions that express the family of (broadly) moral concepts. Likewise ‘see’ is a broadly phenomenal term, and can be seen as part of the same family as ‘conscious’: the family of broadly phenomenal expressions. So even without asymmetrical structure, one can hold that these families count as bedrock families in virtue of being ineliminable from certain disputes, in that there is no correlative dispute that involves expressions outside the family. And once one has gotten this far, it is natural to hold that certain expressions play an especially fundamental role within these families. For example, it is natural to hold that the debate over whether a mouse is conscious is more fundamental than a correlative dispute over whether a mouse can see (in a phenomenal sense of ‘see’), and that while moving from the latter to the former may clarify a debate about whether a mouse can see, moving from the former to the latter does not really clarify a debate about whether mice are conscious.

In effect, the notion of a bedrock dispute requires a sort of asymmetrical structure among the space of disputes, such that (for example) disputes over ‘conscious’ are more basic than disputes over ‘see’, and disputes over ‘ought’ are more basic than disputes over ‘murder’. This asymmetrical structure does not fall out of the very notion of a verbal dispute. Instead, it is suggested by the character of the relations among disputes that we find on examination.

Of course this picture is not compulsory. An opponent might hold that there is no such asymmetrical structure, and might hold that the notion of bedrock is ill-defined. On this view, one can still employ the method of elimination, as long as it is not constrained by the asymmetrical structure: for example, one might clarify disputes over ‘conscious’ into disputes over ‘see’, disputes over ‘ought’ into disputes over ‘murder’, and so on. This will lead to a view on which there is no privileged stopping point for the method of elimination. Instead, there will be just a linked web of related expressions and sentences, with many different starting and stopping points.

We might think of this sort of opposition as “holistic”. It differs markedly from the earlier opposition that rejected the very notion of a broadly verbal dispute: that sort of opposition tends to downplay internal relations among disputes, and so might be called “atomistic”. By contrast, the holistic opposition accepts the notion of a broadly verbal dispute, and can accept rich internal relations among disputes, while denying the sort of asymmetrical structure that privileges some of

those disputes as especially basic.²⁰

While I argued against atomistic opponents earlier, I do not have conclusive arguments against holistic opponents. But I think that the phenomena we have examined at least tend to suggest a view with more structure than the holistic picture allows. We have seen that even without appealing to asymmetrical structure, one can motivate the idea that some families of expressions are bedrock (in virtue of being ineliminable in relevant disputes), as above. Once one has gotten to bedrock families, it is natural to find further structure within these families. In particular, all the ineliminable families above appear to have certain specific concepts at their heart, which is at least suggestive of a distinctive role for those concepts. Of course an opponent might hold that any apparent structure here is a mere artifact of our psychology or our attitudes. But I will adopt the idea that there is some privileged structure among disputes at least as a promising working hypothesis.

7 Bedrock Concepts

At this point, it is natural to think in terms of primitive concepts. We can suppose that there is a special class of *bedrock concepts*: concepts that generate bedrock disputes.²¹ We can define a bedrock concept as a concept that can be expressed by a bedrock use of an expression. Then bedrock expressions will be expressions that (sometimes or always) express bedrock concepts. Every bedrock dispute will involve a bedrock expression that both parties use to express the same bedrock concept: then we can say that the dispute is bedrock with respect to the concept. In effect, bedrock concepts are concepts so basic that we cannot clarify substantive disputes involving them in more basic terms.

If one adopts the holistic view in the previous section and rejects the notion of a bedrock dispute, one will also reject the notion of a bedrock concept. But one may still be entitled to the

²⁰Of course a holistic “web of belief” picture is also sometimes used by Quineans to cast doubt on conceptual connections and not just on conceptual priority. So one might think that this picture could also be used to cast doubt on or deflate the notion of a verbal dispute, and not just the notion of bedrock disputes. A reviewer suggests that a holist might say that the more similar two webs of belief are, the closer a dispute is to being verbal: perhaps a verbal dispute is one that has little impact on the rest of the web? It is not obvious how this picture will handle the apparent phenomenon of verbal disputes that have significant impact (on a group of linked expressions, for example), or of substantive disputes that have little impact. But if the picture can be elaborated, it may well be compatible with the claims about verbal disputes in the first half of this article.

²¹I construe concepts as abstract objects: entities that can serve as contents of simple expressions and as constituents of the contents of judgments. Much of what I say might still apply on a view that construes concepts as types of mental representation, however.

notion of a bedrock family of concepts, as above. For example, the family of moral concepts or of phenomenal concepts may count as bedrock families, in virtue of being ineliminable from relevant disputes that involve them. If one rejects the holistic view, then there will also be internal structure to these families: the concept *conscious* might be the most basic concept in the phenomenal family, for example, or the concept *ought* might be the most basic concept in the moral family.

What are the bedrock concepts, or the bedrock families? This can be seen as a successor to other questions about primitive concepts, such as: what the concepts in terms of which all other concepts can be defined or analyzed? But unlike that question, it does not require that definitions and analyses are possible. Still, as we seen earlier in this book, questions concerning which concepts (or families) are primitive are among the deepest questions in philosophy. One's answers to these questions may be deeply interwoven with one's other philosophical commitments, and often reflect the deepest issues between rival philosophical commitments. The same goes for questions about which concepts are bedrock. So I cannot hope to resolve these questions here.

For example, phenomenal realists (including dualists and a posteriori materialists) may regard 'consciousness' as bedrock, while deflationists about consciousness (including analytic functionalists and eliminativists) may hold that any substantive debates over 'consciousness' are resolvable into debates about associated functional and dispositional features such as discrimination, verbal report, and higher-order thought. Indeed, Dennett (1991) argues that once we have settled questions about functional features such as these, any further questions about consciousness are empty questions, presumably generating verbal disputes. By contrast, phenomenal realists like myself hold that there are further substantive questions here. Both sides have arguments for their view (for example, phenomenal realists can invoke conceivability and knowledge arguments to make a case for there being a further question), and these arguments can be seen as arguments about which concepts are bedrock. But of course no such argument receives unanimous agreement.

Likewise, some moral theorists may regard 'ought' as bedrock, while others may regard 'right' or 'good' as bedrock. This issue about priority between disputes closely reflects the traditional question of the relative priority between deontological and evaluative notions. Some other theorists may hold that none of these notions is bedrock, and that moral disputes are ultimately to be resolved into disputes about specific moral properties such as virtues and vices. And some moral subjectivists may hold that no moral concepts are bedrock: moral disputes among those who agree on the nonmoral facts are themselves verbal disputes depending on whose preferences are at issue. Other moral anti-realists, such as expressivists and fictionalists, are likely to hold that these disputes are not verbal (perhaps invoking an open question argument), so that some moral concepts

are bedrock.

Similarly, an ontological realist may regard ‘exists’ as bedrock, while an ontological anti-realist may not. A non-Humean may regard something in the vicinity of ‘law’ or ‘cause’ as bedrock, while a Humean may not. Some inflationists about truth may regard ‘true’ as bedrock, where deflationists do not. Many epistemological theorists may take one or more epistemic ‘ought’s to be bedrock, while others may take other epistemic terms as bedrock: for example, some primitivists about the concept of knowledge may regard ‘know’ as bedrock. Some theorists of free will may regard ‘free’ as bedrock, while others will deny this, taking the relevant bedrock concepts to include moral concepts, or perhaps practical concepts concerning what agents are able to do. Phenomenalists may hold that only phenomenal expressions are bedrock. Verificationists may hold that only observational expressions are bedrock. And so on.

It is worth noting that at least some of these issues may arise even on the holistic view. For example, a holistic phenomenal realist will hold that the family of phenomenal expressions is ineliminable from phenomenal disputes, while a holistic phenomenal anti-realist will not. A holistic moral theorist might reject the question of the relative priority among various specific moral notions, while still holding that the family of moral expressions as a whole is ineliminable from disputes involving them.

As in earlier chapters, questions about which concepts are bedrock should be distinguished from questions about which properties are metaphysically fundamental. Some type-B materialists may hold that the concept of consciousness is bedrock even though consciousness is metaphysically derivative, for example. An expressivist may hold that moral concepts are bedrock even though they do not pick out properties at all. In the reverse direction, one might reasonably hold that although spin and charge are metaphysically fundamental, the concepts of spin and charge are not bedrock. Disputes about a particle’s spin can be resolved into disputes about the role that the particle and its properties play, so that once one fixes a Ramsey-sentence characterization of the particle and its properties, disputes about its spin are verbal.

How do we decide which concepts are bedrock? I do not have any foolproof method. Some familiar tools can get us some distance: for example, conceivability arguments and open-question arguments can make the case that certain families of concepts are ineliminable. The process of conceptual analysis, even if it does not yield a precise analysis, can make a case that a concept is derivative. If one finds that a concept frequently appears in analyses of other concepts and resists analysis itself, that is some evidence for bedrock status. Likewise, if one cannot find a disagreement that plausibly underlies a dispute, that is some evidence that the latter dispute is bedrock. All

of these methods are fallible (perhaps one has just not yet found the relevant underlying notion yet), so diagnosis of a concept as bedrock is almost always tentative and controversial.

As with our earlier discussion of primitive concepts, some candidates for bedrock concepts include: phenomenal concepts (consciousness, specific phenomenal qualities?); normative concepts (certain moral and epistemic oughts?); some logical and/or mathematical concepts (negation, existential quantification?); some nomic or modal concepts (nomological necessity?); spatiotemporal concepts (relative location, temporal order?); some indexical concepts (I, now?) and just possibly, the concept of explanation (it is striking that many philosophical disputes resolve themselves into disputes over what explains what). Other philosophers will add other concepts: those of secondary qualities, intentional states, freedom?

Viewed through this lens, the method of elimination can be seen as a method for resolving philosophical disputes into two basic kinds of disputes: first, disputes involving bedrock concepts (or at least, putative bedrock concepts); second, disputes over which concepts are bedrock. If we could agree on the answers to these questions, then answers to all other important philosophical disputes would be within our grasp. Sadly (or happily), this does not provide a silver bullet for solving philosophical problems: both sorts of residual dispute are as hard as any dispute in philosophy. Still, the residual disputes will at least take a certain regimented form, and many disputes will dissolve or clarify along the way. That is a form of philosophical progress.

I have defined bedrock concepts wholly in terms of disputes and their properties. Are bedrock concepts also primitive in some other sense? If one embraces a definitional view of concepts, one could see them as the basic concepts from which all other concepts are composed. But if one rejects this picture, as I do, there may be other ways to explicate their primitiveness.

In the eighteenth excursus, following this chapter, I argue that the present framework is compatible with an inferentialist view of concepts and of content, on which all concepts and all contents are individuated by their relations to bedrock concepts. In effect, the bedrock concepts serve as the primitive elements in which all concepts are anchored. On this picture, we grasp most concepts through inferential connections to other concepts, but we do not grasp bedrock concepts this way. This suggests that we grasp them in a way that is especially direct.

This picture also helps to analyze the strengths and limitations of another method sometimes used to diagnose verbal disputes: the method of charity. Hirsch (2005) suggests that we should interpret expressions in a way that minimizes certain errors on the part of speakers and use this method of charity to detect verbal disputes. On the anchored inferentialist picture, there may be a role for charity for non-bedrock expressions: if expressions (such as ‘go round’) get their content

through inferential connections to other expressions (such as spatial expressions), then at least insofar as speakers' assertions ('something goes round something iff...') reflect these inferential connections, then we should expect that these assertions are a guide to the content. However, there is no analogous role for charity for bedrock expressions: because these expressions do not get their content through their inferential role, it is much easier for the inferences and assertions that they are involved in to be largely mistaken.

For example, if we applied a general method of charity to the term 'conscious', we would naturally reach the conclusion that communities who say 'Only humans are conscious' and that communities who say that 'Flies are conscious' (where both communities are competent, reflective, and well-informed empirically) mean different things by 'conscious'. This might be reasonable if we know that only non-bedrock uses of 'conscious' are involved: it might be evidence that one person is invoking reflective awareness and the other is invoking mere perceptual discrimination, for example. But if there are bedrock uses of 'conscious' (say, for phenomenal consciousness), then it is perfectly possible for two people to use the same bedrock concept in these very different assertions. In such a case, the method of charity yields the wrong result.

Likewise, following Hirsch (2005; 2009), one might apply the method of charity to suggest that communities who say 'Only simples exist' and communities that say 'Mereological sums exist' (where both communities are competent, reflective, and well-informed empirically) mean different things by 'exist'. This may be reasonable if we know that only non-bedrock uses of 'exist' are involved. But if there is a bedrock concept of existence, as many ontological realists hold, then the method of charity will not apply to it: the two communities may be using the same bedrock concept, despite their different judgments. On this analysis, the dispute between ontological deflationists such as Hirsch and their opponents will then come down to the prior question of whether there is a privileged nondefective bedrock concept of existence (see Chalmers 2009).²²

²²I think that a similar diagnosis applies to another method occasionally used to diagnose verbal disputes: the method of translation. For example, it is sometimes held that if one can translate everything that a presentist says into eternalist language and vice versa (translating presentist 'exists' into eternalist 'exists and is present' and translating eternalist 'exists' into presentist 'existed, exists, or will exist') so that both parties accept corresponding sentences, then the dispute is verbal (see Sider 2006). Of course this method works best if the translation preserves content. In some cases involving non-bedrock expressions, such a claim is plausible: the similarity between the inferential roles of the translated expressions (especially with respect to bedrock expressions) suggests similarity of content between these expressions. But in cases involving bedrock expressions, one cannot infer similarity or difference in content from similarity or difference in these inferential roles. So insofar as presentists and eternalists take themselves to be having

All this tends to suggest that the sort of conceptual pluralism discussed earlier (and also in the seventeenth excursus) is less apt for bedrock concepts. Of course terms such as ‘consciousness’, ‘ought’, and ‘exist’ can be used in multiple ways, yielding verbal disputes. But insofar as these are bedrock expressions, there is a privileged concept in their vicinity. The considerations above tend to suggest that bedrock concepts function as “concept magnets” of a sort: many different patterns of thought are consistent with possessing the same bedrock concept. Of course one can still chart other concepts nearby, but there is a sense in which bedrock concepts may serve as the basic axes in conceptual space.

8 Analyticity and Translucency

The notion of a verbal dispute is connected in various ways to the notion of analyticity. I have not explicitly presupposed a notion of analyticity in what has gone before. But if one accepts the notion of a verbal dispute, one might use it to characterize certain notions of analyticity in turn. This project might be seen as analogous to the project of the last section, in which we used the character of verbal disputes to characterize primitive concepts in dialectical terms. Where the previous project required an asymmetrical structure of the space of disputes, however, the current project requires only the notion of a broadly verbal dispute.

Most straightforwardly, one might suggest that a sentence S is *dialectically analytic* iff necessarily, any dispute over S is broadly verbal. Or better: S is dialectically analytic iff S is true and any dispute over S involving at least one competent user of S is verbal.²³

The current framework is neutral on whether there are any dialectically analytic sentences. From the fact that some disputes are verbal and some are not, it does not follow that any sentences or utterances are such that all disputes involving them are verbal. It is quite possible to hold that for every sentence S , some disputes involving a competent user of S are verbal and some are not. So the question of whether there are any dialectically analytic sentences or utterances is itself a substantive question. This brings out a sense in which the current framework is not automatically committed to a strong notion of analyticity. Still, it is useful to explore the relationship between dialectical analyticity and other notions of analyticity.

a non-verbal dispute using the same bedrock notion of existence, the translatability argument is no argument against them.

²³Here I understand competence as full competence: a competent user has a full grasp of the public meaning of a term, as opposed to the sort of incomplete grasp exemplified by the ‘arthritis’ cases in Burge 1979.

Among extant notions of analyticity, notions of epistemological analyticity (Boghossian 1996) are perhaps most closely connected to dialectical analyticity. Let us say that S is epistemologically analytic iff necessarily, every competent user of S who considers S accepts S . Epistemological analyticity plausibly entails dialectical analyticity. If S is epistemologically analytic, any dispute over S must involve one party who is not a competent user of S . Given that the other party is competent, the dispute over S will then be explained by the metalinguistic difference and so will be verbal. It is not obvious whether dialectical analyticity entails epistemological analyticity: this depends on whether two competent speakers can verbally disagree over a dialectically analytic S , perhaps because of semantic externalism.

Even if the two notions are not quite the same, the existence of dialectically analytic sentences arguably stands or falls with the existence of epistemologically analytic sentences. Suppose that (in the spirit of Williamson 2007) one denies that there are any epistemologically analytic sentences, holding that any sentence, even ‘Vixens are female foxes’, can be competently rejected. One might well hold that such cases can always be construed to involve nonverbal disagreement with those who accept the sentence. If so, one will then deny that there are any dialectically analytic sentences.

One could also use the current framework to define a notion of analyticity of an utterance. Here we can say that an utterance of S is dialectically analytic if necessarily, any dispute over S involving that utterance is verbal. It is perhaps best to also require that the subject is a competent user of S , to avoid cases of analyticity grounded in false beliefs about meaning. Dialectically analytic utterances will always be true, given the plausible claim that any untrue utterance of a sentence S can always be involved in a nonverbal dispute over S . If S is a dialectically analytic sentence, all competent utterances of S will be dialectically analytic. But when S is not dialectically analytic, the definition allows that some but not all competent utterances of S may be dialectically analytic. For example, given a context-dependent term such as ‘bald’, some but not all utterances of ‘Someone is bald iff they have no hairs’ might be dialectically analytic.

The thesis that there are dialectically analytic utterances is significantly weaker than the thesis that there are dialectically analytic sentences, and the view is correspondingly easier to defend. For example, if I introduce the term ‘glub’ by saying stipulatively ‘Something is a glub iff it is a purple book’, then it is arguable that anyone who denies this claim will be verbally disagreeing with me, although the matter is not entirely clear (perhaps an opponent might have unusual views about biconditionals). Still, stipulative utterances like this are relatively uncommon. And even terms introduced by stipulations rapidly take on lives of their own, to the extent that it is not obvious that

disputes between later users of the terms must be verbal in the same way. So it is arguable that dialectically analytic utterances are rare.

Importantly, nothing in these doubts about dialectical analyticity or epistemological analyticity does much to cast doubt on the notion of a verbal dispute itself. So although it is sometimes said that the notion of a verbal dispute presupposes an analytic-synthetic distinction, the analysis here suggests that this is not quite right. Even if there are no instances of the sorts of analyticity above, there may still be broadly verbal disputes. This raises the possibility that one can use the notion of a verbal dispute to do at least some of the work that notions of analyticity have been used to do.

For example, one can use the notion of a verbal dispute to introduce a useful notion of *translucency*. Let us say that a sentence *S* is translucent with respect to a set of sentences *T* iff, necessarily, if two competent users of the sentences in *T* agree (nonverbally) on all these sentences, then any dispute over *S* (where at least one uses *S* competently) is verbal. Defining translucency for utterances is harder: a version of the previous definition would apply only to cases in which the speaker already has attitudes to all the sentences in *T*. But in a similar spirit, we might say that a competent utterance *S* is translucent with respect to a set of sentences *T*, iff necessarily, any substantive dispute involving that utterance must be accompanied by a potential substantive dispute over a sentence in *T*. Here two parties have a potential dispute over a sentence in *T* iff, were they to consider that sentence (and be competent users of it), they would dispute it. The idea again is roughly that if two speakers agree substantively over all sentences in *T*, any dispute over *S* must be verbal.

I have already in effect invoked the notion of translucency in earlier discussions. For example, I have argued that ‘physicalism’-involving sentences are translucent with respect to non-‘physicalism’-involving sentences, that ‘car’-involving sentences are translucent with respect to non-‘car’-involving sentences, and so on. Likewise, in the squirrel case, certain utterances *S* of ‘The man goes round the squirrel’ are arguably translucent with respect to an appropriate set *T* of non-‘round’ involving statements about the man, the squirrel, and the tree. If two parties disagree nonverbally over *S*, they must also potentially disagree over some sentences in *T*; if they agree over all the sentences in *T*, then they disagree only verbally over *S*.

There may be cases in which some but not all utterances of a sentence *S* are translucent with respect to a set *T*. Suppose that both of James’s proffered meanings of ‘go round’ are available to competent users, and suppose that one party uses it in the north-east-south-west sense while another uses it the front-left-back-right sense. Suppose also that *T* specifies relative motion of the man and the squirrel in north-east-south-west terms but not in front-left-back-right terms. Then it

may be that an utterance of ‘The man goes round the squirrel’ by the first party is translucent with respect to T while an utterance by the second party is not.

Translucency is closely related to dialectical analyticity. In particular, a true sentence or utterance is dialectically analytic iff it is translucent with respect to the empty set. (The same does not hold for false sentences, which can be translucent on the current definition but cannot be dialectically analytic.) But even if no sentences are translucent with respect to the empty set, there may still be many other cases of translucency. For example, even if there are no dialectically analytic truths involving ‘car’, it remains plausible that ‘car’-sentences are translucent with respect to non-‘car’-sentences.

What is the epistemological status of translucency? Here, it is useful to compare it to an epistemological notion of transparency. By analogy with epistemological analyticity, we can say that S is *transparent* with respect to T if any speaker who uses S and T competently, who supposes that the sentences in T are true, and who considers S will know the truth-value of S under the supposition. Transparency plausibly entails translucency, but the reverse is not the case. If there are cases of translucency in which two competent speakers who agree on T verbally disagree on S , perhaps for reasons of context-dependence or semantic externalism, S will not be transparent with respect to T . So translucency does not build in an epistemological notion of transparency.

Still, the term ‘translucency’ is used to suggest that while the notion is weaker than that of transparency, it has relevantly similar effects. For example, ‘The man goes round the squirrel’ may not be transparent with regard to underlying sentences specifying the man and the squirrel’s movements, relative orientation, and so on, because of semantic externalism. But the sentence is plausibly translucent with respect to some such set of sentences. So if we are concerned only with the first-order domain and not with language, then agreement on the underlying sentences in effect enables us to know everything important that we need to know in asking this question. The residual verbal issue can be set to one side.

Once we have accepted the notion of translucency, it is natural to hold that it can do at least some of the work that an analytic/synthetic distinction might do. When a truth S is translucent with respect to a set of truths T , for example, it is natural to hold that the status of S with respect to T is in some sense a matter of language. When a truth S is not translucent with respect to T , it is natural to hold that the status of S relative to T is in some sense a matter of substance. Here, translucency is akin to a sort of relative analyticity.

Furthermore, one can argue that translucency has at least some epistemological bite. Let us say that S is *positionally transparent* with respect to T if any speaker who uses S and T competently

and without semantic deference is in a position to know the truth-value of S under the supposition that the sentences in T are true. Then, by generalizing the “no-hidden-fact” argument in section 5 it is arguable that translucency entails positional transparency.

The notion of positional transparency is a relative of the notion of *positional analyticity*. We can say that S is positionally analytic iff necessarily, anyone who uses S competently and without semantic deference is in a position to know S . If “in a position to know” involved the extreme idealizations earlier in this book, positional analyticity would collapse into a version of apriority. But instead we should understand it in a more standard way, so that it involves only a modicum of idealization away from what one is disposed to accept.

Positional analyticity and positional transparency are not threatened by Williamson-style cases in which competent users reject apparently analytic claims. Even if sentences such as ‘Something is a vixen iff it is a female fox’ are not epistemologically analytic, they are arguably positionally analytic. Even if philosophically stubborn subjects do not accept this sentence, they still appear to be in a position to know it. Insofar as Williamson’s key cases here all involve cases in which subjects do not know what they are in a position to know (or perhaps cases in which subjects use expressions with semantic deference), they do not provide an argument against positional analyticity. (Williamson himself appears to concede this, suggesting instead that notions invoking “position to know” are not genuine forms of analyticity or conceptual truth.) If this is right, then they also do not provide an argument against positional transparency.

The thesis that translucency entails positional transparency might nevertheless be denied by someone (an epistemic theorist of vagueness, say) who holds that there are sources of semantic externalism other than semantic deference and inadequate reflection, resulting in hidden facts about meaning. Still, there is significant plausibility to the thesis, which in effect gives an epistemological role to translucency.

Those who accept an analytic/synthetic distinction may hold that (at least in the case of context-independent sentences used without semantic deference) where there is translucency, there is analytic entailment, and vice versa. Likewise, those who accept an a priori/a posteriori distinction may hold that where there is translucency, there is a priori entailment (although probably not vice versa). But even if one rejects these notions, one may still hold that translucency is a successor notion to analyticity and apriority, one that can do a significant amount of the work that these notions are supposed to do. It is not uncommon for even Quineans to employ notions akin to translucency in their work, so it is not obvious that a Quinean must reject the notion. Of course, once we accept that translucency can play certain core roles of the analytic/synthetic distinction, a

dispute over whether it vindicates that distinction may itself be verbal.

Translucency relations provide a foundation for the method of elimination. One might see the method as being grounded in a thesis that for any non-bedrock expression E , all E -truths are translucent with respect to non- E -truths. Once this thesis is established, it follows that any nonverbal disagreement involving E must be associated with a nonverbal disagreement in a non- E -vocabulary, whereas verbal disagreements involving E need not be. In a similar vein, one might hold that among the non- E -truths, F -truths (for a non-bedrock expression F) are translucent with respect to non- F -truths. Repeating the process, there will be a chain of translucency relations grounded ultimately in bedrock truths: truths involving only bedrock expressions. This structure is at the core of the method of elimination.

One might say that S is translucently settled by T iff either (i) S is a member of T or (ii) S is translucent with respect to a set T' all of whose members are translucently settled by T . Translucent settling is weaker than translucency proper, at least to the extent that translucency relations are nontransitive. The reasoning above suggests that all truths are translucently settled by bedrock truths.

The thesis that all truths are translucently settled by bedrock truths is strongly reminiscent of the Primitive Analytic Scrutability thesis, holding that all truths are analytically scrutable from truths involving primitive concepts. The analogy suggests once again a picture on which all truths are conceptually anchored in some fashion in bedrock truths. And it suggests once again a picture on which links between bedrock and non-bedrock are relatively superficial (these links will be analogous to analytic truths) and on which the issues involving bedrock concepts are highly substantive (the bedrock truths in the base will be analogous to empirical or synthetic a priori truths). In this way, even without presupposing an analytic-synthetic distinction, one might reconstruct some of the spirit of this traditional picture.

It is also related to the A Priori Scrutability thesis. If we accept the thesis linking translucency and a priori entailment, the translucency thesis (conjoined with the thesis that there is a limited class of bedrock truths) entails the scrutability thesis. As with analytic scrutability, however, translucent settling is a stronger relation than a priori entailment. For example, it may be that normative truths are a priori entailed by non-normative truths, without being translucent with respect to non-normative truths. Something similar may apply to mathematical truths. So it may be that while a scrutability base need not include normative truths or mathematical truths, a translucency base will include these truths.

So translucent settling is more akin to analytic scrutability than to a priori entailment. The class

of expressions in a translucency base is more akin to the expressions in an analytic scrutability base than in an a priori scrutability base. And whereas no truths in a minimal scrutability base will be a priori, some truths in a translucency base may be a priori: for example, certain mathematical truths, and certain normative truths. These truths might be seen as the counterparts in the present framework of synthetic a priori truths. Seen this way, all bedrock truths will be synthetic truths: either empirical or synthetic a priori. By contrast, truths connecting bedrock concepts to higher-level concepts will be akin to analytic truths. If we combine these observations with the earlier observation that the deepest questions in philosophy involve the status of bedrock truths, one can draw a version of the conclusion that the deepest questions in philosophy concern the synthetic and not the analytic.

9 Carnapian Conclusion

The picture developed in this chapter has many connections to Carnap's logical empiricism. I am inclined to think that attention to verbal disputes has the potential to lead to at least a partial vindication of various Carnapian projects.

Most obviously, some elements are reminiscent of the *Aufbau*. We have seen that reflection on verbal disputes gives us some grounds for accepting a distinctive class of primitive concepts, as well as something that can play the role of an analytic/synthetic distinction. All this helps to reveal something of the conceptual structure of our conception of the world. Where the *Aufbau* held that all truths were definitionally entailed by primitive truths, here we have a different thesis that is similar in spirit: all truths are translucently settled by bedrock truths.

One can even see the logical empiricists' verificationism (Carnap 1936) as a special case of the framework. Verificationists in effect hold that a dispute is substantive only if it involves a disagreement about the observational. This leads naturally to a version of the framework on which all bedrock truths are observational truths. On the view as I have developed it here, bedrock truths may include nonobservational truths such as normative truths, causal truths, truths about other minds, and so on, so there is no verificationism per se. But there is an interesting structural parallel.

We also have a Carnapian pragmatism about conceptual frameworks. On this view, instead of focusing on existing words and the concepts they express, we should focus on the role one needs them to play. Carnap's (1934) principle of tolerance stresses that everyone is at liberty to build their own form of language. In Carnap's later work (1950a, 1950b), he rejects traditional conceptual

analysis for a project of explication that may be revisionary and that gives the fruitfulness of a concept a central role, and he holds that the choice between different conceptual frameworks depends on our purposes rather than on matters of fact. My pragmatism does not run as deep as Carnap's: Carnap would almost certainly have rejected the claim that there is a privileged basis of bedrock concepts, holding instead that one can equally start from a phenomenal basis or a physical basis (among others), depending on one's purposes. Still, on his project and on this one, the philosophical interest of most concepts lies in the work that they can do for us.

Finally, Carnap and other logical empiricists hoped to provide methods to clarify and resolve philosophical disputes once and for all. I have not provided anything like that. Still, the method I have outlined has the potential to clarify many or most philosophical disputes, and to resolve some of them. It will certainly not resolve them all, but it often gets us closer to the heart of the dispute, and leaves us with a clearer understanding of the fundamental issues underlying a dispute and of what remains to be resolved. That is a form of philosophical progress.

Seventeenth Excursus: Conceptual Analysis and Ordinary Language Philosophy

This book has yielded a mixed verdict on the prospects for conceptual analysis. On one hand, the framework earlier in the book suggests that conceptual analysis is possible. On the other hand, the arguments of chapter 9 suggest that its philosophical importance is limited.

We can start by stressing the positive. In the twentieth century, many argued that conceptual analysis is impossible: partly by arguing against the definitional model, and in partly arguing against an analytic/synthetic and apriori/aposteriori distinction. Here I have argued that a sort of conceptual analysis can survive these critiques. In particular, I have argued that a Carnapian intensional conception of conceptual analysis is unthreatened by Quine's critique of analyticity and by the failure of the definitional program.

On the intensional model of conceptual analysis, concepts are represented in effect by verdicts about cases. More specifically, they are represented by functions from scenarios to extensions. Nothing in the critique of definitions or Quine's critique compromises our ability to make armchair judgments about scenarios, at least given ideal reflection. With enough work and sufficient rational reflection, we can chart the intension for any concept that we possess and can map out its general structure.

Construed as a search for perfect definitions, contemporary conceptual analysis has been a failure. But construed as an attempt to better understand various specific concepts, it has arguably been a success. We have certainly come to understand the contours of concepts such as 'know', 'cause', and 'lie', for example, much better than we did previously. In part, this is because the intensional model can at least yield better and better approximate definitions (chapter 8), and better and better ways to chart the broad patterns of application that a concept involves. All this in effect yields a way to chart the the structure of a concept.

Following this mode, we might understand conceptual analysis as a quasi-scientific process of conjecture, refutation, and refinement. A purported analysis is put forward, counterexamples are found, and a refined analysis is discovered that avoids the counterexamples. Further counterexamples can be found in turn, and we do not in practice reach a perfect definition, but we make progress all the same. The phenomenon is analogous to that found in the high-level sciences, in which counterexample-free principles and laws are almost never found, but in which one can nevertheless state approximate principles that get most cases approximately right, and in which one can refine these principles via successive improved principles, gradually converging on the

truth. If there are converging approximate definitions for our concepts, then we should expect that something similar may be possible in the process of conceptual analysis.

All this may be useful for various purposes in the philosophy of language and mind. It allows us to theorize about the meaning of expressions and the contents of concepts in a general way, and it also allows us to investigate the meaning of many specific expressions and the content of many specific concepts.

Still, the project of conceptual analysis has often been associated with stronger claims than this. Conceptual analysis of specific expressions has sometimes been seen as the very heart of philosophy, and as the method that drives us to the most important philosophical conclusions. My own view is very different from this. I think that the conceptual analysis of specific expressions can lead us to linguistic conclusions, and can thereby settle linguistic disputes and verbal disputes. But its power to yield substantive philosophical conclusions, those which settle substantive debates, is highly limited.

To see this, we can start by noting that much of traditional conceptual analysis centers on questions of the form “What is X ?” or “What is it to be X ?”. Much of contemporary naturalistic philosophy centers on these questions too. In chapter 9, I have argued for a deflationary view of the interest of these questions. Some component of these questions is inevitably verbal, and the non-verbal residue can be found without using ‘ X ’.

On the picture I favor, instead of asking “What is X ”, one should focus on the roles one wants X to play, and see what can play that role. The roles in question here may in principle be properties of all sorts: so one focuses on the properties one wants X to have, and figures out what has those properties. But very frequently, they will be causal roles, normative roles, and especially explanatory roles.

For example, instead of asking “What is semantic content?” and expecting a determinate answer, one can instead focus on various explanatory roles one wants semantic content to play. One can then say, here are some interesting properties (of sentences or utterances): S_1 can play this role, S_2 can play these roles, S_3 can play these roles. Not much hangs on the residual verbal question of which is really semantic content.

Likewise, instead of asking “What is a belief? What is it to believe?” and expecting a determinate answer, one can instead focus on the various roles one wants belief to play. One can then say, here are some interesting states: B_1 can play these roles, B_2 can play these roles, B_3 can play these roles. Not much hangs on the residual verbal question of which is really belief.

The picture is also deflationary about the claims of some opponents of conceptual analysis.

Such opponents often say: “I don’t care about the concept of *X*. I care about what *X* really is. Even though *X*₁ doesn’t mirror our concept of *X*, *X* really is *X*₁.” For example, when Ruth Millikan (1983, 73) gives her teleological theory of meaning, she says that she is not analyzing the concept of meaning, and that she is instead giving a theory about the natural nature of meaning. Likewise, when Hilary Kornblith (2002, 1-2) gives his naturalistic account of knowledge, he says that he is not analyzing the concept of knowledge, and is instead giving an account of knowledge itself.

I think these proposals about what *X* “really is” are often implausible, as the concept of *X* places constraints on what it picks out. But in any case, I think that these theorists’ point could be put much more plausibly and powerfully by saying: “I don’t care about what *X* is. I just care about the associated explanatory role. And *X*₁ can play such-and-such crucial parts of that role.” On this way of putting things, it does not matter whether meaning really is teleological content (so the intuition that a swamp creature could speak meaningfully is irrelevant even if correct). What matters is whether teleological content can explain various phenomena that meaning was supposed to explain.

Of course roles themselves have to be expressed in language, and the danger of verbal disputes also arises in claims about roles. As always, one should always be prepared to give up any particular expression in stating these roles. Still, in practice, the expressions used in stating roles will often be relatively uncontested, where the expressions used in the original question were contested. So in practice, the move to roles can clarify matters greatly. In cases where potential verbal disputes also rise in stating the roles, one can repeat the process, in the hope that one will eventually find common terrain.

The model is not completely deflationary about conceptual analysis. On this model, the analysis of specific linguistic expressions and the associated concepts is relatively unimportant in understanding a first-order domain. But it is still interesting and important to analyze *conceptual spaces*. For example, it is especially interesting to determine which concepts are primitive, and to see how other concepts might be grounded in them. It is also interesting to investigate the space of concepts (and of the entities they pick out) that are relevant to a domain, determining which concepts can play which roles, what the relevant dimensions of variation are, and so on.²⁴

²⁴What about progress in conceptual analysis? One might take the discussion of ‘lie’ in section 2 of chapter 9 to suggest that this progress is always broadly verbal, and therefore of interest largely if we are interested in metalinguistic and metaconceptual matters. I do not think this is quite right: progress in analyzing the concept of a limit was clearly substantive, for example. But I think this progress is substantive rather than verbal precisely because of the many roles

For example, the conceptual metaphysics that I discuss in the nineteenth excursus turns on the investigation of conceptual spaces rather than of specific concepts. Here the aim is to find out which concepts are fundamental—that is, serve as the primitives that ground the entire space of concepts—and thereby to investigate what is metaphysically fundamental. But nothing here turns on claims about which concepts are associated with which expressions. Likewise, the sort of conceivability arguments that I used in *The Conscious Mind* are sometimes described as conceptual analysis, but it is made clear that claims about the meaning of the word ‘consciousness’ are inessential to the main arguments there: what matters is that there is some property we have that is conceivably absent in a physical duplicate, not that that property is consciousness. Importantly, these claims in conceptual metaphysics are robust even when subject to the method of elimination, whereas many claims deriving from traditional conceptual analysis are not.

This view leads naturally to a sort of conceptual pluralism: there are multiple interesting concepts (corresponding to multiple interesting roles) in the vicinity of philosophical terms such as ‘semantic’, ‘justified’, ‘free’, and not much of substance depends on which concept goes with the term. The model also leads to a sort of pluralism about the properties that these concepts pick out. For example, it naturally leads to semantic pluralism: there are many interesting sorts of quasi-semantic properties of expressions, playing different roles. It leads to epistemic pluralism: there are many different epistemic relations, playing different roles. It leads to gene pluralism: there are many different things that deserve to be called “genes”, playing different roles. The same goes for confirmation pluralism, color pluralism, and so on.

I am inclined to think that pluralism should be the default view for most philosophical expressions, at least once we set aside expressions for bedrock concepts. Typically there will be no single privileged role associated with such an expression, and different roles will be played best by different properties. The same holds if we move to sets of roles: different speakers will associate the term with different sets of roles, and different sets of roles will be played best by different properties. It may be that in some cases, a term is uniformly associated with a single role (or set of roles) in our community, or that one role is much more important than all the others, or that one property plays every one of the relevant roles better than any other property. But I suspect that such cases are relatively rare. Setting such cases aside, we should expect there to be many interesting concepts and properties in the vicinity of a given expression.

that the concept of a limit plays for us, for example in characterizing continuous functions, derivatives, and so on. The nonmetalinguistic progress consists in determining just what sort of mathematical property can play these roles. Thanks to Kit Fine for discussion here.

The picture that emerges is somewhat deflationary about the role of ordinary language philosophy in its many manifestations: 1950s Oxford philosophy, Canberra-plan analyses of folk concepts, contextualism and related theses in epistemology, contemporary linguistics-based philosophy, and some parts of experimental philosophy. It is not entirely deflationary about these endeavors: one can certainly move from linguistic premises to substantive conclusions, if one is careful enough about bridging premises and arguments. Still, if one is not careful, one will end up making points that reflect the vicissitudes of one's language rather than deeper philosophical truths.

For a more fine-grained analysis of the role of language, it is useful to focus on a specific sort of argument from ordinary language in philosophy. Perhaps the paradigmatic sort of argument here is a *disquotational argument*. Such an argument takes the following form:

Semantic premise: 'X' means Y.

Disquotational premise: 'X' means X

Conclusion: X is Y.

Here, the semantic premise is established through analysis of language and its usage, while the disquotational premise is a trivial premise about meaning, reference, or truth-conditions. The conclusion, however, appears to be a substantive first-order conclusion.²⁵

For example, something like the following argument is given by Stanley and Williamson (19xx) and Stanley (2011):

(1) 'S knows how to ϕ ' is true iff there is a way w such that S knows that w is a way of ϕ ing. (Semantic premise)

(2) 'S knows how to ϕ ' is true iff S knows how to ϕ . (Disquotational premise)

—————
(3) S knows how to ϕ iff there is a way w such that S knows that w is a way of ϕ ing.

²⁵ A closely related sort of argument adds a third *truth premise* as follows: (1) 'X' means Y; (2) 'X' means X; (3) X; therefore (4) Y. For example: (1) 'John believes something' means 'Exists p: John believes p' (2) 'John believes something' is true iff John believes something; (3) John believes something; therefore (4) Exists p: John believes p. In this way one might use a linguistic argument to establish the existence of propositions. Related arguments with a truth premise have been used to argue for the existence of events (Davidson), possible worlds (Lewis), and the truth of the A-theory of time (Ludlow). Everything I say in what follows about disquotational arguments applies with appropriate changes to these arguments.

Here, the first premise is offered as following from a standard linguistic analysis of “knowing-wh” constructions along the standard claim that “knowing how” is an instance of such a construction. The second premise is just an instance of Tarski’s T-schema. The conclusion, however, appears to be a substantive philosophical claim. It entails the “intellectualist” position that knowing how is a species of knowing that, as opposed to the “anti-intellectualist” position that knowing how differs from knowing that.

There is much to say about arguments of this sort. In effect, these arguments enable us to move from a premise about language to a conclusion about the world, providing a powerful response to the common charge that philosophy is about the world, and not just about language. Still, the arguments can be resisted in various ways. I will briefly discuss two ways of resisting that are closely tied to semantics, before discussing a deeper worry about the upshot of these arguments tied to the discussion of verbal disputes in chapter 9.²⁶

Most obviously, the semantic premise can be resisted by offering an alternative semantic analysis. Where semantic premises deriving from linguistics are concerned, it is especially relevant to observe that linguists’ semantic claims are often subject to fewer constraints than philosophers’ semantic claims. For example, linguists’ analyses often aim to capture inferences that are held to be valid by most or all competent users, whereas a philosopher might instead aim to capture inferences that are in fact valid. Likewise, a linguist might freely invoke abstract objects in their analyses, while some ontologically scrupulous philosophers might not. A philosopher might also give a role to empirical facts about science and naturalness in giving their analyses, while a linguist might not.

The linguists’ project (which we might call *semantics narrowly construed*) has the advantage that it allows the data of semantics to be relatively theory-neutral. But the cost is that it is far from obvious that this project yields genuine worldly truth-conditions. This point is recognized by many linguists. It has often been noted that on a natural model-theoretic account of the semantics of mass terms, derived from respecting speakers’ inferences, ‘water’ has distributive reference, which has the consequence that in all models, every part of water is water. Nevertheless, science suggests that some parts of water (hydrogen atoms, for example) are not water. Some conclude at this point that these model-theoretic analyses reveal more about speakers’ conceptions than about worldly truth-conditions.²⁷

²⁶The next four paragraphs, on the semantic premise and the disquotational premise, are not directly tied to the issues about verbal disputes here and can be skipped, but I think the issues here are interesting in their own right.

²⁷See for example Bunt 1979 and Lonning 1987, and Koslicki 1999 for criticism. Bach (1973) makes the case for

By contrast, the philosophers' project (*semantics broadly construed*) might yield worldly truth-conditions. But here the data will not be theory-neutral: it may be that for a broad semantic analysis of 'X', one first needs a thorough philosophical or scientific analysis of X. To know what 'water' refers to, one has to engage in the science of water; to know what 'knowledge' refers to, one has to engage in the philosophy of knowledge. But to the extent that a semantic analysis (of 'knowledge', say) depends on a philosophical analysis, we will not be able to use it to support that philosophical analysis. Although a disquotational argument may be sound in such a case, one's warrant for the semantic premise will rest on one's warrant for the conclusion, so that one cannot use the former warrant to generate warrant for the latter. More generally, an opponent of the first-order conclusion above, who holds that there are good reasons to hold that X is not Y, may hold that these are likewise good reasons to hold that 'X' does not mean Y, at least when meaning invokes worldly truth-conditions. This is not to deny that claims deriving from semantic analysis may have some independent support and therefore some force, but it suggests that establishing their force is less straightforward than it might seem.²⁸

One can also deny the disquotational premise by finding context-dependence in a key term, or by finding ambiguity or idiolectal variation. If different people use a key term in different ways, as arguably often happens in philosophy, a semantic analysis of one person's use may not apply to another's. Or following the worry above, one might hold that a semantic analysis yields something more like internal truth-conditions or model-theoretic truth-conditions than worldly truth-conditions, so that the disquotational premise does not apply.

The discussion in chapter 9 suggests another reaction to these arguments, though. An opponent "natural language metaphysics" as an area distinct from worldly metaphysics. Thanks to Zoltan Szabo and Brendan Balcerak Jackson for discussion of these matters.

²⁸On the two-dimensional picture deriving from chapter 5 and the eleventh excursus, one may at least be able to establish claims about primary intensions (of 'water', say) independent of empirical scientific analysis (of water, say). But the idealization of primary intensions suggests that a priori philosophical analysis will still play a crucial role here. For example, the same sort of a priori philosophical analysis that goes into answering the question 'What is knowledge?' will also be relevant to the question 'What is the primary intension of "knowledge"'. Furthermore, while primary intensions are a sort of worldly truth-condition, for many or most expressions (in particular, for epistemically nonrigid expressions), they must be combined with empirical truths about the world to yield truths about reference. If one appeals to fine-grained "analytic" meanings to avoid idealization, the latter problem becomes worse: fine-grained meanings may need to be combined with substantive a priori truths about the world to yield substantive conclusions about reference. Perhaps for certain strong sorts of acquaintance concepts, reference will be determined in an especially direct way that does not require a priori reflection; but for these concepts, there may not be much role for semantic analysis to play in revealing reference.

might simply accept the conclusion, while denying that it affects their substantive position. For example, suppose a compatibilist makes a successful argument that ‘free will’ refers to the ability to do what one wants, and infers that free will is the ability to do what one wants. Rather than rejecting the conclusion, an incompatibilist might respond by accepting it, and simply denying that free will (so understood) is tied to moral responsibility. If so, their substantive views about the connection between determinism and moral responsibility will be unchanged, and their position will differ only verbally from their previous one.

Likewise, suppose an intellectualist makes a successful argument that ‘S knows how to ϕ ’ invokes a sort of propositional knowledge and infers that knowing how is a sort of propositional knowledge. Rather than rejecting the conclusion, an anti-intellectualist might react by accepting it, and then saying that knowing how (so understood) is a relatively marginal phenomenon that does not play a central role in intelligent action, and that is grounded in abilities, which play the central role. If so, their substantive views about the character of intelligent action will be unchanged, and their position will differ only verbally from their previous one.

Generalizing the point: the mere fact that existing words like ‘know’ or ‘intentional’ or ‘see’ behave in a certain way does not suffice to settle substantive disputes about epistemology, action, or perception. After all, views based on these data may differ only verbally from views on which ‘know’ or ‘intentional’ or ‘see’ pick out something else, and which endorse apparently different first-order claims that use these words. For example, consider a contextualist view that says ‘know’ has two distinct referents in different contexts: knowledge_{low}, which we often have, and knowledge_{high}, which we never have. A proponent of this view may differ only verbally from a proponent of an apparently skeptical non-contextualist view on which ‘know’ always means know_{high}, or an apparently nonskeptical view on which ‘know’ always means know_{low}. Any substantive differences between the proponents will turn on what they say about further properties of the referents: for example, about the epistemic value and normative role of knowledge_{low} and knowledge_{high}. Assessing these further properties requires going beyond the linguistic data.²⁹

²⁹For more on applying the current methodology to contextualism, see the discussion linked at <http://consc.net/contextualism.html>. See also Sosa (2004), who uses considerations about verbal disputes to draw deflationary conclusions regarding certain aspects of contextualism. For related deflationary approaches to experimental philosophy, see Sosa (2007) and Scanlon (2010). In their “experimental philosophy manifesto”, Knobe and Nichols (2009) suggest that experimental philosophy aims not so much to discover the truth about intentional action, causation, and so on, but rather to discover how people think about these topics. I think that it is not out of the question for experimental philosophy to aim higher: perhaps it might help to establish truths about (what we call) intentional action, causation, and so on. But then the philosophical importance of the truths so established will be subject to the sort of critique given

This brings out a crucial point about disquotational arguments. For their conclusions to be dialectically effective, there must be a tacit *role premise* about the role of X that is accepted by both parties to a dispute.³⁰ In the case of free will, the role premise may be the claim that free will has a certain tie to moral responsibility. In the case of knowing how, the role premise may be the claim that knowing how plays a certain ubiquitous role as the grounds of intelligent action. In the case of knowledge, the role premise may be the claim that knowledge (or the variety of knowledge ascribed in ordinary contexts) has high epistemic value. If that role premise is fixed in place, then agreement on the conclusion of the disquotational argument (that X is Y) will yield agreement that Y (knowing propositions, be able to do what one wants, knowledge_{low}) plays the relevant role. But if the role premise is not fixed in place, an opponent can hold on to their previous position simply by denying it. Knowing how or free will or knowledge (ascribed in ordinary contexts) turn out not to play the role in question, but everything else is unchanged.

There may be some cost to denying the role premise: after all, the claim that free will is a condition on moral responsibility or that knowing how plays the relevant role may seem like common sense, and denying it may involve ascribing error to many ordinary speakers. Still, commonsense is not sacred in these domains. And in any case, some element of common sense may be preserved by noting that while these claims are not literally true, there are nearby claims that are true: the incompatibilist may hold that a certain incompatibilist property (which we commonly take to be free will) is a condition on responsibility, and the anti-intellectualist may hold that certain abilities (which we commonly take to be knowing how) play the relevant role in action.

In such a case, the key substantive issue is whether Y (knowing propositions, being able to do what one wants) plays role R (a role in action or in moral responsibility), or whether some other property Y' (abilities, an incompatibilist property) plays that role. In there are strong first-order reasons to hold that Y' plays role R, it is implausible that the disquotational argument above can defeat these reasons: instead, if one accepts the conclusion that X is Y, the first-order reasons will then give one reason to deny that X plays R. For example, if semantics suggests that commonsense temporal claims (A happened before B) are committed to simultaneity but physics tells us that nothing plays the role of simultaneity, it is clear that the evidence from physics wins out: one will

here.

³⁰The discussion in chapter 9 suggests that in the case of a bedrock expression (such as 'conscious', or 'right', perhaps) there need be no such role premise. But in these cases, it is not clear that semantic analyses do much to help give us a substantive grasp on reference. Furthermore, faced with a putative semantic analysis of such an expression, it will be open to an opponent to deny that the analysis is capturing the use of the expression to express the bedrock concept.

have to either deny the semantic analysis or deny the commonsense temporal claims. The same goes where strong philosophical arguments play the role of physics here. At best, the linguistic evidence deriving from the disquotational argument plus the role premise gives one some weak evidence that X plays R , evidence that may play a role when other first-order evidence is weak.

This leaves some roles for ordinary language philosophy. First, following on the analysis above linguistic claims may play a weak evidential role with respect to nonlinguistic claims. The conclusion of a disquotational argument (say, that knowledge is knowledge₁) may not yet substantively differ from that of someone else who says ‘Knowledge is knowledge₂’. But if we combine it with a role premise saying that knowledge is valuable, it follows that knowledge₁ is valuable. Still, philosophical counterarguments, such as arguments that knowledge₁ is not valuable, may give one reason to question either the semantic premise or the role premise, as we have seen. Linguistic claims are just one sort of evidence among many here, and one always needs to be explicit about just how the evidence relations work.

Second, there is considerable intrinsic interest in understanding the meanings of the words we actually use in philosophy and in understanding the concepts and the patterns of thought that we actually deploy. The language and psychology of philosophy are important topics in their own right. But there is much more to philosophy than the language and psychology of philosophy, and great care is required in moving from the latter to the former. For example, if it turns out that freedom₁ and not freedom₂ is picked out by our ordinary concept of freedom, so that compatibilism is true of the ordinary concept, then that gives some extra interest to freedom₁ and perhaps is evidence that this is the sort of freedom that we really value. But to answer the hardest questions about freedom, one still needs to determine the nonlinguistic properties of freedom₁: is it truly valuable, what is its normative role, and so on. The fact that freedom₁ is picked out by our ordinary word or concept does not settle these questions.

Third, the practice of ordinary language philosophy can lead us to interesting concepts and important distinctions. This is especially likely if the doctrine of “the genius of our tongue” is correct: our language is a tool sharpened by a history of use, so that one can expect that if a concept or a distinction is important, our language may well have captured it already. Still, this point gives ordinary language more of a role in the context of discovery than the context of justification. After using ordinary language to find concepts and distinctions, one still needs to justify substantive claims involving those concepts and distinctions.

Fourth, there are important normative questions about what expressions *ought* to mean. These

questions comprise what Peirce called “the ethics of terminology”. Ideal agents might be unaffected by which terms are used for which concepts, but for nonideal agents such as ourselves, the accepted meaning for a key term will make a difference to which concepts are highlighted, which questions can easily be raised, and which associations and inferences are naturally made. Following the “ameliorative” project of Haslanger (2005), one might argue that expressions such as ‘gender’ and ‘race’ play a certain practical role for us, and that role is played better by some conceptions than others, so ‘race’ and ‘gender’ ought to have certain meanings. The manifestly verbal dispute among astronomers about whether Pluto is a planet is best understood as a debate in the ethics of terminology: given the scientific and cultural roles that ‘planet’ plays, should ‘planet’ be used to include Pluto or exclude it? In philosophy, ‘meaning’ functions as something of an honorific (it attracts people to its study), so if one thinks that meaning₁ is more important than meaning₂, one might hold that ‘meaning’ ought to be used for meaning₁. Likewise, one might argue that ‘free will’ plays a useful practical role in judging the moral status of actions, or a pernicious role in supporting retributive punishment, so that we should use it for freedom₁ (so that we can attribute free will) or freedom₂ (so that we can deny it) respectively. The answers to these normative questions may depend on our purposes and values, but the questions themselves are certainly substantive.

So ordinary language philosophy is not unimportant. But one needs to be very clear about its use. If ordinary language philosophy is practiced in the material mode, as it often is, it is easy to move too quickly from linguistic data to substantive philosophical claims in a way that masks potential verbal disputes. Things work best when ordinary language philosophy is practiced in the formal mode, making claims about ‘freedom’ and not about freedom, for example, and being clear when these claims are descriptive or normative claims about actual usage. If one wants to draw non-linguistic conclusions while avoiding verbal disputes, one needs to be explicit about the bridge (a disquotational premise and a role premise, for example), and ideally one should be prepared to cast the conclusion without using the key expression. Doing things this way minimizes potential verbal disputes and maximizes clarity.

Eighteenth Excursus: Inferentialism and Analyticity³¹

In *The Logical Syntax of Language* (1934), Carnap developed a broadly inferentialist view of language. On an inferentialist view of language, the meaning of a sentence is understood in terms of its inferential role, especially with respect to other sentences, and perhaps with respect to perception and action. The meaning of other expressions is understood in terms of its contributions to inferential role of sentences containing the expression. On Carnap's inferentialist view, the meaning of a sentence is understood in terms of its consequences with respect to certain primitive sentences, as captured by certain transformation rules.

Subsequently, various theorists have developed inferentialist views of both linguistic content and mental content.³² An inferentialist of mental content holds that the content of a mental state (such as a belief or a judgment) is determined or individuated by its inferential relations to other mental states and their contents. Likewise, concepts or their contents are determined or individuated by their contributions to these inferential relations.

The scrutability framework is congenial to an inferentialist view of both linguistic and mental content. The framework of chapter 9 also naturally suggests an inferentialist view, of a somewhat more fine-grained sort. In this excursus I sketch an inferentialist view based on these frameworks, and I draw out a related way to cast an analytic-synthetic distinction. I will focus first on the case of mental content and of concepts.

To finesse the vexed ambiguity of the word 'concept' let us say that an *m-concept* is a concept qua mental item: m-concepts are the constituent of judgments, where judgments are mental acts. An *a-concept* is a concept qua abstract object: a-concepts are the constituents of propositions, where propositions are abstract objects. Propositions are the contents of judgments, and a-concepts are the contents of m-concepts. Then one inferentialist thesis holds that the content of an m-concept is determined by the m-concept's inferential relations to other m-concepts (perhaps along with relations to perception, action, and other entities). A closely related thesis says that a-concepts

³¹This excursus is very sketchy and poorly worked out at the moment. I will probably cut it down to something shorter for the purposes of the book, and expand it later into a separate paper on inferentialism that goes into all these matters in more depth.

³²Inferentialist views of linguistic content have been developed by Sellars (1953), who was strongly influenced by Carnap's inferentialism, Field (1977), Harman (1982), and Brandom (1994). Inferentialist views of mental content have been developed by Harman, Block (1986), and Peacocke (1993), among others. Views of this sort also go by the names "conceptual-role semantics" and "inferential-role semantics". I use "inferentialism" as a generic term here mainly because it is shorter than the other two labels and because it applies more naturally to mental content.

are individuated by their inferential relations to other a-concepts (perhaps along with relations to perception, action, and other entities), where these inferential relations are abstract counterparts of the concrete inferential relations that m-concepts stand in. For clarity, I will henceforth use “concept” (simpliciter) to mean m-concept, and will use “contents” to talk of a-concepts, the contents of m-concepts.

A *holistic* inferentialist holds that all inferences in which a concept is involved play a role in determining its content. A non-holistic inferentialist holds that only a special subclass of inferences do this. My own view is non-holistic: it is only the a priori inferences that matter, and perhaps a constrained subclass of these.

A *pure* inferentialist holds that there are no privileged concepts (or expressions) here: the content of every concept is determined only by its place in the web, and the web as a whole is characterized only by its abstract structure. Pure inferentialism is often regarded as implausible because it invokes a sort of wholly abstract structure that seems too abstract to account for the specific and substantive content of our beliefs and our language. An *impure* inferentialist holds that the contents of some primitive concepts are determined noninferentially (and that the corresponding a-concepts are individuated noninferentially), and that the content of all other concepts are determined at least in part by their inferential relations to these concepts.

Impure inferentialism might also be called *anchored* inferentialism, as the primitive concepts serve as anchors in which the contents of other concepts are grounded by inferential relations. The content of anchoring concepts will be determined in some other way: perhaps by causal or acquaintance relations to objects or properties in the world. The anchored inferential picture shares some of the spirit of a traditional descriptivist picture: there are primitive concepts, and nonprimitive concepts that derive their content in part from relations to them. But on this picture, the relation of deriving turns on inferential rather than compositional relations. Because of this, the picture is not threatened by the observation that most concepts cannot be decomposed into other concepts.

The scrutability framework is especially well-suited for anchored inferentialism. To a first approximation: the concepts involved in a general scrutability base will provide the anchors, and contents for all other contents will be determined by their a priori inferential connections to the concepts in the generalized scrutability base. Things go especially well if the base concepts are all

³²It is also possible that some anchoring concepts, such as the concept of negation, will get their content in virtue of inferential role. Here the relevant aspects of inferential role will be structural, as opposed to inferential connections to other specific concepts.

primitive acquaintance concepts: for such concepts, one can make sense of their contents without appealing to inferential roles. Then all other contents will be determined by their a priori inferential connections to these contents.

This model is illustrated by the way intensions for sentences are defined (in chapter 5 and the eleventh excursus). Given a truth-apt sentence S , its intension is a function from scenarios to truth-values such that S is true at a scenario w if $D \rightarrow S$ is a priori, where D is a canonical specification of w . Canonical specifications are always specifications in the vocabulary of a generalized scrutability base. So the intension of S is being defined in terms of the apriority of conditionals connecting S to base sentences. That is, the content of S is defined in terms of its a priori connections to base sentences. One can define intensions for subsentential expressions in a closely related way. Moving to the level of thought, the intension of an arbitrary concept is defined in terms of its a priori connections to certain thoughts using only primitive concepts. So in effect, the intensions of arbitrary concepts are grounded in a priori connections to base concepts. This is just as the anchored inferentialist supposes.

One might worry that intensions do not really correspond to *inferential* roles, as the definition of intensions requires that certain conditionals be a priori, and not certain inferences. When $D \rightarrow S$ is a priori, an inference from D to S might nevertheless be a bad one; as Harman (19xx) has observed, if one has good reason to believe $\neg S$, one might do better to retain that belief and drop D instead. However, for our purposes what is relevant are *suppositional inferences*: what one should infer on supposing that D is the case. We can say that S is a a priori suppositional consequence of D iff when one supposes D , one considers the question of S , and one reasons a priori, one should come to be certain of S (under an insulated idealization). Then if $D \rightarrow S$ is a priori, S is an a priori suppositional consequence of D , and vice versa. So intensions track a priori suppositional inferential roles, at least, and can in principle be defined in terms of these roles. There may be an approximate correspondence to other sorts of inferential roles, but these roles will not be constitutively relevant.

The inferential role of a concept can be construed as a normative role, constituted by *good* inferences that the concept might be involved in. For $D \rightarrow S$ to be a priori is not for the subject to actually connect D to S , or even to be disposed to do so. Rather, it is that the subject *should* accept this connection is reasoning correctly. So the content of S is given by normative facts about which inferences are good ones, rather than descriptive facts about what inferences the subject performs or accepts.³³

³³This normative character is shared with the inferentialism of Brandom 20xx. Differences include that my infer-

This inferentialist framework gives special weight to *entry inferences*: inferences from thoughts constituted by primitive concepts alone to thoughts involving the concept in question. The intension of S is determined by specifying those D for which $D \rightarrow S$ is a priori, and those D for which $D \rightarrow \neg S$ is a priori, which is in effect to say it is determined by entry inferences from sentences using primitive concepts. We might represent this inferential role as follows:

$$D_1 \vdash S$$

$$D_2 \vdash \neg S$$

...

S will also be involved in exit inferences: inferences from S to sentences involving primitive concepts. Here the simplest way to specify these inference is to take the contrapositive versions of the entry inferences: so $\neg S \vdash \neg D_1$, $S \vdash \neg D_2$, and so on (or alternatively, just two exit inferences with massively conjunctive consequences). Here, the entry inferences will determine the exit inferences, and vice versa, so we need only use one set to determine the sentence's content. I will focus on entry inferences. These entry inferences will determine the sentence's full a priori inferential role, and will determine its intension.

The same goes at the level of 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.

We can use this picture to address certain standard worries for inferentialism.

(1) *What is the relation between inferential role and truth-conditions?* On the current framework, an anchored inferential role determines an intension, which is a sort of truth-condition. Of

inferentialism is non-holistic (privileging a priori inferences), representational (giving a central role to truth-conditions), short-armed (the roles do not stretch into the environment), mind-first (privileging the mental over the linguistic) and impure, whereas Brandom's is holistic, nonrepresentational, long-armed, language-first, and pure (though it can also be understood as short-armed and impure, grounded in perceptual expressions whose content is then grounded in causal connections to the world). Carnap's inferentialism in *Logical Syntax* is arguably normative, non-holistic, nonrepresentational, short-armed, language-first, and impure (focusing on logical roles with respect to primitive sentences). Peacocke's inferentialism in *A Study of Concepts* (1993) is non-normative, non-holistic, representational, long-armed, mind-first, and either pure or impure (grounded partly in perceptual concepts which are themselves grounded in connections to the contents of perceptual experience, which serve as anchors).

course the anchors help here, as we are supposing that they have some sort of truth-conditional content that is not grounded in inferential role. Then inferential roles in effect ground the content of all concepts and expressions in the contents of anchors and inferential roles.

(2) *What is the relation between inferential role and wide content?* On the current view, anchored inferential role determines a primary intension. The relevant role can be seen as an internal (narrow or short-armed) role, so that the content is a narrow content. However, the primary intension of an expression such as ‘water’, on conjunction with the environment, determines its extension (H₂O) and so its wide content. On my view this relation holds quite generally. So wide content is fixed by internal inferential role and facts about the environment.

(3) *What about defective concepts?* The putative expression *tonk* is defined by saying that from A one can infer ‘A tonk B’, and that from ‘A tonk B’ one can infer B. It is commonly held that there is something defective about such expressions and such roles, and that an inferentialism theory needs to explain this and disallow them somehow. It is commonly observed that the problem with roles such as this one is that entry rules and exit rules are not in “harmony” with each other. The current approach avoids this issue by individuating inferential roles by entry rules alone, and making exit rules derive from the entry rules in the fashion discussed above, which ensures that harmony is preserved. I do not think that it is possible to possess a concept with a disharmonious inferential role.

(4) *Which has priority, inferential role or truth-conditions?* I am neutral on the important question of priority. Fixing normative inferential role and the content of anchors fixes the content of all contents and expressions, but this fixing relation may or may not reflect a grounding relation. The key question here is whether the norms of inference are themselves grounded in the content of the concepts, or whether the content of the concepts are grounded in the norms of inference. I am unsure to the answer of this question about the relative priority of the normative and the intentional here.

An anchored inferentialist picture of concepts is also suggested by the model of verbal disputes and bedrock concepts in chapter 9. An anchored inferentialist will hold that the content of an expression such as ‘go round’ is grounded in the content of inferentially related expressions, such as expressions concerning relative location and motion. On the current picture, when one applies the method of elimination to ‘The squirrel goes round the tree’, one is in effect cashing out aspects of the content of ‘go round’, at least insofar as that content matters for dialectical purposes, in terms of the content of various associated expressions, expressing associated concepts. That is just what an inferentialist would expect. If one applies the method repeatedly, one will cash out aspects

of the content of ‘go round’ in terms of the content of various associated bedrock expressions, expressing bedrock concepts. That is just what an anchored inferentialist would expect. In effect, the method of elimination consists in moving to inferentially related concepts within the web of concepts, and ultimately to the primitive concepts that serve as anchors.

The picture in chapter 9 also helps with a worry about the inferentialist picture above. The worry is that a priori inferential roles are too coarse-grained to fix the content of concepts and expressions. After all, two distinct concepts can be a priori equivalent: *the 17th prime* and 59, for example, or perhaps a moral concept (*right*) and a nonmoral concept that is coextensive with it (*maximizes expected utility?*). These will have the same a priori inferential roles, but they seem to be different concepts with different contents.

Here we could take the line that a priori inferential roles at least fix a certain coarse-grained content for m-concepts, while saying that separate considerations fix fine-grained content. Or we could appeal to structure in the m-concepts to yield fine-grained structured content, as we did in the linguistic case in the eleventh excursus. But it is especially attractive to appeal to a constraint on inferential roles stronger than apriority, more akin to a kind of analyticity.

Here the picture in chapter 9, with all truths translucently settled by bedrock truths, is suggestive. On this picture the anchoring concepts are more numerous than on the a priori scrutability picture: they include normative/evaluative and some mathematical concepts, for example. And the inferential relation between bedrock concepts and other concepts is tighter than just a priori inference—it is more akin to a sort of analytic inference. At least the phenomena in chapter 9 can be regarded as evidence for an inferentialist model on which the bedrock concepts serve as anchors, and on which all other concepts are connected by a sort of quasi-analytic inference.

We need not presuppose a notion of analyticity here (although later we will say how this picture might be used to define a notion of analyticity). Instead, we can take the phenomena of chapter 9 as evidence as strong evidence that all concepts (in the sense of a-concepts) are individuated by certain inferential connections to bedrock concepts, where the inferential relations are stronger than mere a priori connections. They will be individuated by patterns of entry rules, like those discussed earlier but more fine-grained, and corresponding patterns of exit rules. Now there is no reason to suppose that the pairs of a priori equivalent concepts above must have the same entry and exit rules. For example, *right* may be a bedrock concept, while *maximizes expected utility* will be a non-bedrock concept itself connected to probabilistic, mathematical, and evaluative concepts.

There will be many different ways to represent *S*’s inferential role. Ultimately we will still want to be able to determine entry rules for *S* from the (now enriched) base language alone. But

there will typically be much simpler representations of the entry rules than those that go scenario-by-scenario. In many cases, there will be much shorter sufficient conditions for S that can be specified in the base language, in effect covering a huge range of scenarios at once. One can also specify entry rules without using the base language, as when we specify the meaning of ‘vixen’ by entry rules such as

x is a female fox $\vdash x$ is a vixen

x is not a female fox $\vdash x$ is not a vixen

As long as there are further entry rules that ground expressions such as ‘female’ and ‘fox’ in other expressions and ultimately in the base language, these rules will collectively determine entry rules for ‘vixen’ starting from the base language. This sort of representation, when it is available, will give a much better representation of the structure of the concept and the core inferences it is involved in.

Of course most concepts are not well-behaved enough to be summed up in a couple of clean inferences such as those above. For a typical concept, as we have seen, such inferences are very often defeasible and subject to counterexample. But this leaves open the possibility of specifying the inferential role of a concept using defeasible inferences. For example, one might specify the inferential role of *knowledge* as follows:

x is not a belief $\vdash \neg K(x)$

x is not true $\vdash \neg K(x)$

x is not justified $\vdash \neg K(x)$

x is a justified true belief $\Vdash K(x)$

x is inferred from a falsehood $\vdash \neg K(x)$

x is true by luck $\Vdash \neg K(x)$

...

Here the idea is that we have a hierarchy of conclusive inferences (marked by ‘ \vdash ’) and defeasible inferences (marked by ‘ \Vdash ’). To classify a given case as knowledge or not, one starts at the top of the list and goes down until one comes across an entry rule whose left-hand side is satisfied by the case. If that rule is conclusive, one classifies the case according to the right-hand side and stops. If that rule is defeasible, one tentatively classifies the case according to the right-hand

side and keeps going. If the case is eventually classified conclusively, that classification counts; otherwise the last defeasible classification counts. The rules should be such that every case is classified eventually (alternatively, those that are not are indeterminate). The left-hand sides will likely get longer and longer, and the list may be infinite, as the list of scenario-by-scenario entry rules are. These rules will ultimately determine the scenario-by-scenario rules, at least on conjunction with entry rules for the other expressions in question. But this specification will indicate the core inferential structure of the concept much more transparently than the scenario-by-scenario rules.

This model is doubtless oversimplistic and would need to be refined in various ways, for example to capture more complex dependencies among conditions of application for a concept. But it gives the idea of a way of representing entry rules on an inferentialist model like the current one that allows some recognition of conceptual connections between high-level concepts will still recognizing the inadequacy of a definitional model. As a bonus, this model fits well with the dynamic program of conceptual analysis set out in the previous excursus. In effect, that model might be seen as gradually articulating the constitutive conclusive and defeasible entry rules for a concept, achieving ever-better approximate conceptual analyses.

What sort of psychological reality does this model have? That is, is any set of rules, however gruesome, that ultimately yields the same conditions of application an equally valid representation of the concept? Or is there something privileged about some representations, such as the one above? I am not sure of the answer to this question, it is attractive to hold that some inferences at least have a sort of normative privilege: they are somehow directly warranted by conceptual structure alone, perhaps, where other more gruesome inferences that the concept is involved in are not.

This model goes naturally with a certain understanding of analyticity that invokes the framework of warrants discussed in the fourth excursus. We can say that *S* is *warrant-analytic* for a subject if there is a *conceptual warrant* for the subject to accept *S*. We can say that *S* is warrant-analytic (simpliciter) 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.

If we follow the model above, in some cases constitutive inferential roles will be defeasible: the inference from *x is a justified true belief* to *x is knowledge*, for example. In these cases, subjects will have at best a defeasible warrant for the relevant inferences, and any further claims grounded in these inferences—*all justified true beliefs are knowledge*, say—should not count as analytic. But in the case of conclusive inferences, it is reasonable to count the resulting claims as warrant-analytic.

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 deriving their warrant from concepts as more primitive states.

A final question is whether we might be able to use the inferentialist picture outlined here to

help provide a reductive account of mental content. I do not have such an account to offer, but if something like the account above is correct, it at least suggests certain directions to pursue. We have seen that once we fix the content of anchoring concepts and certain normative facts about inferential connections between these concepts and non-anchoring concepts, then the content of other concepts is fixed. So one might at least try out the suggestion that the latter content is grounded in the former contents and in the normative facts. Once that model is accepted, it reduces the question of the grounds of mental content to the questions of (i) the grounds of the relevant normative facts about inference (in particular, the fact that there is an a priori warrant for certain suppositional inferences between anchoring concepts and other concepts) and (ii) the grounds of the content of anchoring concepts.

Regarding (i), there are various options. First, one might simply take normative facts as primitive. Second, one might endorse Brandom's "normative phenomenalism" according to which certain inferences count as good ones roughly in virtue of our treating them as good inferences. Third, one might try to reduce the normativity by giving a reductive analysis of rationality: we could then at least approximate the normative inferences by equating them with the inferences that one would make if one were ideally rational. Fourth, one might try to ground the normative facts about inference in certain non-normative facts about inference—most obviously, the fact that we are disposed to make certain inferences—along with further factors in order to avoid problems concerning error and incompleteness here. Fifth, the normative facts might be grounded in prior facts about the content of the non-primitive concepts, along with quite general norms of inference between related concepts.³⁴ I am inclined to give the most weight to the fourth option (which will then yield as a reductive dispositional inferentialism, not just a nonreductive normative inferentialism), but I take all the options but the first seriously.

Regarding (ii), there are also various options. I am inclined to divide the relevant concepts into structural concepts and acquaintance concepts. For structural concepts, such as logical concepts and perhaps nomic and fundamentality concepts, it is not out of the question that these will themselves be grounded in some sort of structural inferential role: not grounded in connections to other bedrock concepts as much as in certain structural forms of inference. For example, there are stan-

³⁴Of course the fifth view then raises the question of where these prior contents are grounded. Presumably these are not grounded in norms of inference. Holding that they are grounded in inferential dispositions in effect reduces this option to the fourth option. On my view this content is narrow, so it will not be grounded in causal connections to the thinker's environment. A remaining possibility is that it is grounded in the phenomenology associated with the non-primitive concept.

standard inferentialist approaches to logical concepts. There are broadly inferentialist approaches to normative concepts, grounding them in certain connections to judgment and action. One might try an inferentialist approach to nomic concepts, for example, grounding these in norms of inference from observed cases to laws? One might also try something similar for fundamentality concepts, perhaps centering on their role in explanation.

Acquaintance concepts may include phenomenal concepts and observational concepts: primitive concepts of phenomenal properties, spatiotemporal properties, and secondary qualities. Here I think it is plausible that our concepts of these properties derive from our relation to these properties in experience: from the way spatiotemporal properties and secondary qualities are represented in perceptual experience, for example, and from the way that experiences are presented to us in introspection. Of course that then shifts the question to the grounds our relations to these properties in experience. One might hold that causal relations play a role here, or one might hold that Russell-style acquaintance is central. It is also not out of the question to hold that somewhere around here, intentional content (for example, primitive awareness of the relevant Edenic properties) is primitive and not grounded in anything more basic.

I return to these issues about the naturalization of inference-grounded content briefly in the twentieth excursus. Of course both the inferentialist view here and the speculation about naturalization need much more development to be taken seriously as theories of the grounds of mental content. But the scrutability framework at least provides some promising avenues to pursue.