

# Chapter 4

## Substantivity

**S**TRUCTURE IS THE KEY to understanding an elusive notion of substantivity.<sup>1</sup>

### 4.1 Nonsubstantive questions

Is Robinson Crusoe a bachelor? Is a water glass a cup? Is a protrusion from the floor of a deep ocean, whose tip is a tiny island, a mountain (Hawthorne, 2006b, vii.)? Is “some nonsense made out of sour green apple liqueur”, served in a V-shaped glass, a martini (Bennett, 2009)?

Philosophers like to argue about such things at bars, but even they regard the questions as being, in some sense, *nonsubstantive* (shallow, nonobjective, conventional, terminological). By contrast, the question of whether electrons repel one another is substantive (deep, objective, nonconventional, about the world).

This is intuitively clear; but in what does the lack of substantivity consist? The answer is not straightforward. It’s not a matter of mind-dependence, for example. Suppose, for the sake of argument, that Hawthorne’s protrusion is indeed a mountain. Then, even if no people had existed, the protrusion would still have been a mountain; and if everyone had said ‘Let such protrusions not be mountains’, the protrusion would still have been a mountain. To paraphrase Lincoln, calling a tail a leg wouldn’t make it one.

(The protrusion wouldn’t have been *called* a ‘mountain’ if people had not

---

<sup>1</sup>Related work includes Chalmers (2011), Fine (2001) on nonfactualist discourse, and Sidelle (2007) on verbal disputes.

existed, or had spoken or acted in certain ways. The linguistic fact that the sentence ‘The protrusion is a mountain’ is true *is* mind-dependent, since whether the sentence is true depends on what it means, and what it means is mind-dependent. But the same goes for every true sentence, even ‘Electrons repel one another’. So the mind-dependence of the linguistic fact isn’t what makes the question of whether the protrusion is a mountain nonsubstantive.)

Nor is the lack of substantivity due to the questions being *about* words or concepts—not in the most straightforward sense anyway. Questions such as “Is ‘attorney’ synonymous with ‘lawyer’?” and “Does the concept of a sport include gymnastics?”, which name or quantify over words or concepts, are in a straightforward sense about words or concepts; but our nonsubstantive questions are not about words or concepts in this sense. The question of whether martinis can be made of sour apple liqueur is about *martinis*, and what they can be made of. It is no more about ‘martini’, or the concept of a martini, than the question of whether electrons repel one another is about ‘electron’ or the concept of an electron.<sup>2</sup>

Nor is a question’s nonsubstantivity due to its having no answer—to there being no fact of the matter as to what its answer is. For the nonsubstantive questions considered above might well have answers. Gettier (1963) uncovered a surprising feature of our concept of knowledge: that justified true belief does not suffice for its application. For all we know, clever gettiers could convince us beyond the shadow of a doubt that our concepts of bachelor, cup, and the rest yield answers to our questions. (As the case of knowledge shows, this possibility isn’t ruled out by the failure of a few hours or centuries of armchair reflection.) But the questions still seem nonsubstantive.

What if there will be no Gettier-like successes in conceptual archaeology? It’s arguable that our questions then have no answers. The main opposition to this conclusion comes from the epistemists about vagueness (Sorensen, 2001; Williamson, 1994), who say that even without the facts that gettiers uncover—at least in part, surprising facts of usage that nevertheless are readily appreciated once noticed—the questions have answers.<sup>3</sup> But if we reject epistemicism—as I think we should—then we may well want to say that the questions without lurking gettiers have no answers. Where our conceptions of bachelors, cups, and mountains run out, so to speak, there are no “further facts” to be discovered.

---

<sup>2</sup>Compare Williamson (2007, chapter 2).

<sup>3</sup>Though they may have no “determinate” answers, given an epistemic reading of ‘determinate’.

Failure to have an answer is thus sometimes symptomatic of nonsubstantivity. But since it is not necessary for nonsubstantivity, it is not the underlying cause.

The failure of the questions to be substantive, then, is hard to pin down; it's neither mind-dependence, nor a conceptual or linguistic subject matter, nor the lack of an answer. This might remind one of a problematic familiar from discussions of antirealism.<sup>4</sup> Begin, for example, with expressivism about morality, according to which the function of assertions about right and wrong is to convey the speaker's attitudes, rather than to "depict objective facts". But make the expressivism sophisticated, so that it applies to logically complex sentences; and further, adopt a deflationary approach to truth and related concepts. Then you will be willing to say things like "moral sentences can be true" (since "'killing is wrong' is true" boils down to "Killing is wrong", and you're willing to say the latter because you're down on killing), "there are moral facts" (since "it's a fact that it's wrong to kill" boils down to "Killing is wrong"), and "there are moral properties (since "killing has the property of being wrong" boils down to "Killing is wrong"). In what, then, does your moral antirealism consist? There are indeed analogies here with my examples. The nonsubstantivity of my questions does not consist in discourse about bachelors (cups, etc.) failing to be truth-apt, or there failing to be facts about bachelors, or there failing to exist a property of being a bachelor. But the analogy quickly breaks down, since an expressivist semantics for discourse about bachelors, cups, and the rest is clearly a nonstarter. Nor is J. L. Mackie's (1977) error theory a good model: our beliefs about bachelors, cups, and the like are surely not radically mistaken (not because of their nonsubstantivity, anyway).

## 4.2 Substantivity characterized

I think that our questions are nonsubstantive because their answers turn on which of a range of equally good available meanings we choose for the words in those questions.<sup>5</sup> Better (though this will need further refining): for one or more expressions  $E$  (e.g. 'bachelor') in a nonsubstantive question, the semantic candidates for  $E$  (unmarried-adult-male, unmarried-adult-male-eligible-for-marriage, etc.) are such that i) each opposing view about the question comes out true on some candidate; and ii) no candidate carves at the joints better than

---

<sup>4</sup>See Blackburn (1993); Fine (2001); Rosen (1994); Wright (1992).

<sup>5</sup>The account requires there to be some way of making sense of quantification over "meanings", but not that such quantification be fundamental.

the others. If  $E$  uniquely means one of these candidates, this is not because that candidate is intrinsically privileged. It is only because our linguistic community happened to select that candidate rather than one of its mates as the meaning of  $E$ .<sup>6</sup> The situation may be depicted thus:

$$E \quad \left. \begin{array}{l} m_1 \\ m_2 \\ m_3 \\ m_4 \\ m_5 \end{array} \right\} \begin{array}{l} \text{no candidate carves at the joints} \\ \text{better than the rest; the} \\ \text{question's answer turns on} \\ \text{which candidate is adopted} \end{array}$$

Given what we happen to mean by  $E$ , the nonsubstantive question may have an answer, but we could just as easily have meant something else by  $E$ , something equally good, one of the other non-joint-carving candidates in the vicinity, in which case the question would have had a different answer.

A substantive question is one that is not nonsubstantive. The answer to a substantive question is not sensitive to a choice amongst equally joint-carving candidates. Now, one way for this to happen is for the question to be cast in perfectly joint-carving terms (and for none of the expressions in the question to have *multiple* perfectly joint-carving candidates). Each expression  $E$  in the question may then be depicted as follows:

$$E \quad \left. \begin{array}{l} m_1 \\ m_2 \\ \mathbf{m} \\ m_4 \\ m_5 \end{array} \right\} \begin{array}{l} \text{One candidate is special; it} \\ \text{carves at the joints much better} \\ \text{than the others} \end{array}$$

But this is not the only way for a question to be substantive. I drank no alcohol on January 1, 2011; thus, even though the question of whether I had a martini that night is not cast in particularly joint-carving terms, it is substantive because it has the same answer (no) under all candidates for its terms. Being cast in joint-carving terms is thus not the sole determiner of substantivity. Being wholly cast in perfectly joint-carving terms is normally sufficient for substantivity (since

---

<sup>6</sup>How did we select the candidate? Perhaps it is the most charitable candidate given our use of  $E$ ; perhaps only it is relevantly causally related to our use of  $E$ ...—the answer depends on the truth about metasemantics.

rarely does an expression have multiple joint-carving candidates), but it's not necessary.

Rival conceptions locate nonsubstantivity in some defect of the proposition in question—for example, mind-dependence, failure to have a determinate truth-value, and so on. On my conception, even questions without such defects can be nonsubstantive. The question of Hawthorne's protrusion is nonsubstantive even if ‘the protrusion is a mountain’ expresses a true mind-independent proposition about the physical world. The nonsubstantivity is not due to any defect of this proposition, but rather to the proposition’s “metasemantic surroundings”—the other propositions that the sentence could have expressed had we meant a different candidate for ‘mountain’. It is the failure of the actually expressed proposition to stand out from its surroundings (and the fact that it differs in truth-value from some of the surrounding propositions) that generates the nonsubstantivity. Put another way, the nonsubstantivity results from our process of selection of the proposition, rather than being intrinsic to the proposition itself.

What follows is a series of refinements and amplifications of this account. This process is not intended as a conceptual analysis of **substantivity**, thought of as pretheoretically given. The aim is rather to introduce a concept that sheds light on the phenomena. Also, this concept is not intended to apply to everything that might justly be called “nonsubstantive”. For example, it isn't meant to apply to equivocation between distinct lexical meanings (as in a dispute over whether geese live by “the bank”, in which one disputant means river bank and the other means financial bank), or disputes involving expressivist language. Nor is it meant to capture the shallowness of inquiry into whether the number of electrons in the entire universe is even or odd (an inquiry that is substantive in my sense, but pointless). My goal is simply to identify one distinctive—and neglected—type of nonsubstantivity.

On to the amplifications and refinements. First, I will speak of substantivity for many items: sentences, questions, disputes, and so on. The definition may be adjusted in obvious ways for these items. Examples: a question (construed as the set of sentences that are its possible answers—this was implicit above) is nonsubstantive iff the candidates of some expression are equistructural and each answer comes out true under some candidate; a sentence is nonsubstantive iff the candidates of some expression are equistructural and both the sentence and its negation come out true under some candidate.

Second, I said earlier that being cast in perfectly joint-carving terms normally suffices for substantivity. But being cast in *highly* albeit not perfectly

joint-carving terms—a common occurrence in the special sciences—also normally suffices for substantivity. Except for questions that strain the boundaries of taxonomy (a relatively uncommon occurrence), special-science questions normally fall into one of the following two categories: i) each expression has a candidate meaning that carves far better than all other candidates; or ii) each expression has a range of candidates that carve far better than do other candidates not in the range, and the question’s answer is insensitive to choices of candidates within these ranges. In either case, the question is substantive.

Third, some disputes don’t seem substantive even though there is a unique joint-carving candidate, if that joint-carving candidate is of the wrong sort. Recall Hawthorne’s physically significant line in the Ural mountains (section 3.2), and consider the question of whether a certain location near the line is part of Europe. The answer to this question turns on which of several candidates we mean by ‘Europe’. Now, given the line, one of these candidates carves at the joints far better than the rest, and so my account classifies this dispute as substantive. But that seems wrong. The line through the Urals, though physically distinguished, isn’t geographically or politically distinguished, and therefore seems irrelevant to the boundaries of Europe. The physically significant line is indeed a joint in nature, but it’s the wrong sort of joint in nature, relative to a dispute over ‘Europe’.

What is the “right sort”? I don’t have a finished answer, but here is the beginnings of one. Suppose an expression,  $E$ , is a theoretical term, in the sense of section 3.2—a term intended to stand for a joint in nature. Suppose, further, that there is a single joint-carving meaning,  $m$ , that satisfies enough of the “core theory” that is collectively associated with  $E$  by the participants in the dispute. Then  $m$  is the right sort of joint-carving meaning. For example, in a dispute over whether any electrons are located in a certain region,  $R$ , of space, the term ‘electron’ is a theoretical term, and the property of *being an electron* will satisfy enough of the core theory associated with ‘electron’ by the disputants—a theory saying that electrons are subatomic particles that orbit nuclei, have negative charge, a certain mass, and so on. Thus, this property will be the right sort of joint-carving candidate meaning, and the dispute is substantive.

It’s important that the right sorts of joint-carving meanings needn’t satisfy *all* of the core theory, since there can be substantive disputes even when the core theory is somewhat mistaken. The dispute over whether any electrons are located in  $R$  remains substantive even if scientists are mistaken about the mass of an electron. The property *being an electron* remains the right sort of

joint-carving meaning because it satisfies enough of the core theory. Also, there can be substantive disputes over which central principles electrons satisfy—over what the mass of an electron is, for example. In such cases the core theory that is “collectively” associated with ‘electron’ will be drawn from the principles not under dispute. To have a substantive dispute about electrons, there must remain enough common ground about what electrons are so that all disputants can be regarded as talking about the same thing.

What is the “wrong sort” of joint-carving candidate? Hawthorne’s line is the wrong sort to render a dispute about ‘Europe’ substantive because ‘Europe’ is not a theoretical term; ‘Europe’ isn’t intended to stand for a joint-carving meaning at all. It might be objected that ‘Europe’ *is* a theoretical term, albeit one of political science or geography rather than physics. But then the physically significant line will still be of the wrong sort, since it won’t satisfy the associated core theory. That core theory will be a political or geographical theory, not a physical theory, and the physical line won’t play a role in any distinctively political or geographical laws or explanations. (If it did, then it wouldn’t be merely “physically significant”, and the dispute would then be substantive.)

The revised account, then, says that a nonsubstantive question is one containing an expression  $E$  whose candidates are such that i) each opposing view about the question comes out true on some candidate; and ii) no candidate carves at the joints *in the right way for E* better than the rest. A candidate  $c_1$  carves better “in the right way for  $E$ ” than another candidate  $c_2$ , to a first approximation anyway, iff  $E$  is a theoretical term,  $c_1$  satisfies enough of the core theory associated with  $E$ , and  $c_1$  carves better than does  $c_2$ .

One might worry about the fact that this definition makes *all* differences in joint-carving for candidates of nontheoretical terms irrelevant to substantivity. Recall the question of whether I had a martini on January 1, 2011, which is intuitively substantive (and has the answer *no*) because I drank no alcohol that night. The current definition counts this question as *nonsubstantive* if its answer is sensitive to the choice of candidates for its nontheoretical terms, regardless of distinctions of joint-carving amongst those candidates. But won’t there be Putnam-like candidates, based on arbitrary permutations, under which the question’s answer is yes? Perhaps it would be enough to reply that Putnamian semantic values are not candidates because they just couldn’t be meant by any linguistic community. Or perhaps—the more likely case, I fear—a subtler definition of “the right sort” is needed, no doubt based on a subtler distinction than that between theoretical and nontheoretical terms.

Fourth, ‘candidate’ needs to be clarified. On the one hand, it shouldn’t

be taken too narrowly, to include only what supervaluationists call “precisifications”. Precisifications of  $E$  are, intuitively, those semantic values that our usage of  $E$  doesn’t definitely rule out. But suppose that a gettier could show us that our usage of ‘bachelor’ definitely excludes Crusoe. Then semantic values under which ‘Crusoe is a bachelor’ is true aren’t precisifications of ‘bachelor’, but they should count as candidates for ‘bachelor’ (since whether Crusoe is a bachelor is paradigmatically nonsubstantive). So candidatehood is consistent with a certain degree of “mismatch with usage”. On the other hand, match with usage isn’t *irrelevant*. The property of being an electron fails to be a candidate for ‘bachelor’ precisely because it doesn’t come close to fitting our usage of that term. A candidate meaning  $m$  needn’t perfectly match our usage of  $E$ ; but the mismatch can’t be too severe. If a linguistic community, roughly in our circumstances, *could* have used  $E$  to mean  $m$  without seeming “semantically alien”—could have used  $E$  to reach “the same semantic goal” as we use  $E$  to reach, albeit perhaps by a different route—then  $m$  is a candidate for  $E$ .

This is admittedly pretty vague. Now, it’s fine if ‘substantive’ and related notions are rough and ready, come in degrees, admit borderline cases, and so on. Still, to firm up the notion a bit, consider another example (there will be more later). Suppose, for the sake of argument, that i) causation does not carve at the joints; and ii) in English, ‘cause’ has a counterfactual analysis. Now consider a linguistic community  $L$  in which ‘cause’ has a covering-law analysis, but in which the *conceptual role* of ‘cause’ is similar to its role in English, in the following sense. Causation, in both our community and  $L$ , stands in a complex network of conceptual relations to concepts of moral responsibility, control, and myriad others. Although the members of  $L$  use ‘cause’ differently from how we do, there are also corresponding differences in their usage of ‘responsible’, ‘control’, and other such terms, so that the network of relations is preserved. I want to count this linguistic community as not being (too) semantically alien. Given the similarity of conceptual role, speakers of  $L$  use ‘cause’ with the same semantic goal as we do with our word ‘cause’. Thus the covering-law meaning is a candidate for ‘cause’; and as a result, assuming that causation doesn’t carve at the joints, the question of whether effects counterfactually depend on their causes is nonsubstantive.

Fifth, my notion of substantivity is a metaphysical one, to be contrasted with a notion of conceptual substantivity. Many expressions that fail to carve at the joints are embedded in our conceptual lives in important ways; and questions involving such expressions can have a sort of conceptual substantivity even when my analysis counts them as (metaphysically) nonsubstantive. Suppose again

that the fact that causes counterfactually imply their effects is nonsubstantive in my sense. Still, discovering that causation is counterfactual dependence might reveal something important about our conceptual scheme, in contrast with a discovery that cups are glasses, which we would regard as an intellectual trifle. Given the network of connections between causation and other concepts, the discovery about causation would have far-reaching implications. And causation matters to us in ways that the concept of a cup does not—partly because of the network. True, in learning that causes counterfactually imply their effects, we are primarily learning something about ourselves; that's the metaphysically nonsubstantive part. But we're learning something important about ourselves.

Sixth, my notion of substantivity is essentially metalinguistic. Consider:

- (1) Drinks with sour apple liqueur are not martinis.

In my opinion, (1) is true; ‘martini’ means *drink made of gin or vodka and vermouth with such-and-such proportions*. (If you disagree, replace (1), mutatis mutandis, with its negation.) But we could have used ‘martini’ differently, so as to include sour apple liqueur drinks, without being semantically alien or carving worse at the joints. So (1) is nonsubstantive. But now consider:

- (2) Drinks with sour apple liqueur are not drinks made of gin or vodka and vermouth with such-and-such proportions.

(2) is substantive. Although the terms in (2)—‘sour apple liqueur’, ‘drink’, ‘gin’, and so on—do not carve at the joints, they have no candidates under which (2) comes out false. But (2) is just the result of substituting for ‘martini’ in (1) an expression that has the same meaning (though not the same candidates). Moral: sentences that express the same proposition can differ in substantivity. This is a feature of my account, not a bug. Notions like substantivity, depth, objectivity, and the rest have proved so elusive partly because philosophers have been looking in the wrong place: in what we say with the disputed vocabulary.<sup>7</sup> Substantivity, or lack thereof, is not intrinsic to semantic values. We might put this by saying that we should look to metasemantics, rather than semantics, to reveal substantivity.

---

<sup>7</sup>Thus I deny Gideon Rosen's (1994, p. 301) claim that “if the facts in the contested class can simply be read off in a mechanical way from the facts in an uncontroversially objective class, then there can be no grounds for denying the same status to facts in the contested area.” Adherence to this principle leads Rosen to his pessimistic conclusion that “it adds nothing to the claim that a certain state of affairs obtains to say that it obtains objectively” (p. 279).

Seventh, substantivity can depend on the world, and not just on the words involved and what candidates they have. Consider the question of whether there is lithium on Mars. It is natural to think that some expression in ‘There is lithium on Mars’ has a range of equally joint-carving candidates corresponding to the “fuzziness” in Mars’s spatial boundaries. For the sake of definiteness, suppose this to involve a range of mostly overlapping candidate referents of the name ‘Mars’, which differ from one another by including slightly different parts near Mars’s vague boundary. The question of whether there is lithium on Mars is nonsubstantive, then, if and only if some, but not all, of these candidate referents contain lithium. Now, if there is, in fact, plenty of lithium in Mars’s core, so that all the candidate referents contain lithium, then the question is substantive, since its answer is yes under each candidate. But if the only lithium in the vicinity of Mars is a single lithium atom at its vague border, then the question is nonsubstantive. So the substantivity of the question depends on a fact about the world: the location of lithium. Still, regardless of the location of lithium, the question “admits” nonsubstantivity because without any change in meaning or candidates, the question could have been nonsubstantive: there might have been lithium just at Mars’s vague borderline. In contrast, “Are there charged particles?” does not admit nonsubstantivity, and “Is it possible for a martini to be made of sour apple liqueur?” does not admit substantivity.

Eighth, it’s natural to think of substantivity as depending on the interests of disputants. Suppose two scientists look at an atom through a telescope and disagree over whether it is lithium. In fact, the atom *is* lithium. Also, no other lithium atom is in the vicinity of Mars, and the scientists know this. Finally, the atom at which they are looking is in Mars’s vague border. But the scientists don’t know this last fact; indeed, they falsely believe that the atom is definitely on Mars’s surface. As a result, they phrase their debate thus: “Is there any lithium on Mars?”. For they believe (falsely) that the answer to this question is definitely *yes* if and only if the atom seen through the telescope is lithium. Now, my official account says that this question is nonsubstantive (its answer turns on which candidate for ‘Mars’ is meant). But intuitively, what is at issue is substantive, since the scientists don’t care whether the atom is on Mars; they’re just using the question of whether there’s lithium on Mars to get at the question of whether the atom seen through the telescope is lithium. We might account for this by distinguishing the question the scientists actually ask (‘Is there lithium on Mars?’) from the question they really care about (‘Is the atom seen through the telescope lithium?’), and applying the official account, as-is, to these questions. Alternatively, we might alter the official account by treating

substantivity as a property of question–context pairs, rather than a property of questions simpliciter. The idea would be to make the set of relevant candidate meanings for the question depend on what issues are treated as important in the context. In the context of the scientists, different candidate meanings for ‘Mars’ are not relevant because the scientists don’t care whether the atom is located on Mars. The difference between these two approaches seems insignificant.

Ninth, when nonsubstantive disputes have multiple expressions that fail to carve at the joints, sometimes the source of the nonsubstantivity can be localized to a proper subset of those expressions. For example, suppose ‘lithium’ has two candidates, lithium<sub>1</sub> and lithium<sub>2</sub>. (Lithium is like jade.) And suppose that Mars is shot through with lithium<sub>1</sub>, but has no lithium<sub>2</sub>. Then the question of whether there is lithium on Mars is again nonsubstantive. But even though both of the terms involved, ‘lithium’ and ‘Mars’, have multiple candidates, we can place the blame for the nonsubstantivity solely on ‘lithium’. For under any candidate for ‘Mars’, the question’s answer remains sensitive to the choice of a candidate for ‘lithium’, whereas it’s not the case that for every ‘lithium’-candidate, the question’s answer is sensitive to the choice of a ‘Mars’-candidate (indeed: for *every* ‘lithium’-candidate, the answer is insensitive to the choice of a ‘Mars’-candidate).

Tenth, a question can be substantive “along some dimensions” but not along others. Contrast the following two predicates:

$$Fx =_{\text{df}} x \text{ is more massive than all bachelors}$$

$$Gx =_{\text{df}} x \text{ is wittier than all bachelors}$$

(‘=<sub>df</sub>’ means “means by definition that”). There is, I will say, just one dimension along which *F* has multiple candidates and generates nonsubstantive questions, whereas there are two such dimensions for *G*. *F* and *G* are alike in having multiple candidates and generating nonsubstantive questions. But in the case of *F*, these facts come from a single source: the fact that ‘bachelor’ has multiple candidates. (Let’s assume for the sake of argument that ‘all’ and ‘more massive than’ carve at the joints.) *F*’s candidates are generated by those of ‘bachelor’; and there’s just one way for it to be nonsubstantive whether a given thing is *F*: namely, for there to be some candidates bachelor<sub>1</sub> and bachelor<sub>2</sub> of ‘bachelor’ such that the thing is more massive than all bachelors<sub>1</sub>, but not more massive than all bachelors<sub>2</sub>. Along the dimension of ‘more massive than’ (and ‘all’ as well), we may say, the question of whether a given thing is *F* is substantive. In contrast, *G*’s candidates are generated both by candidates of ‘bachelor’ and

by candidates of ‘wittier than’. There are, correspondingly, two ways for the question of whether a given object is  $G$  to be nonsubstantive—nonsubstantivity “along two dimensions”.

Given this terminology I can clarify some cryptic remarks at the end of section 3.3. In that section I argued that even a subjectivist about epistemic value might regard the distinction between reasonable and unreasonable prior probability functions as being objective. According to this position, I said, judgments about epistemic value are subjective “along one dimension” but not along another. What I meant can now be stated more fully. Terms of epistemic evaluation do not carve at the joints. However, some of their “components” *do* carve at the joints—just as some of  $F$ ’s components (‘more massive than’, ‘all’) carve at the joints even though  $F$  itself does not carve at the joints. If, for example, the terms of epistemic evaluation are defined in part by “simplicity” constraints on prior probability distributions, then the joint-carving components in question would be those that state the simplicity constraint in joint-carving terms. Any nonsubstantivity in the question of whether a prior probability distribution is reasonable would *not* be due to candidates for these components (since they carve at the joints—reasonably well, anyway), but would rather be due to candidates in the rest of the components.

Eleventh, suppose some expression  $E$  has a candidate  $c$  that carves at the joints (in the right way) much better than all its other candidates, but that the actual meaning of  $E$  is not  $c$ , but rather some other candidate  $c_{@}$  that carves much worse at the joints. (This could happen if reference magnetism is false, or if for some other metasemantic reason the joint-carvingness of  $c$  is outweighed by other factors.) Now, suppose that the truth-value of some sentence  $S$  is sensitive to the fact that  $E$  means  $c_{@}$ , rather than  $c$ . My account treats  $S$  as substantive (assuming that none of its other expressions have appropriately varying candidates), since one of  $E$ ’s candidates (namely,  $c$ ) carves at the joints much better than the others. Nevertheless, there is an intuitive sense in which  $S$  is nonsubstantive since, intuitively, its actual truth-value is a mere reflection of a linguistic choice, not the world’s structure.  $S$ , we might say, is “actual-verdict nonsubstantive” because it could have had a different truth-value *without carving worse than it actually does*—because, that is,  $S$  has a different truth-value under some candidate for one of its expressions that carves at the joints no worse than that expression’s actual meaning. Something substantive is at stake here, but the actual verdict is nonsubstantive.

### 4.3 Conventionality

The next two sections discuss two species of nonsubstantivity. The first is a sort of conventionality. A sentence exhibits this sort of conventionality when it involves, in a sense to be explained, an arbitrary conceptual choice.<sup>8</sup>

The word ‘convention’ generally signifies an arbitrary choice amongst equally good ways to achieve a certain goal by collective action. The USA had a goal of safely organizing its motorways; that goal could have been achieved either by everyone driving on the right-hand side or by everyone driving on the left; the convention to drive on the right was a more or less arbitrary choice between these two alternative solutions.

Turning to language, there is a sort of conventionality that is both familiar and banal: the choice of symbol or sound to represent a given content. Call this: symbol-conventionality. A second sort of conventionality—call it content-conventionality—is exhibited by sentences that are *about* conventions. The notion of aboutness is admittedly slippery, but obvious examples of content-conventionality include ‘there are some conventions’, “Snow” refers to snow in the dominant language of North America in 2011, and so on.

The sort of conventionality I have in mind is different. Sometimes we have a certain semantic goal; we need to introduce a word in order to accomplish that goal; and there are a number of different candidate meanings, each such that the goal would be accomplished equally well if that candidate were chosen as the meaning of the word. The choice of one of these candidate meanings to be the meaning of the word exhibits what I’ll call *candidate-selection* conventionality (sometimes just “conventionality”, when there’s no danger of confusion).<sup>9</sup>

To illustrate, consider the word ‘inch’. The purpose of ‘inch’ is to be a convenient measure for smallish things, the kinds of things we can hold in our hands. But there is a range of very similar lengths that would each have served this purpose. We chose one of these to mean by ‘inch’, but that choice was arbitrary; any of the others would have served our purposes equally well. This choice was one of candidate-selection convention.

By saying that each length in the range would have served our purposes equally well, I have in mind two things. First, the lengths in the range carve at the joints equally well.<sup>10</sup> And second, adopting any of the alternate lengths

<sup>8</sup>Sidelle (2009) discusses a similar sort of conventionality.

<sup>9</sup>Skow (2010, section 4) gives a similar account.

<sup>10</sup>Or if any carves better than the others, it is of the wrong sort for ‘inch’.

would have “achieved the same semantic goal”. All length-words achieve a general semantic goal of allowing speech of absolute and relative sizes, but ‘inch’ has a more specific goal: to be a convenient measure of smallish things. This goal could have been achieved by many lengths within a certain range. But if ‘inch’ had meant *mile*, it would not have achieved exactly this goal, since measuring smallish lengths in miles would be inconvenient. And if ‘inch’ had meant something other than a length—for instance, if it had meant happiness—then it would not have achieved anything like its actual semantic goal.

All words for units of measure are conventional in this way. The boundaries of countries provide a further example. When countries are formed or resized, an arbitrary choice is sometimes made about where the new border will go. The corresponding choice of semantic values for words about the country is candidate-selection conventional.

Call a sentence candidate-selection conventional when its truth-value turns on a candidate-selection conventional choice. Sentence (C) is an example:

(C) My computer screen measures exactly 15 inches.

(C) is true, but would have been false if ‘inch’ had meant a slightly different candidate length. This is not content-conventionality: (C) is not about conventions in any interesting sense. Nor is it mere symbol-conventionality. Of course, (C) *is* symbol-conventional; all sentences are. But not all sentences contain a word for which we could have chosen an alternate meaning that would have equally well suited our purposes for that word, and which would have given the sentence a different truth-value.

It might seem odd to call (C) conventional (except, of course, in the sense of symbol-conventionality). Facts about measurable quantities like length, after all, are as objective as can be. But remember that my account of substantivity, and hence of candidate-selection conventionality, is metalinguistic. It is the sentence, not the fact, that is conventional. Moreover, recall the interest-relativity of substantivity (the eighth refinement of section 4.2). It seems odd to call (C) conventional because in typical contexts where (C) is disputed, the focus of the disputants is not on which candidate is meant by ‘inch’, but rather on the length of the computer screen. However, consider a context where the disputants *do* focus on which length counts as being “one inch” (and let the length of the computer screen be common ground). Then the dispute does seem nonsubstantive, and it seems natural to call the sentence conventional.<sup>11</sup>

---

<sup>11</sup>As with substantivity in 4.2, I’m neutral on whether to relativize conventionality to context,

And if we consider sentences about measurable quantities where the second sort of context is more common, the label ‘conventional’ no longer seems odd at all. A dispute between an American and someone from England over whether a certain container of milk measures “one gallon”, where the container’s volume is not under dispute, is quite naturally thought of as nonsubstantive; and it’s natural to call sentences in the dispute as being conventional, since their truth turns on the conventional decision of whether to adopt the U.S. or imperial standard for the gallon.

(Even paradigm nonsubstantive sentences exhibit the sort of relativity just discussed. ‘The pope is a bachelor’ could be the subject of a substantive-seeming dispute if, say, the disputants knew little about Catholicism and were in effect arguing about whether the Pope is married.)

Our definitions ensure that conventionality (of the candidate-selection variety) implies nonsubstantiality; but the converse does not hold. Both involve a sentence whose truth-value depends on which of several equally joint-carving candidates is assigned to one of its terms. But for conventionality, there must be a selected candidate (or a vague selected range, in the case of vague conventionality), and that selection must be made by arbitrary choice. If no selection is made (whether because of vagueness or some other form of semantic indeterminacy), or a selection is made nonarbitrarily (see section 4.4), there is no conventionality. Also, ‘conventional’ seems most apt when the arbitrary choice is made more or less consciously, when alternative choices stare us in the face, and when those choices accomplish *exactly* the same semantic goal; it seems less apt when the choice has been made implicitly and collectively, over time, when no one thinks much about the alternatives, and when the alternatives accomplish slightly different semantic goals. Supposing the question of whether Crusoe is a bachelor to have an answer, perhaps we should not call it “conventional”.

Everyone agrees that the boundaries of countries and units of measure are in some sense conventional. But claims of “conventionality” have been made in more controversial domains: for physical geometry (recall section 3.4), morality, and so on.<sup>12</sup> Candidate-selection conventionality is a useful tool

---

or to say instead that although (C) is conventional in every context, for some length,  $l$ , what the disputants really care about in the first context is not (C), but rather the nonconventional sentence ‘the screen has length  $l'$ . (The variable ‘ $l'$ , under a given assignment, is intended to lack alternate candidates.)

<sup>12</sup>The geometric conventionalism of philosophers like Reichenbach should not be confused with conventionalism about units of measure. The former concerns even unitless length

for articulating such doctrines.<sup>13</sup> For some of these conventionalists do not want to claim that sentences about the target domain are *about* conventions, nor do they wish to merely make the trivial claim that the sentences exhibit symbol-conventionality.

Consider, for example, the view that moral sentences are candidate-selection conventional. According to this view, society had an interest in introducing moral vocabulary, attached to some norms or other; but within certain limits, various norms would have equally well served the purpose. Thus, there were various candidate meanings available for normative words, each of which would have achieved those words' semantic goal. Moreover, none of these candidate meanings carves at the joints better than the others. One of these was selected, more or less arbitrarily—a candidate-selection convention. Now, this view is not a mere claim of symbol-conventionality. By claiming that the meanings we have actually selected for moral language are on a par with alternate meanings—both metaphysically and concerning the satisfaction of the goal of morality—this view really does downgrade morality (a part of it, anyway) in a way that a mere claim of symbol-conventionality would not. But nor does this view imply that moral sentences are about conventions. Thus it allows that (for example) murder would have been wrong even if we had chosen different norms (since the conventionality of the choice of norms is not built into the propositions we express using moral words).

## 4.4 Subjectivity

The second species of nonsubstantivity may be called subjectivity. Like conventionality, it occurs when a linguistic community chooses one of several candidates. But in the case of subjectivity, the choice is not arbitrary; rather, it reflects something important about the linguistic community.

---

predicates such as Tarski's 'congruent'.

<sup>13</sup>Not in all cases, however. Some such doctrines are best understood as claims of content-conventionality. Conventionalism (or social constructionism) about works of art, artifacts, or gender and race might be construed as the view that sentences about these subject matters express propositions that are in some sense about social conventions. Were our conventions different or nonexistent, there would exist different (or even no) sculptures, tables and chairs, men and women, and so on, since what it is to be a work of art, etc., involves social conventions. (Compare Haslanger's (1995, p. 98) notion of constitutive construction.) Other such doctrines are not clearly instances of any of my three sorts—the logical conventionalism discussed in section 6.5, for instance.

As with conventionality, this sense of subjectivity must be distinguished from others. Consider the following toy semantic theories of aesthetic sentences:

**Expressivism** By uttering ‘ $x$  is beautiful’, a speaker communicates no proposition, but rather gives expression to a certain positive aesthetic attitude,  $A$ , to  $x$ .

**Indexicalism** By uttering ‘ $x$  is beautiful’, a speaker,  $S$ , communicates the proposition that  $S$  bears attitude  $A$  to  $x$ .

Aesthetic sentences express attitudes, given the first semantics, and communicate propositions about attitudes, given the second. Either way, there is a straightforward kind of subjectivity. But the kind of subjectivity I have in mind is different. It is brought out by a third semantics:<sup>14</sup>

**Projectivism** By uttering ‘ $x$  is beautiful’, a speaker,  $S$ , communicates the proposition that  $x$  is  $P$ , where the property *being P* is a certain physical property that is the linguistic meaning of the predicate ‘is beautiful’ in  $S$ ’s language; *being P* is the linguistic meaning of ‘is beautiful’ because members of  $S$ ’s linguistic community bear attitude  $A$  to  $P$ s.

Under this semantics, aesthetic sentences are not subjective in the straightforward sense. Accordingly, they pass common tests for “objectivity”. For example, sentences about beauty make mind-independent claims: beautiful mountains would still have been beautiful even if no humans had ever existed. Nevertheless, there remains a clear sense in which the aesthetic is subjective. For *which* physical properties aesthetic predicates stand for is determined by the attitudes of the speaker’s community, and any attitudes are as good as any other.<sup>15</sup> Even though the properties ascribed by aesthetic predicates are wholly mind-independent, response-independent, and so on, the *selection* of *these* features as the features to be expressed by aesthetic predicates is accomplished solely by our having the attitudes that we do. Imagine a range of linguistic communities, each with different aesthetic attitudes. The predicate ‘is beautiful’ expresses different—equally joint-carving—properties in these different lin-

---

<sup>14</sup>I am using ‘projection’ nonstandardly; it usually signifies the mistaken attribution of mental features to the external world.

<sup>15</sup>Compare Street (2006, section 7).

guistic communities.<sup>16</sup> Where we call a mountain beautiful, speakers of another language withhold that predicate, and instead call the mountain ‘ugly’. In such cases, everyone speaks truly; no one is making a mistake. The mountain has the property  $B_E$  that is expressed in English by ‘beautiful’, and lacks the property  $B_O$  that is expressed in the other language by ‘beautiful’.<sup>17</sup> The appropriateness of the language of subjectivity in this case is manifest.

Shine a light down on a piece of paper suspended over a table. If the paper has a geometric shape cut from it—a circle, say—the light will shine through the hole and project that shape onto the table below. Let  $X$  be the illuminated portion of the table.  $X$  is circular. Moreover,  $X$ —that portion of the table—would still have been circular even if a square rather than a circle had been cut from the paper (though  $X$  would not then have been exactly illuminated). But there is nothing special about  $X$ . The square region that would have been illuminated, had the cut-out hole been square, is just as good a region. The illuminated region is a projection of the hole. An observer of the illumination will learn something more important about the hole than about the table, even though circularity is an intrinsic feature of the illuminated part of the table. Aesthetic features are like the shape of the illuminated region  $X$ , the selection of meanings for aesthetic predicates like the selection of the shape of the hole in the paper. If projectivism is true, then aesthetic features are not “about” us, and would persist even if there were no humans. But assuming there are no aesthetic joints in nature, nothing beyond our aesthetic attitudes constrains the aesthetic categories picked out by aesthetic language; an observer who watches us label things ‘beautiful’ and ‘ugly’ will learn as much about our attitudes as they will about the things, even if these predicates ascribe intrinsic properties of the things. Aesthetic predicates express the properties they do because of our attitudes; aesthetic features are projections of our attitudes.

I do not claim that projectivism is true; it is, I suspect, an overly simplistic model of the semantics and metasemantics of ‘beautiful’. Its point, rather, is to establish a general fact about the nature of subjectivity: there is a kind of subjectivity that results, not from statements in the target discourse being about

---

<sup>16</sup>If predicates are individuated by the properties they express, then we cannot say that each community uses the same predicate ‘beautiful’; what we must say instead is that each uses a predicate that plays the role that ‘beautiful’ plays in our language. Allow me the liberty of speaking of each community as having “aesthetic” predicates, “aesthetic” attitudes, and so on.

<sup>17</sup>This variation in the meaning of ‘beautiful’ is not just the mundane sort of variation that is possible because of symbol-conventionality. In each language, ‘beautiful’ plays the same role; it is a word for the things picked out by the aesthetic attitudes of that language’s speakers.

our values, but rather from our values selecting one from a range of equally good meanings. The importance of this fact for metaethics should be clear. To earn titles of objectivity and realism (together with the associated imagery of externality, discovery, and so on), it is not enough that evaluative language be assigned “objective content”. For projectivism assigns *physical* contents, which are as objective as can be; yet it merits neither the names nor the imagery. The objective content must also stand out from its metasemantic surroundings. It must enjoy some sort of privilege over alternate candidate contents.<sup>18</sup>

A sentence is subjective, then, in the sense illustrated by the projectivist semantics, if and only if its truth-value depends on which of a range of equally joint-carving candidates is meant by some term in the sentence, where the candidate that we in fact mean was selected in a way that is not arbitrary, but rather, reflects something important about us, such as our values.<sup>19</sup> As with conventionality, it is appropriate to speak of this sort of subjectivity only if no one candidate carves at the joints (in the right way) better than the rest. If there *were* aesthetic joints in reality—vindicating the very strongest form of aesthetic realism—then one of the communities from our earlier example might match those joints with their usage of ‘beautiful’. This lucky community would then be *uniquely right* about aesthetics, and talk of “subjectivity” would seem out of place. They might be uniquely right in the straightforward sense of being the only community that speaks truly, if ‘beauty’ in every community has the same, joint-carving sense (this might happen because of reference magnetism). But even if the communities mean different things by ‘beautiful’, so that every community speaks truly, the lucky community remains uniquely right in the sense that only their term ‘beautiful’ gets at the distinguished aesthetic structure of the world. The other languages are metaphysically second-rate. In such a circumstance, the language of subjectivity again seems out of place, even though

---

<sup>18</sup>Moral realists sometimes implicitly support such a privilege, even if their explicit focus is on the contents of moral sentences (see, for instance, Boyd (1988) on homeostatic property-clusters and Railton (1986) on feedback). Note also that the appropriateness of the language and imagery of realism and objectivity comes in degrees. Even if contents for moral language are selected by facts about us, morality seems more realistic and objective if those facts are counterfactually robust and universal across different societies—if they reflect the human condition rather than historical accident.

<sup>19</sup>What if our actual values are so hard-wired into our brains that it would be difficult or even impossible to adopt others? I still want to count other values as picking out “candidates” for the term in question; those candidates are not “semantically alien” even if we could not adopt them. Thus it’s not always right to say (as I have been) that “we could just as easily” have adopted alternate candidate meanings.

the unlucky communities all speak truly given what they mean by ‘beautiful’.

The kind of subjectivity we have been discussing results from the projection of our values. But perhaps ‘subjective’ is also appropriate when we project important features of ourselves other than values. Suppose causation does not carve at the joints, and that there are a number of candidate semantic values for ‘cause’, none of which is metaphysically distinguished. Suppose that one of these is in fact the actual semantic value, for some reason that reflects something important about us. Suppose, for instance, that it’s essential to the role that ‘cause’ plays in our conceptual scheme that it have a counterfactual analysis. Like ‘beautiful’, the semantics of ‘cause’ reflects an important feature of ourselves, not an arbitrary semantic decision, but unlike ‘beautiful’, the feature has nothing to do with value.

Whether or not ‘subjective’ is appropriate here, this sort of projection is opposed to “objectivity”. If, contrary to what I think, there is a joint-carving relation of causation, then it would be natural to describe this by saying “there are objective facts about causation”, or “one description of the facts of causation is objectively correct”. Objectivity is opposed to subjectivity, conventionality, or any other sort of nonsubstantivity.

Philosophers often make inchoate claims that are best understood as concerning this notion of objectivity. To take one example, consider the criticism made by Armstrong (1983, pp. 40–59) and others that Lewisian laws of nature are not genuinely necessary and cannot explain regularities. Everyone agrees, after all, that mere regularities do not explain regularities, and need not be necessary; but Lewisian laws are just glorified regularities, regularities that are integrated into the simplest and strongest system. Lewis’s reply always seemed elusive and unconvincing:

Some familiar complaints seem to me question-begging ... If you’re prepared to agree that theorems of the best system are rightly called laws, presumably you’ll also want to say that they underlie causal explanations; that they support counterfactuals; that they are not mere coincidences; that they and their consequences are in some good sense necessary; and that they may be confirmed by their instances. If not, not. It’s a standoff—spoils to the victor. (1994, pp. 478–9)

Lewis’s response is in essence to define ‘nomic necessity’, ‘explanation’, and so on, in terms of ‘law’, and then to claim that if you don’t think that necessity and explanation, thus defined, are *genuine*, that must be because you are just resisting his analysis of lawhood. What is really going on, I think, is that Armstrong is

bothered by the fact that on Lewis's account, laws (and so everything defined in terms of them) are not objective. For Lewis there is no distinguished structure in the vicinity of laws of nature. (This is the whole point of Lewis's Humeanism.<sup>20</sup>) This rankles some; they believe that laws "run deep"; they cut at the joints; they are objective. Likewise for necessity, explanation, and the rest: Lewis provides analyses of these concepts that to some degree fit our ordinary concepts, but are not particularly objective. This is what really bothers Armstrong (and Lewis's reply is not responsive).

## 4.5 Epistemic value

I have been arguing for connections between joint-carving and a range of concepts. Many of these connections can be unified by a single thesis about epistemic value: it's *better* to think and speak in joint-carving terms. We ought not to speak the 'grue' language, nor think the thoughts expressed by its simple sentences.

The goal of inquiry is not merely to believe truly (or to know). Achieving the goal of inquiry requires that one's belief state reflect the world, which in addition to lack of error requires one to think of the world *in its terms*, to carve the world at its joints. Wielders of non-joint-carving concepts are worse inquirers.

Recall the community that divides the red–blue world along the diagonal plane (section 1.1). They are *missing out*; they do worse than we do as inquirers. Of course, if they explicitly form false beliefs *about* structure, then there is a perfectly obvious sense in which they are doing worse. But even if neither they nor we form such beliefs, they still do worse, simply by thinking of the world in the wrong terms.

Joint-carving thought does not have merely instrumental value. It is rather a constitutive aim of the practice of forming beliefs, as constitutive as the more commonly recognized aim of truth. Nor is joint-carving a conscious goal, at least not of most inquirers. It is rather a standard by which beliefs and believers may be evaluated, whether or not it is consciously acknowledged.

---

<sup>20</sup>See Lewis (1986c, introduction), Lewis (1994). In this respect, Lewis is just like the traditional regularity theorist. Each agrees that the world is *fundamentally* anomic; the difference is that Lewis does a better job of approximating in extension the ordinary notion of a law of nature. Matters are very different for Armstrong (1983), whose relation of nomic necessitation is a universal, and so is part of reality's fundamental structure.

(Strong and weak versions of this thesis about epistemic value can be distinguished.<sup>21</sup> The strong version is what I have been advocating: it's worse to employ non-joint-carving concepts. The weak version says merely that it's worse to believe non-joint-carving *propositions*, where a proposition is joint-carving to the extent that it can be simply expressed using joint-carving concepts (given some appropriate notions of simplicity and proposition). (The strong version implies the weak, more or less, given the perhaps plausible additional assumption that anyone employing joint-carving concepts ought to believe propositions simply expressible in terms of them.) The weak version allows one to employ ‘grue’ and ‘bleen’, provided one does not believe propositions like *all emeralds in this room are grue*, but rather believes instead propositions like *all emeralds in this room are either grue and first observed before 3000 A.D. or bleen and not first observed before 3000 A.D.*—the idea is that this latter proposition is relatively joint-carving because it can be expressed by “all emeralds in this room are green”. Both versions are correct, I think, and so I stick with the strong. But one might reject the strong version while accepting the weak; and much of what I go on to say could be reformulated to appeal only to the weak.)

In his book *Dividing Reality*, Eli Hirsch argues effectively that it is intuitively compelling that we ought to speak and think in joint-carving terms; but he also argues effectively that this position is difficult to support. Although the epistemic value of joint-carving inquiry is, I think, a basic one and hence not derivable from other values, I do think it can be supported in several ways. First, the aim of joint-carving can be seen as having the same source as the aim of truth: beliefs aim to *conform to the world*. Here is a simplified but intuitive picture. The realist about structure thinks of the world as coming “ready-made” with distinguished carvings. By analogy with the notion of a mathematical structure, think of The World as a structure: a set  $E$  of entities together with a set  $R$  of relations over  $E$  (think of the relations here simply as ‘tuples of members of  $E$ ’). Now, ignoring partial belief, it is natural to think of the beliefs of a subject,  $S$ , as consisting of the representation of a structure: the subject represents there being objects,  $E_S$ , together with a set  $R_S$  of relations over  $E$ . Given this picture, it is utterly natural to think of full conformity to The World as requiring  $\langle E_S, R_S \rangle$  to be identical to  $\langle E, R \rangle$ . Conformity requires the believer to represent the structured world exactly as it in fact is, and thus requires the represented relations  $R_S$  to be identical to the world's structuring relations  $R$ . Thus if belief aims to conform to the world, and if belief and the world are both

---

<sup>21</sup>Thanks to Steve Steward.

structured, belief aims not just at truth, but also at the right structure—truth in joint-carving terms.

Second, we think of scientific discovery as satisfying the aims of inquiry particularly well; why? Answer: it is because scientific discoveries are phrased in particularly joint-carving terms. Relatedly, we think of truths that are stated in extremely non-joint-carving terms—for example, the scrambled propositions of section 2.6—as being comparatively worthless. Relatedly, imagine (or recall) first coming to believe that morality, beauty, justice, knowledge, or existence is a mere projection of our conceptual scheme—that the truth in these domains is conventional, subjective, or otherwise nonsubstantive. Why does that feel so deflating; why does it diminish the urgency of finding the truth; and why does it diminish the value of the truth once found? Answer: though we might not put it exactly thus, our original picture in these lofty domains is that of joint-carving. Morality, beauty, and the rest are built into the nature of things, we naïvely think, rather than being mere projections. Giving up on objectivity means giving up on joint-carving, and hence diminishes the value of truth.

Third, consider the following series of scenarios. Scenario 1: the physical world is pretty much the way we think it is; it includes physical objects in addition to spacetime. Scenario 2: the physical world consists of nothing more than propertied points and regions of spacetime. Scenario 3: the physical world consists of a wave function in a massively dimensional configuration space. Neither three-dimensional space nor four-dimensional spacetime exist, fundamentally speaking. Scenario 4: our ordinary beliefs are caused by *The Matrix*, a computer simulation that directly stimulates our brains while our bodies lie in stasis.<sup>22</sup> Scenario 5: I am a disembodied brain floating in an utterly empty space; the changes in my brain that give rise to my “mental states” happen purely by chance. In Scenario 1 my ordinary beliefs about myself and my surroundings are true. The same holds in Scenario 2, I think, though some would disagree. This is less clear in Scenario 3, since the structures in the world that ordinary beliefs would need to pick out in order to be true—patterns in the wave function—are so far from the joints in reality. Matters are worse in Scenario 4: our ordinary beliefs would be true only if interpreted as picking out aspects of the computer program, which are (we may stipulate) quite distant from reality’s joints in the world that houses the program. And once we get to Scenario 5, if I can be said to have any mental states at all, nearly all of them are clearly untrue. They would be true only if they had contents defined

---

<sup>22</sup>Compare Chalmers (2005).

on the world's empty space (or parts of my brain); but no assignment of such content could be regarded as being more correct than other assignments on which the mental states would come out false. Now, in this series of scenarios, the match between our beliefs and reality's joint-carving structure is gradually eroded. The erosion is severe enough to disrupt truth only late in the series. But even earlier in the series, at stage 3, say, much of what we care about has been lost, even if what we normally say is still true. We are a partial intellectual failure if we live in configuration space or *The Matrix*, even if we believe truly. Moreover, even if the transition from truth to falsity in our ordinary beliefs is abrupt, what we care about in inquiry seems to be more continuously eroded in the series. These facts suggest that what we care about is truth in joint-carving terms, not just truth.

The thesis about the value of joint-carving inquiry unifies many of the connections I have forged between joint-carving and other concepts. The unification takes the form of a recurring theme, rather than a derivation from first principles. Here are a few sweeping lines of thought.

Duplication: Lewis defined duplicates as objects whose parts have the same natural properties and stand in the same natural relations. But why *care* about duplication? A partial answer emerges from this section's thesis: if it's good to classify objectively alike things together, then duplication should be of interest to us. (This is only a partial answer, because it does not address why duplicates in particular, rather than merely objects that share natural properties, would be of interest. In particular, why are we interested in objects whose *parts* have the same natural properties and stand in the same natural relations? The answer here, I think, somehow concerns the distinctive status of parthood in our thought.<sup>23)</sup>)

Explanation: here we have a direct connection to the thesis of this section. Explanation is an epistemic achievement. There are many views about explanation, and no doubt many species of explanation; but on all views, explanation is an attempt to improve our epistemic position in some way, to make the world more intelligible. But the epistemic achievement will be diminished if cast in non-joint-carving terms, given this section's thesis.

Laws: given the Lewisian theory of laws, on which laws must be phrased in joint-carving terms, it follows that laws are propositions in which we ought to take an interest. Concern with what science says is not epistemically optional.

Reference: Lewis himself defended reference magnetism by saying that it is

---

<sup>23</sup>See Sider (2007b).

constitutive of reference that reference goes to the most natural candidate. But this just invites the question of why we should care about reference. What's better about reference than "reference\*", which does not always go to the most natural candidate? Answer: reference carves better at the joints than reference\*; given the thesis of this section, that is why we should care about it.<sup>24</sup>

Induction: in Bayesian terms, the problem of induction is to characterize the correct prior probability distributions. Now, here is a way of thinking about this notion of correctness. In worlds like ours, human beings have survived, and thrived, in part because of how they formed beliefs about their environment. Taking for granted that they updated by conditionalization (more or less), they survived because they had appropriate priors. So an explanation of why humans survived will cite certain general features of human prior probability functions. These general features cannot be too specific if they are to explain why humans in general (as opposed to on this continent or that) survived; but they cannot be too general either if they are to robustly explain the survival. They will surely include such features as this: humans tend to project reasonably joint-carving properties (like blue) more than badly non-joint-carving properties (like grue). Note further that having these features explains why humans survive *in worlds like ours*. If our world contained few regularities involving joint-carving properties, then projecting such properties wouldn't lead to survival. So: there are general features of human priors that explain why we survived. These features, according to this line of thought, just are the features that make priors correct. Correctness just is: having the features that explain humans' survival. This is just one crude story, at best incomplete and at worst on the wrong track; but my point could be made under refined or alternate stories. My point is this: epistemic notions—such as the notion of a correct prior probability distribution—have their source in what explains certain facets of our epistemic success. Thus those notions must respect nature's joints in certain ways, since explanations quite generally must be cast in reasonably joint-carving terms. Given the crude story imagined above, for example, the features of our prior probability distributions cited by the explanation of our survival would need to be cast in reasonably joint-carving terms. (Similar points could be made under alternate stories in which the explanation of our epistemic success cites features of the present, say, rather than our evolutionary past.)

Space and time: the intrinsic structure of space and time is given by the spatiotemporal notions that carve at the joints. But why is uncovering this

---

<sup>24</sup>Thanks to Robbie Williams.

structure the goal of science? Because science—like all inquiry—aims to correctly represent the world, and this requires representing the world “in its own terms”; it requires carving at the joints.

Objectivity: why do we prize substantivity and objectivity, and downgrade the subjective, conventional, nonsubstantive? Because it’s better to think in joint-carving terms; and the more one does, the more one’s meanings “stand out from the metasemantic background”, yielding substantive, objective claims.

## 4.6 Objectivity of structure

We have uncovered a web of connections between structure and various notions. This web of connections yields the primary argument against Goodman’s claim that structure is merely the projection of our interests or biology: subjectivism about structure leads to subjectivism about the other notions in the web. We could not formulate an appropriately objective form of Bayesianism. We could not rebut conventionalists about physical geometry. We could not believe in objective semantic determinacy.

But the most significant fallout from Goodmania, to my mind, arises from structure’s connections with epistemic value and with objectivity.

Epistemic value: joint-carving languages and beliefs are better. If structure is subjective, so is this betterness. This would be a disaster. Recall section 2.6. If there is no sense in which the physical truths are objectively better than the scrambled truths, beyond the fact that they are propositions that we have happened to have expressed, then the postmodernist forces of darkness have won.

Objectivity: whether questions are substantive, nonconventional, objective, and so on, depends on whether they are phrased in terms that carve at the joints. Given subjectivism about structure, we would have subjectivism about substantivity, depth, conventionality, objectivity. No discourses would be objectively objective. Another disaster.

The knee-jerk realist thinks that the world is “out there”, waiting to be discovered rather than constructed—all that good stuff. Everyone agrees that this picture rules out views according to which all truth is mind-dependent in the crudest ways, but it requires more. After all, under the projectivist semantics (section 4.4) one can truly say that “the mountain would still have been beautiful even if humans had never existed”. The realist picture requires the “ready-made world” that Goodman (1978) ridiculed; it requires the world to *really* be as

physics says; it requires objectivity; it requires objective distinguished structure. To give up on structure's objectivity would be to concede far too much to those who view inquiry as being merely the investigation of our own minds.

# Chapter 5

## Metametaphysics

METAMETAPHYSICS is inquiry into the status of metaphysics. It is of the nature of the beast that one is led to ask: are metaphysical disputes substantive? Are they objective, genuine, deep? Or are they nonsubstantive in some way: conventional, subjective, merely verbal or conceptual? Must there be a fact of the matter about who is right?

The answers to these questions depend on which metaphysical dispute is at issue. The crucial factor is whether the dispute is phrased in terms that carve at the joints. This connection to metametaphysics is the final part of the role I envisage for the notion of structure.

### 5.1 The challenge of metaphysical deflationism

Metaphysics has always had critics. The most extreme base their critique on sweeping views about language (logical positivism, ordinary language philosophy) or knowledge (naïve empiricism). But such views are in trouble on independent grounds. Their oversimplified conceptions of how we make semantic and cognitive contact with the world notoriously threaten the science of unobservables as much as they threaten metaphysics.

A more formidable challenge comes from a more modest critic, whom I'll call a (metaphysical) *deflationist*.<sup>1</sup> A deflationist about question *Q* says that *Q* is in some sense merely verbal or conceptual. Its answer (if it has an answer)

---

<sup>1</sup>Not to be confused with a deflationist about truth such as Paul Horwich (1990). Although my metaphysical deflationist is not any specific person, Eli Hirsch's work on metaontology—certain facets of it anyway—may be kept in mind as a model.

is determined by linguistic or conceptual rules. What is at issue is not “the world”, but rather *us*—how we think and talk. (We will need to refine this rather vague characterization, but it will do for now.) This critique, moreover, is based, not in a sweeping conception of language or epistemology, but rather in considerations specific to *Q*.

My deflationist has a favorite “go-to move”. The move is schematic, and can be made in service of various antimetaphysical conclusions, but in each case its form is the same:

1. The deflationist observes a certain metaphysical dispute, in which one of the contested views is expressed by a certain sentence *S*.
2. He argues that there is an interpretation of the language of *S*—a way of assigning meanings to the sentences of that language—under which *everyone* can agree that *S* is true.
3. And he argues for a certain *parity* between this and rival interpretations.

The go-to move may be summarized thus: “There’s a perfectly good way to talk under which *S* is clearly true.”

Before discussing the meaning of ‘parity’ in step 3 and how the move might lead to antimetaphysical conclusions, let me give some examples. Confronted with a dispute over whether the relation of part to whole is reflexive, the deflationist might point out that the following defines a reflexive relation:

$$x \text{ is part}^* \text{ of } y =_{\text{df}} x \text{ is part of } y \text{ or } x = y.$$

Even if ‘part’ in English is not reflexive, ‘part\*’ is. So under an interpretation of English words that is just like actual English except that ‘part’ means part\*, the sentence ‘Everything is part of itself’ obviously comes out true. Moreover, this interpretation is very similar to English since ‘part’ and ‘part\*’ have exactly the same extension except perhaps for pairs of the form  $\langle x, x \rangle$ —parity.

Second example: with respect to the same dispute, the deflationist might construct a related interpretation, also very similar to English, in which ‘Nothing is part of itself’ comes out true, by letting ‘part’ mean part\*\*:

$$x \text{ is part}^{**} \text{ of } y =_{\text{df}} x \text{ is part of } y \text{ and } x \neq y.$$

Third example: confronted with a dispute over whether holes exist, a deflationist might construct an interpretation of quantificational sentences in which

‘There is a hole in object  $x$ ’ means that  $x$  is perforated. Even an opponent of holes will agree that some such sentences are true in this interpretation, since the opponent agrees that some things are perforated.

The deflationist can use the go-to move within various dialectical strategies. Many of the strategies appeal to the fact that meaning is largely determined by use. For example, a large part of why ‘bachelor’ means (something like) *unmarried adult male* is that we tend to call something a ‘bachelor’ if and only if it is an unmarried adult male. This is oversimplified; but rather than trying to refine it, let’s just employ the code: “Our use of ‘bachelor’ favors the hypothesis that it means unmarried adult male.” Here, then, are some of the strategies:

**Common-sense strategy** Suppose a metaphysician argues for a noncommonsensical position. Argue first, using the move, that sentences expressing a more commonsensical rival position can be truly interpreted. Argue, second, that our use of the crucial terms favors the commonsensical interpretation over any interpretation on which the noncommonsensical position comes out true. Conclude that the noncommonsensical position is not true. (For example, our use of sentences like ‘There is a hole in that piece of cheese’ clearly favors interpretations under which the sentence comes out true whenever the piece of cheese in question is perforated. So, nominalism about holes is false—and this conclusion was secured simply by reflecting on language.)

**Indeterminacy strategy** Argue first, using the move multiple times, that each of a range of views about some metaphysical question comes out true under some interpretation. Argue, second, that our use of the crucial terms does not favor any of these interpretations over the others. Conclude that it is indeterminate which view is true. (For example, our use of ‘part’ favors neither part\* nor part\*\* over the other; so it’s simply indeterminate—and therefore pointless to debate—whether parthood is reflexive.)

**Deflationary strategy** As with the indeterminacy strategy, argue first that each of a range of views about some metaphysical question comes out true under some interpretation. But now, rather than taking a stand on whether actual usage favors one interpretation, simply conclude that the question is “merely conceptual”, because which view is true depends on which interpretation is favored by our usage. (For example, both the question of whether parthood is reflexive, and the question of whether

there are holes, are merely conceptual. Since each view about these questions comes out true under some interpretation, the only question is which interpretation fits English usage.)

I think of the third strategy as the core of metaphysical deflationism, since it seems presupposed by the first two strategies, and is potent even if unaccompanied by either of the first two.

The deflationist's arguments are pretty unimpressive if step 3 of the go-to move is omitted. Let  $p$  be a particle in some distant galaxy. The sentence 'particle  $p$  is an electron' is a paradigm of the kind of sentence for which no form of deflationism is true. And yet, we can construct an interpretation on which it is true and an interpretation on which it is false, each of them very similar to English: let the first be like English except with  $p$  removed from the extension of 'electron', and let the second be like English except with  $p$  added. Step 3 is crucial; the rival interpretations must in some sense be on a par with one another and with other competing interpretations (such as English).

If the deflationist is a realist about structure then he can construe step 3 as requiring at least that the constructed interpretation must carve at the joints (in the right way—recall section 4.2) as well as its rivals. The rival interpretations of 'electron' do not lead to deflationism about whether particle  $p$  is an electron because, whereas the actual, English, meaning of 'electron' carves at the joints, one of the two constructed meanings does not. But when a deflationary stance is correct, the rival interpretations carve at the joints as well as one another and as well as English.

A deflationist who rejects realism about structure, on the other hand, will need some other way to explain why the go-to move has a deflationary upshot for metaphysical questions but not for the question of whether particle  $p$  is an electron; and he will need some other way to construe step 3. It is hard to see how this could be done.<sup>2</sup> So I think the way to defend a targeted deflationism

---

<sup>2</sup>Hirsch tries to target his deflationism by applying it only when the disputed metaphysical proposition is regarded by one disputant as being a priori necessarily equivalent to one undisputed proposition, and by another disputant as being a priori necessarily equivalent to another undisputed proposition. But the disputants need make no such claims about a priority and necessity. Most metaphysicians nowadays regard their views as theoretically supported conjectures, not as propositions knowable by anyone who understands them. And some of Hirsch's targets regard their subject matter as contingent (Cameron, 2007). Moreover, a priori necessary equivalence seems relevant only if one thinks of it as a kind of sameness of meaning, whereas surely it is no such thing. For a more comprehensive (and very insightful) discussion of Hirsch, see Hawthorne (2009).

is to be a realist about structure. At any rate, this is the form of deflationism on which I'll focus.

The deflationist should also, I think, take step 3 as requiring that the rival interpretations not be “semantically alien”, in the sense of section 4.2. That is, although the rival interpretations might determinately differ from English, speakers using those interpretations should use the words in question to accomplish the same semantic tasks as do English speakers. For the deflationist about a question  $Q$  wants to conclude that the answer to  $Q$  depends on how we use words, but not in the trivial sense in which *all* truth depends in part on meaning. The idea should rather be that  $Q$ 's answer depends on usage in the nontrivial sense that it depends on which *nonalien* interpretation is actual.<sup>3</sup>

Thus what the deflationist is saying is in essence that  $Q$  is *nonsubstantive*, in the sense of chapter 4. The deflationist's nonalien interpretations result from multiple *candidates* for the crucial terms in  $Q$ . For example,  $\text{part}^*$  and  $\text{part}^{**}$  are candidates for ‘part’, and generate interpretations under which different answers to the question “Is parthood reflexive?” come out true. If each carves at the joints as well as the other (and as well as any other candidate for ‘part’), then this question is nonsubstantive. As we saw, this is not to say that  $Q$  is *about* how we use words. The question of whether appletinis are martinis is not about words, and yet it too is nonsubstantive in the same sense.

What are the ways for a metaphysician to respond?

Sometimes metaphysicians should embrace deflationism. And sometimes this requires admitting that the debate is just silly, and should be discontinued. (Even a partisan needn't fight *every* battle.) But not all nonsubstantive debates are silly. As we saw in section 4.2, questions that are nonsubstantive in my metaphysical sense may yet be conceptually deep and thus important if their answers reveal important facets of our conceptual scheme. Perhaps questions about causation are like this.

But sometimes the true believer in metaphysics will want to oppose the deflationist. This is particularly true when the believer wants to defend a revisionary position, since such positions are hard to regard as reflecting our ordinary conceptual scheme.<sup>4</sup> It is also true when the believer regards herself as doing fundamental metaphysics.

In the case of fundamental metaphysics, the most straightforward way to

---

<sup>3</sup>Note also that if the interpretations are alien, then there is no hope of carrying out the common-sense and indeterminacy strategies.

<sup>4</sup>Hard but not impossible; perhaps certain internal tensions in our conceptual scheme are best resolved by preserving some aspects at the expense of others.

resist deflationism is to claim that the crucial expressions in the debate carve perfectly at the joints. As we saw in section 4.2, the relation between substantivity and joint-carving is a complex one. However, a *sufficient* condition for substantivity (or near enough, anyway) is that the dispute be cast in perfectly joint-carving terms. Such a dispute concerns the nature of fundamental reality. In such a dispute, the existence of alternate interpretations has no more deflationary import than it had in the question of whether particle  $p$  is an electron.

(Why is the condition only “near enough” sufficient? Because there might be multiple joints in the vicinity, and because the joints might be the “wrong sort”, in the sense of the third refinement of section 4.2. I’ll mostly be ignoring these complications.)

There are other ways to resist the deflationist, but they are unappealing. One might try to argue that the alternate interpretations simply do not exist. For example, the deflationist about the ontology of holes claimed to produce an interpretation under which ‘There is a hole in  $x$ ’ means that  $x$  is perforated. But the dispute over holes does not involve that sentence form alone; it involves sentences with many different syntactic forms, for example:

There is a circular hole in  $x$ .

There are fifteen holes in  $x$ .

The hole in  $x$  is identical to the hole in  $y$ .

In light of this, it might be argued that there is no way to give a *general* anti-nominalist interpretation of quantification over holes. But this is a slender reed on which to rest one’s hopes. (I argue in chapter 9 that in the case of the ontological deflationist, the hope is in vain.)

It is also possible to grant the existence of multiple equally joint-carving interpretations, but claim that those interpretations are semantically alien and thus not candidates. Consider the question of whether “Magnesium is more plentiful on Earth than carbon.” This sentence comes out true if ‘magnesium’ means oxygen, but this does not support deflationism about the question since this alternate meaning for ‘magnesium’ is not a candidate (despite carving at the joints). The semantic goal we are trying to achieve with ‘magnesium’ is not so unspecific that it could just as well have been achieved by letting ‘magnesium’ mean oxygen. However, this kind of response to deflationism seems inapplicable in cases of philosophical interest. For what is distinctively

puzzling about philosophical questions is closely connected with the fact that we *can* imagine ourselves speaking in any of a number of different ways using the disputed term, without thereby being semantically alien.

One might instead make a big deal out of the fact that the candidate meanings for the crucial term differ from its actual meaning:

“Who cares whether there are ways to interpret ‘There is a hole in  $x$ ’ so that it comes out true? What I care about is whether *there is a hole in x*; and if you define ‘There is a hole in  $x$ ’ to mean that  $x$  is perforated, it no longer says that there is a hole in  $x$ !”

This performance is oblique, since it rejects none of the deflationist’s claims. It does not challenge the claim that the interpretations exist, nor does it deny that they carve at the joints equally well, nor does it claim that they’re semantically alien. And the deflationist never claimed that ‘There is a hole in  $x$ ’ still means that there is a hole in  $x$  when it means that  $x$  is perforated. What is conveyed by the agitated speech is that the “deflationist’s” claims do not deserve the name ‘deflationism’, that the status quo in metaphysics is unthreatened by the existence of the interpretations.

This is a hard attitude to maintain. Most metaphysicians at least sometimes think of themselves as *not* being engaged in conceptual archaeology. (From this point of view, Peter Strawson’s (1959) “descriptive metaphysics” is a near oxymoron.) Instead, they think of their project as being rather like speculative science. This self-conception isn’t always articulated, but it is often subconsciously present. It reflects itself in the willingness to take noncommonsensical positions seriously, and to be guided by theoretical virtues that are prized in science (such as “simplicity”). This self-conception cannot survive the admission that rival answers to one’s question come out true under equally joint-carving candidate interpretations. Imagine we live on the shore of a gigantic lake, around which there are numerous other isolated linguistic communities. None of these is semantically alien with respect to the others, but the different communities do use ontological language differently: in one, ‘There are holes’ is true, in another it is false; in another ‘There are numbers’ is true; in another it is false; and so on. And suppose further that none of these languages carves better at the joints than any of the others. If we metaphysicians learned of all this, it would seem perverse for us to continue to regard our ontological questions with the original, quasi-scientific attitude. Granted, in some of these other languages, ‘There are holes’ does not mean that there are holes. But so what? Speakers of those languages could make parallel true statements about

us in their language: “Your sentence ‘There are holes’ does not mean that there are holes.” Look out over the lake—the parochiality of our conception of there being holes is staring us in the face. The metaphysical attitude requires a more transcendental view.

## 5.2 Personal identity, causation

Let’s make all of this a little more concrete. Consider a group of proponents of the doctrine of temporal parts who disagree over the criterion of personal identity. Their dispute is over which relation between person stages is the “unity relation” for persons—that is, over which relation holds between all and only the person stages that are part of some one continuing person. Some of them think that the unity relation is that of psychological continuity; others think that it is the relation of bodily continuity (i.e., spatiotemporal continuity under the sortal: human body).

Never mind which group has the right answer. Let us instead ask the metametaphysical question: Is the dispute substantive?

A deflationist about personal identity could argue as follows. First, there is no perfectly fundamental unity relation over person stages. There are numerous relations over person stages in the vicinity: the relation of psychological continuity, the relation of bodily continuity, and so on. But none of these carves at the joints perfectly, and none carves better than the others. Moreover, these are candidates for our talk of personal identity. Neither a community who spoke of personal identity as being governed by psychological continuity, nor a community who spoke of it being governed by bodily continuity, would be semantically alien. So the question of personal identity is nonsubstantive. There is no objective, substantive, deep answer as to whether I would survive certain transformations—those in which my psyche is transferred to a new body, for example, or in which I abruptly lose all my memories. Which answers are correct is largely a question about our conceptual scheme, not a question of reality’s fundamental structure (and our conceptual scheme might be silent about some cases).<sup>5</sup>

For another example, return to causation. If there is a perfectly fundamental relation of causation, then there must be objective, deep answers to questions like: Is causation two-place or four-place? Does it relate facts or events? But

---

<sup>5</sup>See Sider (2001a).

otherwise, deflationism about some questions about causation will presumably be true, since linguistic communities differing over whether ‘cause’ expresses a two-place or four-place relation, for example, do not seem linguistically alien. The questions about causation may have no answers at all; and insofar as they do have answers, those answers will be a mere reflection of our concept of causation.

The distinction between conceptual and metaphysical substantivity (section 4.2) is particularly important in metametaphysics. Even if questions about personal identity and causation are not metaphysically substantive, they clearly are conceptually substantive, given how deeply the concepts of sameness of person and causation are embedded in our conceptual scheme. These concepts play a role in many central aspects of our thought: thought about ourselves, moral responsibility, deliberation, control, prediction, explanation, and myriad others. Even if this entire edifice rests on metaphysical sand, understanding its inner workings is a crucial part of understanding ourselves.

Disputes over causation, personal identity, and the like are conceptually deep, even if metaphysically shallow; this is a significant part of their philosophical interest. Conversely, and ironically, some of the metaphysically deepest disputes—certain ontological disputes, for example (see chapter 9)—are conceptually shallow in that they have few implications outside rarified metaphysics. This, I suspect, contributes to the common distrust of those disputes.<sup>6</sup>

### 5.3 The metaphysics room

Sometimes fundamental metaphysics can be conducted in ordinary language. But not always. Metaphysicians need a plan B.

Suppose we attempt to ask a question of fundamental metaphysics using some ordinary, natural-language expression *E*. Suppose further that there is some joint-carving meaning *m* “in the vicinity” (see below). The problem is that *E* might not mean *m*. Various metasemantic scenarios discussed in section 3.2 could have this result; here are three. 1. Reference magnetism is false in all its forms, and *E* is not a theoretical term—like ‘amulet’ it is not “trying” to stand for a joint-carving meaning. 2. Some form of reference magnetism other than simple charity-based descriptivism is true, and *E* is not a theoretical term. 3. Simple charity-based descriptivism is true, but *m* fits badly with

---

<sup>6</sup>Thanks to Eric Funkhouser here.

the use of  $E$ . (According to simple charity-based descriptivism, the correct interpretation must maximize the combination of charity and eligibility; thus a highly joint-carving interpretation might nevertheless be incorrect if it is too uncharitable. Not all words carve at the joints, after all!)

If  $E$  does not stand for  $m$ , then it might instead stand for some non-joint-carving meaning that reflects a more-or-less arbitrary choice of usage that our linguistic community has made—a choice of one amongst a range of equally non-joint-carving candidate meanings. Thus the question would be nonsubstantive in that its answer would turn on linguistic usage, not reality’s structure. (More accurately, the answer would be actual-verdict nonsubstantive in the sense of section 4.2.)

That is, the *ordinary, natural language* question, phrased in terms of the ordinary, natural-language expression  $E$ , would be nonsubstantive. But we could discard  $E$ , and enter the metaphysics room, so to speak. We could replace the ordinary expression  $E$  with an improved expression  $E^*$  that we stipulate is to stand for the joint-carving meaning in the vicinity. The question we ask in the metaphysics room, cast in terms of  $E^*$  rather than  $E$ , is substantive. Indeed, it is *superior* to the original question, for it concerns reality’s fundamental structure, rather than its merely conventional or projected aspects. This is plan B.

Early on in philosophy we are taught not to abandon ordinary language on the battlefield. If a novel language had to be stipulated in order to carry on a debate, we’re warned, there could not be open questions about what is true in that language, since the answers would need to be settled by stipulation. Arguing about what is true in the novel language would be like arguing about how the queen ought to move in a new version of “chess” whose rules are unconstrained by the existing rules. The traditional debate over whether freedom is compatible with determinism, for example, would be trivialized if we had to stipulate a meaning for ‘free’. If we stipulated that ‘free’ means ‘undetermined by the laws and past’ then there would be nothing worth debating: “freedom” thus understood is obviously incompatible with determinism. And if we instead stipulated that ‘free’ means ‘not in chains’, then again we would have nothing worth debating; “freedom” thus understood is obviously compatible with determinism. The only way to have a meaningful debate, so we are taught, is to abandon such stipulations, and mean by ‘free’... *freedom!*—freedom in the ordinary sense.<sup>7</sup>

---

<sup>7</sup>Qualification: *some* stipulation might be allowed, if it is ultimately grounded in ordinary terms (‘free’ could be stipulatively tied to ‘moral responsibility’ in its ordinary sense).

Abandoning ordinary language is indeed often a bad idea, but when it is, that is because there is no *other* way to anchor the debate, no other way to explain the meanings of the crucial words without trivializing the debate. But joint-carving meanings give us another anchor. We can then introduce new words with a *minimum of semantic pressure*, with only minimal stipulations on their behavior. Of course the stipulations must be strong enough to pick out unique meanings (by picking out the relevant “vicinity”); but this aside, we stipulate nothing that settles disputed claims. We nevertheless succeed in securing unique meanings because joint-carving meanings are sparse—there is usually at most one joint-carving meaning in a given “vicinity”. (To suppose otherwise would be to needlessly attribute complexity to the world.) Picture introduction of terms with minimal semantic pressure as something like definition by ostension. We are saying “Let us introduce words for *these* meanings, so that we can disagree about how *they* behave.” If there are indeed (unique) joint-carving meanings in the vicinity—the anchor—then the words in question will have determinate meanings even though the stipulations were so minimal. The answers to the disputed questions will not be settled by the stipulations, and inquiry into those questions—questions about reality’s fundamental structure—will be worthwhile.

Let’s consider an example. Suppose there is a fundamental relation *C* that is a lot like causation, except that it holds only between events at the subatomic level. Macro events, such as the throwings of rocks and the breakings of windows, never stand in *C*. Now, the ordinary English term ‘cause’ may well not mean *C*. For i) *C* fits terribly with ordinary usage of ‘cause’ (or at least with usage of simple causal sentences such as ‘the throwing of the rock caused the window to break’); and ii) ‘cause’ may well be a nontheoretical term in English. Rather than standing for *C*, ‘cause’ may instead stand for that non-joint-carving relation that best fits our usage of ‘cause’. A debate involving ‘cause’ would then not be (actual-verdict) substantive. But we could enter the metaphysics room, and coin a new term, ‘cause\*’, for the joint-carving relation in the vicinity of causation. ‘Cause\*’ will stand for *C*—fundamental causation, we might call it—and our new debate about causation\* will be substantive.

How, exactly, will we fix the meaning of ‘cause\*’? It is to stand for the joint-carving meaning “in the vicinity of” causation. Thus its metasemantics should be like that of ‘cause’ except that joint-carving is paramount. Any general metasemantic presumption that non-joint-carving candidates may be assigned is suspended in the case of ‘cause\*’; carving perfectly at the joints is an absolute requirement. ‘Cause\*’ should stand for that meaning which i) carves at the

joints (perfectly); ii) fits our use of ‘cause’ better than any other joint-carving meaning; and iii) fits our use of ‘cause’ well enough. If nothing satisfies all three of these conditions—if, that is, there is no fundamental causation, or if there is more than one sort—then ‘cause\*’ stands for nothing, and debates about causation\* are ill-posed. Conditions ii) and iii) are admittedly vague, but harmlessly so. What might we actually do to coin a new term with this metasemantics? I see no reason to deny that the following performance would do the trick: “Let ‘cause\*’ be a theoretical term for that perfectly joint-carving relation (assuming there is such a thing) that is in the vicinity of the ordinary notion of causation, but which (since it carves at the joints) may differ somewhat from that ordinary notion.”

The metaphysics room gambit requires that successful stipulations of the envisaged sort be possible. It may be objected that new languages or terms can be introduced only by translation into existing natural languages or terms.<sup>8</sup> But that can’t be right. Natural languages themselves had to bootstrap; they had to somehow latch onto the world in the first place. So why can’t we bootstrap now? It may be thought that the stipulation could not succeed if reference magnetism is false, or worse, if some radically nonreferential conception of meaning—some form of deflationism or inferentialism, perhaps—is true. But even such views must allow for the introduction of new theoretical terms in science. If they don’t, then they are thereby refuted. And if they do, then given realism about joint-carving they surely also allow the introduction of theoretical terms in the metaphysics room. The cases really are parallel: in each case we introduce new terms with minimal semantic pressure, we intend to mean “whatever is in the vicinity”, and semantic determinacy is achieved (if it’s achieved) primarily because the world contains an appropriate meaning, not primarily because of facts about us.

The gambit also requires that semantic interpretation—at least for terms stipulated to carve at the joints—be under our control. But it does not require that grammar be under our control in the same way. Consider, for example, the fact that natural language quantifiers are grammatically binary, whereas first-order logic’s quantifiers are monadic. It might be argued that for this reason, the first-order quantifiers are in some sense grammatically impossible for us. Even if we try to pronounce “ $\exists x$ ” in a monadic-looking English form, “For some  $x\dots$ ”, we nevertheless continue, despite ourselves, at some psychologically basic

---

<sup>8</sup>Hofweber’s (2009) critique of “esoteric” metaphysics—metaphysics done using distinctively metaphysical rather than ordinary concepts—is related.

level, to represent the quantification binarily: “For some *object*  $x$ ...”. Binary quantificational grammar is built into our minds in a way that’s difficult or even impossible to change. But even if this view is correct, it is no obstacle to introducing the language of first-order logic in the metaphysics room and giving it a monadic *semantics*. The formal sentences of this language would describe fundamentally monadic facts; it’s just that we couldn’t think about those facts “natively”. There would be a mismatch between the structure of the facts and the structure of our thoughts (though not between the facts and the sentences of the formal language).

## 5.4 Substantivity in nonfundamental disputes

The simplest way to regard a dispute as substantive is to regard it as being phrased in terms that carve *perfectly* at the joints. But sometimes one wants, instead, to regard a dispute as substantive because it concerns reality’s less-than-perfectly fundamental joints.

To illustrate, it is natural to wonder whether anything substantive is at stake in the dispute over the location of the semantics/pragmatics border. Everyone agrees that context supplies an enormous amount of information that is relevant to communication. But is that information part of the *semantics* or the *pragmatics*? For there to be any interesting issue here, we must fix the notions of semantic and pragmatic without begging any disputed questions. For example, if it is to be an open question whether the semantic contents of sentences are (complete, truth-conditional) propositions, we must not simply define the semantic content of a sentence as a proposition that is conventionally associated with that sentence.

To secure substantivity, the debate should be regarded as being about the way to carve the subject matter of communication at its natural joints—about the form that a good theory of communication ought to take. Two participants in this debate might advocate their positions as follows. *Participant one*: “A good theory of communication divides the processing of communicated information into two stages. At stage 1, a fairly simple level of content is associated with each word. This level of content is memorized by the competent speaker. Stage 2 combines the output of stage 1 with all the information provided by the context, resulting in truth-conditions of the utterance. In computing stage 2 content, the speaker appeals to her general world-knowledge. Further, stage 1 associates a complete proposition—something with truth-conditions—with

each grammatical sentence.”<sup>9</sup> *Participant two*: “I agree about the two stages, but I disagree with your final claim. Stage 1 does not associate a complete proposition with each grammatical sentence; sometimes it instead associates what Kent Bach (1994) calls a ‘proposition radical’.” We can think of this dispute as being over whether semantic contents are always complete propositions, since we can think of stage 1 as semantics and stage 2 as pragmatics. And the dispute isn’t terminological since ‘semantic’ and ‘pragmatic’ have been picked out by their role in a good theory of communication, and not by a definition that begs the question of whether semantic contents are propositions.

The story I just told in effect picks out the semantics/pragmatics divide via its role in the Ramsey sentence for a theory of communication. This pattern—a term of philosophical interest is picked out by its role—recurs all over philosophy.<sup>10</sup> Think of debates over perception. How much is contained in “the content of visual experience”, it is asked.<sup>11</sup> Information about real-world objects, or just their appearances? Causal information? This debate can seem puzzling to outsiders; how is the meaning of ‘the content of visual experience’ to be fixed so that the questions remain open? The answer must be that the participants in this debate are attempting to carve the subject matter of perception at its joints. A good theory of this domain, they believe, will appeal to a certain notion of the content of visual experience; and their debate is about *that notion*—the notion that appears in a good theory of perception. (The role might be expansive. Some, for example, pin down the notion of the content of experience by its role in epistemology.)

Construing a dispute in this way doesn’t on its own secure substantivity, because it might be that no joint-carving theory takes the shape that the disputants envisage. In the dispute over the border between semantics and pragmatics, a third participant might deny the claim shared by participants one and two, that a good theory of communication has anything like the two stages. If she is right, then in a sense there would be no substantive questions about semantics and pragmatics, since the distinction is based on a false theory. Either our uses

---

<sup>9</sup>The two-stage process is oversimplified; a more realistic account would divide linguistic processing into multiple stages (phonological, syntactic, logical form...), and contextual information might intrude at multiple points. The point still stands: there can be substantive dispute about the semantics/pragmatics border if the disputants can agree on a sufficiently detailed skeleton of a theory of linguistic processing so that they can disagree over where, in that skeleton, truth-conditions figure in.

<sup>10</sup>See also Chalmers (2011).

<sup>11</sup>See Siegel (2006).

of ‘semantic’ and ‘pragmatic’ are empty of meaning, or they are massively indeterminate, or their meanings are determined simply by how they are casually used in ordinary speech.<sup>12</sup> (Even so, a substantive three-way debate over the shape of a good theory of communication would remain.)

So: substantivity, in these questions of perception and meaning, turns on the nature of reality’s joints. (Likewise for questions of substantivity in the special sciences, although this is less often in dispute.) But reality presumably contains no *perfectly* fundamental perceptual or meaning structure. So we must make sense of joint-carving in domains that are not perfectly fundamental. We will revisit this issue of imperfect fundamentality in section 7.11.

## 5.5 A test case: extended simples

To illustrate my approach to metametaphysics, let us look in detail at the recent dispute over “extended simples”.<sup>13</sup> My guess is that many readers suspect that this dispute is not substantive (as I myself did initially). Is this suspicion justified?

“Extended simples” are defined as spatially extended objects without proper parts. For example, if *I* am an extended simple then *I* have no proper parts. Not only do I lack “philosophers’ parts” such as a “right half”, I also lack such ordinary parts as a head, arms and legs, internal organs, molecules, atoms, and subatomic particles. There is, in my vicinity, but a single thing—me—which occupies an extended, person-shaped region of space.<sup>14</sup> A limiting case would be existence monism, in Jonathan Schaffer’s (2007; 2010c) terminology, according to which the entire world is a single extended simple.<sup>15</sup> The dispute is over whether there are, or could be, such things.

The status of this dispute is tied up with general questions about space and ontology. Are ontological questions in general substantive? Do there exist

---

<sup>12</sup>The following can happen. (a) A philosophical debate began when there was no decent scientific theory of a certain domain. (b) Subsequently a good scientific theory was developed. (c) No good rival philosophical theory was developed; the concepts of the original philosophical debate were simply badly chosen; they carved nature at its joints badly. Nevertheless, (d) the philosophical debate continues uncritically, in the old terms.

<sup>13</sup>See Hudson (2006, chapter 4); Markosian (1998, 2004a,b); McDaniel (2003, 2007a,b); Parsons (2003).

<sup>14</sup>In this example we have a qualitatively heterogeneous extended simple, which goes beyond the bare definition of an extended simple. See Sider (2007b, section 1).

<sup>15</sup>See also Horgan and Potrč (2000, 2002).

such entities as points and regions of space? If so, are objects in space to be identified with points or regions? If not, how are objects in space related to space? The answers to these questions affect whether disputes over extended simples are substantive.

To illustrate, suppose for the sake of argument that the following answers to the general questions—call them, collectively, the *occupation picture*—are correct:

1. The notions of logic, mereology, and physical geometry (quantification, identity, parthood, sum, point, region, ...) carve perfectly at the joints.
2. Substantivalism is true: there exist points and regions of space.
3. Supersubstantivalism is false: objects in space (“inhabitants”) are not identical to points or regions of space.
4. Spatial facts about inhabitants emerge from the holding of a perfectly joint-carving relation of *occupation* between inhabitants and space.

Now, the question of whether there are any extended simples, given the occupation picture, becomes a question about the pattern of instantiation of the occupation relation:

Does there exist something that lacks proper parts but occupies more than one point of space (i.e., occupies multiple points or a sum or region of points)?

And given my metaphysics, this question is substantive, since it is phrased in terms that carve perfectly at the joints.

I have seen philosophers roll their eyes when extended simples come up. They regard that topic as “spooky metaphysics” (in a bad sense, it would seem)—as being somehow misguided. But what is the complaint?

In fact, there are a number of complaints that one could be making. The list of possible complaints, I hope, challenges both the extended-simple enthusiasts and the eye-rollers. The enthusiasts must check that they avoid each complaint. The eye-rollers must check they can genuinely embrace one of them, and they must specify which one that is (the eye-rolling tends to be scattershot and uncritical).

One complaint is purely epistemic. It admits that the dispute is substantive, but claims that the considerations offered by the enthusiasts are inadequate to

resolve it. We do not know whether there exist extended simples, it is alleged, and the enthusiasts' arguments aren't helping. Such allegations may or may not be right, but one cannot make them without actually entering into the fray. (Eye-rolling tends to come from the sidelines.) Also, this complaint does not demand that the enthusiasts cease and desist. The enthusiasts might instead take it as an exhortation to do better.

A second complaint concedes that the question of whether extended simples *actually* exist is substantive, but notes that some enthusiasts address primarily the question of whether they *possibly* exist. This shift makes a big difference to the tenability of deflationism about the debate. Given my metametaphysics, a deflationist about an issue must locate a nonfundamental term essentially involved in its statement. Once the issue becomes the *possibility* of extended simples, it is open to claim that the nonfundamental term is 'possibly'. In particular, suppose that there are, in fact, no extended simples. One could then claim that the question "Is it possibly the case that there exist extended simples?" is nonsubstantive, on the grounds that its answer is *yes* under some candidate meanings of 'possibly' and *no* under other equally joint-carving candidates.

The theory of possibility to be defended in chapter 12 could sustain such a claim. Very roughly, the theory says that for a proposition to be necessarily true is for that proposition to be i) true; and ii) of the right sort. What is the "right sort"? This is given by a list of properties of propositions, a list including the property *logical truth*, the property *mathematical truth*, and certain others. Moreover, the list is determined more or less arbitrarily, rather than by some deep criterion. *Our* meaning of 'necessity' is tied to *our* list; but other linguistic communities could choose different lists and thus different meanings for 'necessarily'. They wouldn't thereby carve at the joints worse than we do; they would just be different—like a linguistic community that counted the pope as a "bachelor". Given this theory of modality, different candidate meanings for 'necessary' arise from different choices of lists. If the true proposition that there are no extended simples falls under some of these lists but not others, then 'necessarily there are no extended simples' counts as true under some candidates and not others; and so, 'possibly, there are extended simples' counts as true under some candidates and not others.<sup>16</sup> Intuitively: extended simples would depart from actuality in certain ways; the question of whether they are possible is the question of whether these departures are too drastic; whether the departures are too drastic turns on the metaphysically shallow question of

---

<sup>16</sup>As is customary, I define "possibly  $\phi$ " as "not necessarily not- $\phi$ ".

what we have decided to mean by ‘possible’.

Our first two complaints interact. Given the second complaint, the only substantive issue can be the question of the actuality of extended simples. But much of the literature on extended simples has discussed only their possibility. If this portion is omitted, the first complaint becomes more powerful.

A third class of complaints results from rejecting one facet or another of the occupation picture. An *ontological deflationist*, for example, who thinks that quantifiers do not carve at the joints, could claim that the sentence ‘there are extended simples’ is true under some candidate meanings for ‘there are’ and false under others, and that all such candidates carve at the joints equally well. Thus the sentence is nonsubstantive.

There are many questions here. There is first the question of whether ontological deflationism is true; I will argue in chapter 9 that it is not. There is also the question of whether ontological deflationism is behind the eye-rolling; I think that in most cases it is not. Most of the people I have caught rolling their eyes are not skeptical of ontology in general (though perhaps their eye-rolling is a sign of an underlying malady that would lead to ontological deflationism, or worse, if not treated). And there is, finally, the question of whether ontological deflationism is enough to deflate all of the disputes in the vicinity. It’s hard to tell, because ontological deflationists tend not to say what they think the fundamental structure of reality *is*; they merely say what they think it *isn’t* (namely, quantificational). Once the deflationist has developed his fundamental language—which must be able to completely describe the world, including its scientific aspects—who knows? Perhaps we will be able to raise, in its alien terms, a substantive question that is akin to the question of whether there are extended simples.<sup>17</sup>

A fourth and final sort of complaint amounts to simply rejecting extended simples—perhaps all extended simples, or perhaps only certain kinds. This may not seem like a metaphysical complaint, but it feels more like one if the reasons for the rejection are sufficiently general. Consider, for example, what Kris McDaniel (2003) calls “spanners”: extended simples that occupy an extended region, rather than occupying each of the many points in that region. One might object to spanners because they violate a sort of micro-reductionism, namely, the view that all fundamental properties and relations relate mereologically simple entities. (One might, in turn, base this on the view that only mereologically simple entities exist, in the fundamental sense of

---

<sup>17</sup>See also section 9.6.2.

‘exist’; see chapter 13.)

We have seen four different sorts of complaint about extended simples.<sup>18</sup> Notice that, with the exception of the first, each presupposes a substantive thesis of metaphysics. The second depends on the claim that there are no modal joints to reality; the third depends on the claim that there are no quantificational joints to reality; and the fourth assumes micro-reductionism. No complaint wholly sustains the common attitude of the eye-rollers that the extended-simple enthusiasts are making a *methodological* error. The complaints mark a difference of opinion about metaphysics. As we will see in the next section, this is a general feature of metametaphysical critiques.

## 5.6 Metametaphysics as just more metaphysics

One worry about metaphysics is that we have no way to answer metaphysical questions. But is metaphysics so much worse off than the rest of philosophy?<sup>19</sup> Many scattershot critics move from the difficulties of metaphysical epistemology to the conclusion that metaphysics is something like meaningless, without realizing how close this comes to assuming a crude form of verificationism. Often these critics have a blind spot: they are verificationist when thinking about metaphysics, but not when thinking about other matters (especially about

---

<sup>18</sup>A fifth complaint might target the dispute only when extended simples are construed as having “pseudo-parts”—further extended simples that are not genuine parts (i.e., parts in the joint-carving sense) but are located where genuine parts would be located. This dispute might be claimed to be nonsubstantive because the truth of ‘there are extended simples’ turns on whether ‘part’ means genuine parthood or pseudo-parthood. The claim might be correct if ‘part’ is not a theoretical term; for then, even though exactly one candidate meaning for ‘part’ carves at the joints, that candidate might be the “wrong sort”—recall the third qualification of section 4.2. But if ‘part’ is a theoretical term then the dispute remains substantive, for it is then cast in purely joint-carving terms of the right sort. Thanks to a referee. A sixth complaint might be that the question just isn’t significant. (The old joke about angels dancing on the head of a pin combines this and the first complaint.) And a seventh complaint might be that the answer is obvious in some sort of Moorean way. Neither the sixth nor seventh complaint challenges the substantivity (in my sense) of the question. Moreover, each seems wrong. Why wouldn’t a question about the fundamental relationship between objects and space be significant? And why should the nature of this relationship be obvious to “common sense”? (See also the critique of Mooreanism in Sider (2011).)

<sup>19</sup>Timothy Williamson (2007) gives a compelling account of the epistemology of philosophy (and of the sociology of pessimism about philosophy) with which I mostly agree, except that for the epistemology of metaphysics I would stress continuity with scientific explanation rather than counterfactual reasoning.

their own bit of philosophy!). A sensible attitude is that metaphysics, like much of philosophy, is just hard. Its epistemology is hard too, but this is no cause for panic.

A different (though related) worry is the feeling that metaphysicians see substantive issues where there are really just different equally good ways to talk. It is on this strand—metaphysical deflationism—that I have focused.

Details aside, my crucial claim has been that a sufficient condition for substantivity is being cast in joint-carving terms. An important consequence is that metametaphysical critiques are distinctively metaphysical in nature. Whether they are correct is a *function of the facts*—a function of what joints reality in fact has. One cannot do metametaphysics simply by examining metaphysical language and reasoning. For given the sufficient condition, in order to claim that a question is nonsubstantive, one must claim that it is not cast purely in joint-carving terms, and such a claim cannot be supported solely by reflecting on language and reasoning. For example, we saw in the previous section that a sufficient condition for the extended-simples debate being substantive is that a certain metaphysical thesis be *true*: the occupation picture. Thus, in order to decide whether that debate is substantive, one must directly engage in metaphysics.

It may be objected that this conclusion—that metametaphysical critiques are metaphysical in nature—is simply an artifact of *my* way of conceiving of metametaphysics. The objection is partly correct. An austere metaphysician who rejected realism about structure could perhaps introduce some sensible notion of substantivity that doesn’t presuppose the notion of carving at the joints. But even so, his metametaphysical critiques could not be wholly ameta-physical. For imagine a metaphysician who shrugs off a purely methodological or linguistic critique by saying that *she* is a realist about carving at the joints, and that she believes her ideology to carve at the joints. Surely the austere critic couldn’t simply concede these claims. For, I hope, this book provides a *model* of metaphysics; a model of how, in a world with objective structure, language could be attached to that structure and metaphysicians could reasonably speculate on its nature. So even the austere critic should agree that *given* the realism about structure, metaphysics would make sense, both linguistically and methodologically. So the austere critic must oppose the metaphysician’s claims. Granted, he could oppose them by rejecting the realism about joint-carving, rather than by accepting this realism while denying that the metaphysician’s ideology carves at the joints. But this rejection of joint-carving is just more metaphysics.

Many are drawn to metaphysical deflationism because they want the epistemic high ground. They want to rid the world of difficult-to-answer substantive questions. But their very metaphysical deflationism costs them the high ground. For since metametaphysical critiques are just more metaphysics, they raise all the old epistemic questions. This is certainly true if the critic embraces realism about structure and claims that the crucial notions in the targeted debate do not carve at the joints—the epistemology of such a claim is as hard as can be. But it's even true if the critic is austere, since *rejecting* realism about joint-carving raises the same old epistemic issues.

This does not mean that metaphysical deflationism is never reasonable. It's often attractive to avoid some bits of metaphysics, some difficult questions. But it does mean that no one should adopt metaphysical deflationism with the goal of avoiding metaphysics, or difficult questions, in their entirety.

Nor should one adopt deflationism about a metaphysical question merely because the ideology needed to state that question differs from one's own. This should go without saying, but in fact, many critics have a blind spot here too. They notice that the crucial term *T* in a metaphysical question can be understood in multiple ways in terms of their own ideology, which they regard as fundamental, never stopping to think that the participants in the debate regard *T* itself as being fundamental. In the dispute over holes, certain critics keep complaining that they don't see the question. "Is it whether holes are causally efficacious? Is it whether some of their properties fail to supervene on the presence of matter and its arrangement?" No, the question is simply: Do holes *exist*?!

# Chapter 6

## Beyond the predicate

STRUCTURE IS MORE GENERAL than its kin, Armstrong's universals and Lewis's natural properties and relations, along two axes.

The first axis is ontological. Call a predicate "sparse" when it marks a joint in nature. For Armstrong, a predicate is sparse when there exists a corresponding universal; for Lewis, a predicate is sparse when there exists a corresponding natural property or relation.<sup>1</sup> Each assumes the existence of abstracta. But the idea that the world has a distinguished structure—that electrons go together and not together with cows, that it is better to think in terms of electrons than in terms of electron-or-cows, and so on—does not require this assumption. (Nominalists could surely embrace the idea.<sup>2</sup>) The notion of structure is to be free of commitment to abstract entities.

The second—more important—axis concerns the scope of structure. Armstrong and Lewis's accounts are confined to properties and relations. Linguistically speaking, their focus is on the predicate. Structure, on the other hand, is not to be restricted to any particular grammatical category. Just as Lewis and Armstrong ask which predicates get at the world's structure, we can also ask which function symbols, predicate modifiers, sentence operators, variable binders, and so on, get at the world's structure.

One might force these expressions into the Armstrong/Lewis mold, by ana-

---

<sup>1</sup>Lewis (1983b, 347–8) considers replacing his primitive predicate of natural sets with a primitive multigrade predicate of similarity amongst possible individuals. But this proposal does not generalize along the second axis, and it arguably requires Lewis's modal realism.

<sup>2</sup>Not that Armstrong and Lewis claim otherwise. They simply formulate the idea in their own terms.

lyzing them using predicates and then assessing the predicates for sparseness as before. Sometimes this is harmless; one can treat function symbols as standing for relations, for example. But in other cases, it foists unwanted commitments on us. For example, if sentence operators for negation and disjunction are to be predicates, they must presumably be predicates of propositions. Thus there must exist propositions; and further, even friends of propositions might doubt that the most fundamental negative and disjunctive facts are facts *about* propositions. (We will return to this last point.) And in still other cases the predicate strategy is unavailable. Variable binding expressions, for example, seem impossible to treat as predicates. (One can take quantifiers to be predicates rather than variable binders, but this just moves the bulge in the carpet: new variable binders must then be introduced. Stalnaker (1977), for example, treats quantifiers as not binding variables, but his variable-binding operation of complex predicate formation cannot be thought of as predication. Similar remarks apply to Montague's treatment of quantification.<sup>3</sup>)

The two axes are connected. For Armstrong and Lewis, a predicate is sparse when it stands for an appropriate sort of abstract entity. When we move beyond the predicate, for instance to sentential connectives or quantifiers, it becomes increasingly strained to think in terms of abstract entities. So, insofar as one is wedded to abstract entities (first axis), it is hard to move beyond the predicate (second axis).

## 6.1 The reason to generalize

*Why* move beyond the predicate? Because the connections with the notion of structure that we have been exploring are not confined to the predicate.

For example, I have argued that a dispute's substantivity—whether it concerns reality or just how we conceptualize reality—depends on how well its crucial expressions carve at the joints. Sometimes those crucial expressions are indeed predicates; disputes over causation, for example, are usually cast in terms of predicates of facts or events. But other times the crucial expressions are not predicates, and the question of substantivity can be posed nevertheless. In certain disputes over ontology, time, modality, and classical logic, for example, the crucial expressions are quantifiers (or perhaps names), tense operators, modal operators, and negation and other sentential connectives, respectively.

---

<sup>3</sup>See Montague (1973) (Dowty et al. (1981) makes it go down easier).

(These disputes will be discussed in subsequent chapters.) As before, one can attempt to force-fit the predicate mold. But often this can be done only given the existence of disputed entities (such as propositions); and with disputes involving quantification, predicate reinterpretation seems unavailable.

Advocates of restrictive ontologies will be particularly loathe to formulate their distinctive claims using predicates. Rejectors of both events and facts, for example, cannot construe causal claims as involving a predicate of events or facts; they might instead use a sentence operator: “that  $\phi$  causes it to be the case that  $\psi$ ”. Rejectors of propositions cannot regard modal and temporal claims as involving predicates of propositions; they might instead use sentential modal and tense operators. Rejectors of sets cannot define generalized quantifiers such as ‘finitely many’ and ‘most’ in terms of set theory; they might instead take them as primitive.

And even defenders of more permissive ontologies often require nonpredicates in subtle ways. For example, someone who regarded modal and tense operators as really being predicates of semantic entities might still require primitive nonpredicates such as ‘the proposition that  $\phi$ ’ and ‘the property of  $F$ -ing’ to characterize the space of semantic entities. (These are not predicates; their grammar is to turn sentences and predicates, respectively, into terms.)

To assess the substantivity of disputes whose crucial terms are nonpredicates, then, we need to go beyond the predicate. Such disputes also introduce a further need for doing so, within epistemology. As we saw in section 2.3, ideological simplicity is part of what makes a theory choiceworthy. Further, in choosing a fundamental theory we often face the choice of whether to adopt increased ideology or ontology. But theories which forgo ontology at the cost of increased ideology often do so precisely by introducing distinctive non-predicate ideology. The function of predicates, after all, is to combine with singular terms—which denote entities—to make statements; renouncing certain entities can thus call for distinctive modes of expression other than predicates. So to evaluate proposed trades of ontology for ideology—which we must do in order to choose which fundamental theory to believe—we need to speak of joint-carving for expressions other than predicates.

We have seen two ways in which the proposed “applications” for the notion of structure are not confined to the predicate. There are many more. (To mention just two: the doctrine of reference magnetism should be extended to nonpredicates; and in the Lewisian theory of laws, the logical expressions in the lawmaking language must be required to be joint-carving; otherwise cheap simplicity can be obtained with rigged logical expressions just as through

rigged predicates.) The domains in which we need to speak of structure are not confined to the predicate. We need a broader conception than that of Armstrong and Lewis.

But this broad conception of structure often meets resistance. The following sections examine various sources of this resistance, and argue in each case that the resistance is misguided.

## 6.2 Inapplicability of the similarity test

One common source of resistance is the apparent failure of a similarity conception of structure to apply to expressions other than predicates. Armstrong and Lewis tie their accounts to similarity. Lewis, for example, says of his perfectly natural properties:

Sharing of them makes for qualitative similarity, they carve at the joints, they are intrinsic, they are highly specific, the sets of their instances are *ipso facto* not highly miscellaneous ... (1986b, p. 60)

For many, this “similarity criterion” is the only handle they have on Lewisian naturalness. But it seems inapplicable to quantifiers and sentential connectives, for example. Quantifiers and sentential connectives aren’t “shared”, nor do they have instances, so we can’t assess whether their sharing makes for qualitative similarity or whether the sets of their instances are highly miscellaneous. As for “carving at the joints”, that metaphor suggests a similarity-theoretic reading: i) *do* carve where there’s a joint, i.e., assign different natural properties to dissimilar objects; and ii) *don’t* carve where there’s no joint, i.e., assign similar things the same natural properties. And neither i) nor ii) seems to apply to quantifiers or sentential connectives. So it’s easy to see how someone whose only entrée to naturalness is similarity would be baffled by the more general notion of structure.

Worse, insofar as the similarity criterion can be applied in these cases, it seems to deliver the wrong results. Does the existential quantifier carve at the joints? I will argue in chapter 9 that it does. But every two things, no matter how dissimilar, share the feature of *existing*; so doesn’t that mean that the existential quantifier in some sense fails the similarity criterion?

The worry is not that since natural properties are *defined* as similarity-makers, the notion of naturalness cannot be generalized to quantifiers. Lewis

doesn't *define* 'natural' at all.<sup>4</sup> The worry is rather the following. As an undefined theoretical term, 'natural' is understood through its theoretical role (section 2.1). The theoretical role consists of principles specifying how naturalness relates to certain other notions, such as similarity. My term 'structure' has its theoretical role fixed by many of these same principles, and more besides. If the theoretical role of 'structure' is exhausted by its connection to similarity, and if this connection is restricted to predicates, then we have no understanding of how 'structure' could apply to expressions in other grammatical categories.

But the theoretical role that Lewis offers for naturalness is *not* exhausted by the similarity criterion. Still less is the theoretical role for the more general notion of structure exhausted by the similarity criterion. The connection to similarity is just one of a network of theoretical connections that give the notion its life.<sup>5</sup> These are the connections we have been exploring, the connections to laws, explanation, metasemantics, epistemology, physical geometry, substantivity, objectivity, and epistemic value. None of these further aspects of structure's theoretical role relies on similarity. So even if we set aside similarity, structure has a rich enough remaining theoretical role to be intelligibly applied beyond the predicate.

But in fact, we needn't set aside similarity. The connection to similarity can be maintained, even for expressions other than predicates, if that connection is properly understood.

First we must expose a confusion. A putative reason for doubting that the existential quantifier carves at the joints was the observation that every two objects, no matter how dissimilar, share the feature of existence. But this observation is irrelevant. It concerns the *predicate* 'exists', whereas our question concerns the existential quantifier.

Next we must change our focus, from object-similarity to fact-similarity. The connection between similarity and structure, in the case of the existential quantifier, should be understood thus: if existential quantification carves at the joints, then whenever two facts are existential facts, that is a genuine similarity

---

<sup>4</sup>Nor *could* he have defined natural properties as similarity-makers, at least not without some serious chisholming (along the lines of Hirsch (1993, chapter 3, section 3, and appendix 2)). He doesn't count extremely specific intrinsic properties—for example, the property shared by all and only perfect duplicates of a certain porcupine—as being perfectly natural, but their sharing certainly makes for similarity. Indeed, their sharing makes for *more* similarity than the sharing of the properties that Lewis does count as being perfectly natural: the fundamental properties of particle physics.

<sup>5</sup>See also Sider (2004, p. 682).

between them. The fact that *there is a donkey* and the fact that *there is an electron* are genuinely similar in that each is an existential fact. Similarly, to evaluate whether ‘or’ ensures similarity, we must ask whether any two disjunctive facts are *ipso facto* similar; to evaluate whether modal operators insure similarity, we must ask whether all facts of the form *possibly*,  $\phi$  are *ipso facto* similar; and so on.

Language is no more a guide to fact-similarity than it is to similarity in general. Just as the applicability of ‘grue’ to each of a pair of particulars does not guarantee the similarity of those particulars, the recurrence of ‘there is’ in the sentences used to state a pair of facts does not on its own guarantee the similarity of those facts. If all existential facts are indeed similar, this is because of something about quantification: being existential is a genuine, not merely nominal, feature of facts. (We can invent words with the same grammar as quantifiers for merely nominal features of facts. “There schmexists an  $F$ ”, recall, means that the property of being an  $F$  is expressed by some predicate in some sentence of this book. It is a fact that there schmexists a donkey, and also a fact that there schmexists a brilliant Republican president born in New Haven, but these facts share nothing in common.)

Why this move to fact-similarity? The idea of a genuine similarity is that of a real commonality. Here we have a negatively charged thing; there we have another negatively charged thing; has something in nature recurred? Is there a real commonality between the negatively charged things? We think so. The recurrence was within objects in this case, but that is inessential to the legitimacy of the question. There exists a donkey; there exists an electron—has anything recurred? Is there something in common between there existing a donkey and there existing an electron? If I am right that quantification carves at the joints, then the answer is yes; these facts share a real commonality. But if I am wrong (as defenders of “quantifier variance” think—see chapter 9) then the answer is no; quantificational facts do not particularly “go together”.

It has been convenient to speak of facts in order to introduce this broad conception of similarity, but the similarity judgments in question don’t really require reifying facts.<sup>6</sup> The idea of a genuine commonality, of recurrence, requires no *things* in which the recurrence occurs. Such nonontic similarity judgments can be regimented using a sentence operator: “When  $\phi$ , it’s like when  $\psi$ ”.<sup>7</sup> Thus we can say, for example, that when there exists a donkey, it’s

---

<sup>6</sup>Better: it does not require that there are facts in the fundamental sense of ‘there are’; see section 9.3.

<sup>7</sup>Jason Turner suggested this regimentation, pointing out that we can think of the ‘it’ as being like the ‘it’ in ‘It is raining.’ (The suggestion is not that this sentence operator is

like when there exists an electron. I don't claim that we already speak this way (though it's not clear to me that we don't; consider "When it sleets, it's like when it snows"). What I do claim is that there are coherent judgments that can be thus expressed, and that they're similar enough to ordinary ontic similarity judgments to be thought of as being in the same species.

### 6.3 No entities

We have already touched on a second source of resistance to going beyond the predicate. Armstrong/Lewis sparseness—the most familiar game in town—is “entity-based”. For Lewis, in order to evaluate whether a predicate carves at the joints we must look at a certain entity—the set of the predicate's actual and possible instances—and ask whether the entity is natural. And for Armstrong, we must ask whether a certain entity exists—a universal corresponding to that predicate. To extend their strategy to expressions like quantifiers or sentential operators, we would need to identify corresponding entities. But there seem to be no such entities.

One might reply that the entities do exist after all. In the usual model theory of first-order languages, quantifiers and sentential connectives are taken to be syncategorematic: the definition of truth in a model fixes their contributions to truth-conditions without appealing to entities associated with them. But one can instead associate entities with these expressions as their semantic values. One can treat quantifiers as denoting second-order properties of (or relations between) properties (Montague, 1973), and one can treat sentential connectives as denoting truth functions or relations between propositions.<sup>8</sup> And then joint-carving for quantifiers and sentential connectives can be treated in Armstrong and Lewis's way, as turning on the existence or nature of denoted entities.

But this does not get to the heart of the issue—at least, not if we are looking for an account of what structure *is* (please read with the appropriate cadence), as opposed to merely seeking a systematic way to talk about structure. For the treatment of quantifiers as expressing second-order properties or relations, however appropriate in linguistic theory, does not ring true at a metaphysical level. Let  $\exists^2$  be the second-order property of *having at least one instance*. That

---

fundamental.)

<sup>8</sup>Indeed, one can treat all expressions as standing for entities, as in categorial grammars (Gamut, 1991, chapter 4), although then there is a meaningful syntactic operation—concatenation—that stands for no entity.

there exist cows does not seem to be a second-order fact. It surely concerns the concrete world directly, rather than through abstract intermediary entities like  $\exists^2$ . But if quantificational facts are not about  $\exists^2$ , then surely facts about quantificational structure are not about  $\exists^2$  either.

Similarly, it may be appropriate in linguistics to treat ‘and’ as a relation between propositions; but metaphysically speaking, the fact that I am human and I am typing surely concerns neither propositions nor relations between them. I’d even like to say the same about predicates. Even if linguistics is right to associate ‘is human’ with a semantic value, the ultimate metaphysics of my being human has nothing to do with this semantic value.<sup>9</sup> And if conjunctive and predicational facts don’t involve the semantic values of ‘and’ and predicates, then surely the corresponding facts about structure don’t involve these semantic values either.

The view that structure facts do not concern semantic values is certainly intuitively compelling; but it’s best not to rely solely on intuitive compulsions. Fortunately, there is a systematic consideration as well: semantics is, like other special sciences, not fundamental. Our most fundamental level of theorizing should no more recognize distinctively semantic entities and ideology than it should recognize distinctively economic or psychological entities and ideology.<sup>10</sup> This is not to say that the statements of semantics are untrue, only that they are not fundamental (see section 7.8). But if semantics isn’t fundamental, the facts about carving at the joints can’t fundamentally involve semantic entities.<sup>11</sup>

It’s a mistake, then, to think of structure as concerning semantic values. For similar reasons, it would be a mistake to think of structure as concerning linguistic items. The fact (if it is a fact) that ‘is negatively charged’ carves at the joints isn’t in the first instance a fact about the predicate ‘is negatively charged’. The fact is simply a fact about the concrete, nonlinguistic world—about its “charge aspect”, so to speak. Likewise, the fact (if it is a fact) that the first-order quantifiers carve at the joints isn’t a fact about the linguistic items ‘there is’ and ‘for all’. It’s a fact about the world—specifically, its quantificational aspect.

---

<sup>9</sup>See Melia (1995, 2000).

<sup>10</sup>By “distinctively semantic ideology” I do not mean metasemantic ideology like ‘refers’, which relates semantic entities to human populations (though such ideology is surely not fundamental either). I mean, rather, ideology that gives the intrinsic structure of the domain of semantic entities (for instance: notions of conjunction, entailment, and the like, over propositions; or a fundamental predicate functor ‘the property of  $\phi$ -ing’).

<sup>11</sup>More needs to be said about this argument; see chapter 13 for a fuller discussion.

(Not to reify aspects.)

The no-entities worry can seem inescapable given a certain regimentation of structure-talk, which takes the core locution to be a predicate. In Lewis, for example, the core locution is the predicate ‘is natural’. But predicates must be ascribed to *entities*; for Lewis, the entities were the semantic values of predicates; this then leads to taking the facts about structure to involve semantic values.<sup>12</sup> (And it’s no better to take the core locution to be a predicate of linguistic items.) This suggests a way to answer the no-entities worry: introduce a regimentation in which the core locution is not a predicate, so that we can talk about structure without bringing in entities of any sort.<sup>13</sup>

We want a locution, call it “ $\mathcal{S}$ ”, with which to make statements about structure. What should its grammar be? Here we face an obstacle. Since we are going beyond the predicate,  $\mathcal{S}$  must somehow combine with expressions  $\alpha$  of arbitrary grammatical category—with quantifiers, sentential operators, and so on, as well as predicates—to form sentences. But what kind of expression has this sort of grammar? (The attraction of the predicate regimentation is that it avoids this problem: we first convert each such  $\alpha$  into a singular term,  $t_\alpha$ , to which a predicate for structure may be applied. But what would  $t_\alpha$  be? The only possibility seems to be a quotation name of  $\alpha$  itself, or a name of  $\alpha$ ’s semantic value; but then we’d be back to treating the facts about structure as involving linguistic items or semantic values.)

We might overcome this obstacle by taking  $\mathcal{S}$  to have a very flexible grammar, so that it attaches *directly* to  $\alpha$ , regardless of  $\alpha$ ’s grammatical category. I don’t know of any natural language expressions with this sort of grammar, but I don’t see that as a problem. Some philosophers think that we can understand only what can be defined using the pre-existing resources of natural language; but this stultifying doctrine is inadequate to the evident fact of linguistic innovation within science, as well as to the initial emergence of natural language itself. If the inferential role of a novel expression has been made tolerably clear, and if the world contains structure corresponding to the new expression, then surely the introduction of the novel expression has been successful.

So on this way of overcoming the obstacle, the core locution for talking about structure is an “operator”  $\mathcal{S}$ , which can attach to an expression of any

---

<sup>12</sup>Caveat: Lewis takes properties to be sets of possible individuals; and set-theoretic ontology and ideology can be viewed as earning their keep in physics, not semantics, which answers my argument against taking the facts of structure to involve semantic values. See the fuller discussion of this argument in chapter 13.

<sup>13</sup>Thanks to Robbie Williams for discussion of the following issues.

grammatical category  $\alpha$  to form a grammatical sentence  $\mathcal{S}(\alpha)$ . Thus, we can say “ $\mathcal{S}$ (is negatively charged)”, “ $\mathcal{S}(\exists)$ ” (or “ $\mathcal{S}$ (there is)”), “ $\mathcal{S}(\wedge)$ ” (or “ $\mathcal{S}$ (and)”), and so on. Since nothing in English really matches this regimentation, I’ll tend to revert informally to a predicate of linguistic items or abstract entities; I’ll speak of ‘is negatively charged’, ‘and’, and ‘there exists’—or negative charge, conjunction, and existential quantification—as “carving at the joints” or “being fundamental”. But the facts of structure are more faithfully represented using  $\mathcal{S}$ .<sup>14</sup>

To say  $\mathcal{S}$ (and) is not to say something about an alleged object Conjunction. It is not to say anything about any thing at all. It is nevertheless to say something true, something objective, something about reality. Nowhere is it written in stone that all facts must be entity-involving. In Graham Nerlich’s phrase, “realism need not be ontic”.<sup>15</sup> To be sure, the entity-based ideology of predicate logic is simple, beautiful, and well-behaved, and it’s best to stick to it whenever possible. But the realist about structure, it would seem, cannot live by predicate logic alone.

There are hard questions about  $\mathcal{S}$ , thus taken. I said that it can attach to expressions “of any grammatical category”. What, exactly, does that mean?  $\mathcal{S}$  must at least be able to attach to all primitive expressions of the language in question; but what about complex expressions? (As we will see in section 7.13, we need to be able to query complex expressions for joint-carving.)

We might say that  $\mathcal{S}$  can attach to any “grammatical unit”—intuitively, any string that is either a primitive symbol or a complex symbol that is generated at some point by the language’s recursive formation rules. This has certain limitations, depending on the grammar of the fundamental language we’re using to talk about structure. Suppose we want to ask whether a certain conjunctive predicate, the conjunction of predicates  $F$  and  $G$ , carves at the joints. If our language has a predicate functor  $c$  for predicate conjunction then this is straightforward. Our formation rules will include a clause for complex-predicate formation; one of the expressions formed via that clause will

---

<sup>14</sup>In a fundamental language, all and only primitive expressions carve at the joints. Thus  $\lceil \mathcal{S}(\alpha) \rceil$  will be true iff  $\alpha$  is a primitive expression. One might worry that this somehow makes the  $\mathcal{S}$  operator metalinguistic or trivial. But that would be a mistake. First, it doesn’t follow that  $\lceil \mathcal{S}(\alpha) \rceil$  says that  $\alpha$  is a primitive expression. Second, you can’t learn that  $\lceil \mathcal{S}(\alpha) \rceil$  is true simply by observing that  $\alpha$  is a primitive expression; you would also need to know that your language is a fundamental one.

<sup>15</sup>Nerlich (1982, p. 274). See also Putnam (1975e, p. 70); McGinn (1981, 169–70); Yablo (2000, section IX).

be the conjunctive predicate  $F \mathbf{c} G$ ; so we can form the sentence  $\mathcal{S}(F \mathbf{c} G)$ . But suppose (as I suspect is more likely) that our fundamental language is grammatically simpler. Suppose, for example, that its grammar is that of predicate logic (without function symbols), with the addition of the operator  $\mathcal{S}$ . In that case, there simply are no complex predicates, since the only rules of formation are for the grammatical category of sentence. So the only complex strings we can query for structure are sentences (including those with free variables). But then, we can achieve our goal of evaluating the conjunction of  $F$  and  $G$  for structure only indirectly, by forming the open sentence  $\mathcal{S}(Fx \wedge Gx)$ .<sup>16</sup> And the querying becomes even more indirect for, say, complex operators. Suppose, for instance, that the only primitive sentential connectives in the language are  $\sim$  and  $\vee$ , and we want to query for structure a complex expression with the truth table of the material conditional. We can query the various sentences of the form  $\sim\phi \vee \psi$  (for example,  $\sim Fx \vee Gy$ ,  $\sim Ha \vee \forall x \forall y Rxy$ , and so on); but each of these sentences also queries certain other expressions; namely, the expressions occurring inside  $\phi$  and  $\psi$ .

We might just live with the fact that we can't query complexes directly. We could say that the question of whether “the material conditional carves at the joints” is not a matter of the truth of any one sentence in our fundamental language; rather, it emerges from the totality of sentences of the form  $\mathcal{S}(\sim\phi \vee \psi)$ .<sup>17</sup> Alternatively, we might pursue a different method for overcoming the obstacle to constructing a grammar for  $\mathcal{S}$ .<sup>18</sup> According to this method,  $\mathcal{S}$  no longer has a flexible grammar; now it is a one-place sentence operator. However, we include in our fundamental language a collection of dummy variables. There are dummy variables of every grammatical category: individual dummy variables, sentential dummy variables, predicate dummy variables, sentence-operator dummy variables, and so on. Dummy variables are not bindable. Their purpose is to combine with other expressions to form complete sentences, so that the sentence operator  $\mathcal{S}$  may then be applied. The expressions other than dummy variables in such sentences are those that are queried for carving at the joints. So we no longer query sub-sentential primitive

---

<sup>16</sup>And even this does not fully isolate the complex predicate, since one is also querying the variables.

<sup>17</sup>We might introduce a nonfundamental language with grammatical resources to directly query the material conditional. A metaphysical semantics (see section 7.4) for such a language might assign to the query the truth-condition that each sentence of the fundamental language with the form  $\mathcal{S}(\sim\phi \vee \psi)$  be true.

<sup>18</sup>Compare Sider (2009, section 8).

expressions by directly attaching  $\mathcal{S}$  to them; instead, we attach those primitive expressions to appropriate dummy variables to obtain a sentence, and then attach  $\mathcal{S}$  to that sentence. For example, we query the predicate  $G$ , the name  $a$ , the quantifier  $\exists$ , and the operator  $\square$ , with the following sentences, respectively (dummy variables are in sans serif):

$$\mathcal{S}(Gx)$$

$$\mathcal{S}(Fa)$$

$$\mathcal{S}(\exists xP)$$

$$\mathcal{S}(\square P).$$

Complex expressions are now straightforward to query. We can query the material conditional, for example, using the sentence  $\mathcal{S}(\sim P \vee Q)$ .

I have argued that we should not think of judgments about structure as concerning entities. Judgments about structure concern ideology, not ontology. Let us close with a discussion of an opposing, ontic approach.<sup>19</sup>

This opposing approach is modeled on Armstrong, rather than Lewis. Lewis's approach to regimenting talk of structure was to posit an abundant group of entities (for him, sets of 'tuples of possibilia) and regard only some of them as being structural (natural). The Armstrongian approach, by contrast, does away with the abundant entities, and posits only a sparse group of entities. Given a sparse entity, there is no further question of whether it carves at the joints; sparse entities automatically carve at the joints, so to speak. But given a *linguistic* entity, there is a further question of whether it stands for a sparse entity. The sparse entities, for Armstrong, were universals; thus on his view, a predicate carves at the joints iff it stands for some universal.

The sparse entities of this approach are not to be thought of as semantic values. (After all, most meaningful expressions do not stand for sparse entities.) Thus the approach is immune to the objection that semantics is nonfundamental.

Call this approach ontologism, since it insists that fundamental metaphysical commitments be ontic. Distinguish two forms: methodological and metaphysical. The defender of methodological ontologism says that his opponents are making some sort of methodological or conceptual mistake. It's somehow

---

<sup>19</sup>Thanks to Jonathan Schaffer for discussion here, though the opposing approach should not be attributed to him.

conceptually confused to think of a fundamental metaphysics as being given by anything other than a list of entities. The defender of metaphysical ontologism, on the other hand, puts forward her position as a substantive hypothesis about the nature of fundamental reality.

Methodological ontologism seems hopeless. Earlier in this section I put forward a conception, using the operator  $\mathcal{S}$ , of what non-ontic claims of fundamental metaphysics might look like. It's hard to see how this position is conceptually or methodologically confused. Further, it's tempting to view the defender of methodological ontologism as arbitrarily privileging his own ideology. *His* ideology includes the quantifiers and a special predicate that singles out the sparse entities; *my* ideology includes  $\mathcal{S}$  and various other expressions; what makes his ideology the sole intelligible vehicle for giving fundamental descriptions of reality?

Quine's terminology—to return to my rant—has perhaps contributed to methodological ontologism appearing more plausible than it really is. 'Ideology' suggests a purely arbitrary, conceptual matter; ideology and ontology are supposed to be exhaustive; thus, the only nonconceptual question is that of one's ontology. Quine's terminology is so ingrained that this reasoning can seem built into the very concept of metaphysics.<sup>20</sup> Simply appreciating the possibility of an alternative, of ontology-free but nevertheless worldly metaphysics, should break this spell.

Metaphysical ontologism is a much more likely proposition. The attitude here is that the alternative is inferior, not unthinkable; the best metaphysics of fundamentality is entity-based.

Ontologism could, in principle, sustain much of the project of this book. For in many cases, talk of joint-carving ideology can be replaced with talk of sparse entities. In a discussion of whether causal disputes are substantive, for example, instead of asking whether  $\mathcal{S}(\text{causes})$ , the defender of ontologism could ask whether there exists a causal universal. But problems begin when the defender of metaphysical ontologism tries to go beyond the predicate. Whenever she wants to speak of joint-carving she faces a choice: either posit a corresponding sparse entity, or else do not speak of joint-carving at all.

For example, if she wanted to say that modal operators carve at the joints, she would have to posit an appropriate entity and claim that the modal facts ultimately boil down to the facts about this entity. Candidates for the entity

---

<sup>20</sup>In fact, one can discern a usage of 'ontology'—particularly prevalent in the philosophy of physics literature—as a synonym for 'metaphysics'.

might include i) a modal universal that is instantiated by other universals;<sup>21</sup> ii) a modal universal that is instantiated by propositions (which would themselves need to be reified—a very unArmstrongian move); or iii) a modal entity of a different sort, call it a “modal monad”, that plays a role in instantiation: the necessary possession of a universal  $U$  amounts to the instantiation of  $U$  by a particular *with respect to the modal monad*, whereas the merely contingent possession of  $U$  is the instantiation *simpliciter* of  $U$  by the particular. If she is unwilling to say any of these things—to underwrite the claim of modal structure by positing a sparse modal entity—then she cannot speak of modal structure at all.

I myself do not believe in modal structure (chapter 12), so I don’t regard this limitation as unwelcome. But consider, next, the case of quantification. As we saw, it is artificial to take a quantifier, in one’s ultimate metaphysics, as standing for an entity. So the defender of ontologism must choose between an artificial metaphysics of quantification and forgoing talk of joint-carving for quantifiers. And the latter would be a real limitation: as we will see in chapter 9, the thesis that quantifiers carve at the joints is the best way to defend the substantivity of ontological questions. This limitation is the chief problem with metaphysical ontologism.

The previous paragraph’s line of thought is especially dialectically effective because the defender of ontologism has a particular reason to regard quantificational questions as being substantive. After all, questions about the nature of fundamentality, for her, turn on whether *there are* appropriate sparse entities. She, of all people, cannot acquiesce to the Carnapian thought that talking as if sparse entities do not exist is just as good as talking as if they do. Thus metaphysical ontologism is “unreflexive”: given its strictures on what can be evaluated for fundamentality, its own apparatus cannot be ratified as fundamental.

Armstrong’s own theory is similarly unreflexive. With good reason, Armstrong refrains from positing a universal of instantiation (1978a, chapter 11, section 1). Thus, although ‘instantiates’ is an ineliminable part of his ideology, he cannot recognize ‘instantiates’ as fundamental, in the way that he recognizes fundamental physical predicates as fundamental. But claims specifying which particulars instantiate which universals are clearly part of Armstrong’s fundamental theory of the world. Merely to list the universals and particulars, without specifying which particulars instantiate which universals, would be a woefully partial description of what, according to him, the world is funda-

---

<sup>21</sup>Compare Armstrong’s (1983) own approach to laws of nature.

mentally like. There is, therefore, overwhelming pressure on Armstrong to recognize, somehow, that talk of instantiation is fundamental; but he is barred from doing so by his ontologism.<sup>22</sup>

The defender of ontologism might, in Wittgensteinian fashion, reply that one cannot *say* the whole truth about fundamentality. The whole truth can only be *shown*, by quantifying over entities, saying that particulars instantiate universals, and so on. If this position is uncomfortable, that is a reason to reject ontologism, and to adopt a broader conception of metaphysical commitment.

## 6.4 Unclear epistemology?

A third reason to fear going beyond the predicate is epistemic: how could we ever tell when attributions of structure to nonpredicates are justified? But in fact, the generalized conception of structure raises absolutely nothing new, epistemically. Questions about how much nonpredicate structure the world contains are substantive metaphysical questions, just like the most substantive questions of first-order metaphysics, and can be addressed in the same way.

Section 2.3 presented a broadly Quinean approach to the epistemology of metaphysics, and to the epistemology of joint-carving in particular. It is reasonable to regard the ideology of our best theory—“best” by the usual criteria for theory choice, such as simplicity—as carving at the joints. This approach is not bound to the predicate. Successful theories justify belief in *all* of their ideological posits. I will, for example, argue in chapter 9 that since our most successful theories employ quantification, we have reason to believe that quantifiers carve at the joints—that quantificational structure is fundamental. This argument is exactly parallel to the argument for fundamental spatiotemporal structure: quantification is no less part of the ideology of fundamental physics than are spatiotemporal notions.

When a conceptual decision has become wholly familiar, it is easy to forget that it is nevertheless a decision. Such are the overwhelmingly successful conceptual decisions of modern logic. Thinking in terms of *and*, *or*, *not*, *all*, *some*, and *identical* has led to great strides in the foundations of logic and mathematics, and so, less directly, in the foundations of all other disciplines as well. The success of these conceptual choices justifies belief in the existence of corresponding structures in the world. Once “ideology” is purged of its

---

<sup>22</sup>To be fair, Armstrong is not really asking the same set of questions as I am, so it is somewhat misleading to describe him as accepting ontologism.

psychological connotations, there is no barrier to recognizing a theory's logical ideology as a coequal part of that theory's portrayal of the world, and thus as being as good a candidate for carving at the joints as the theory's nonlogical vocabulary.

## 6.5 Logical conventionalism

A fourth source of resistance to going beyond the predicate, and in particular to speaking of joint-carving logical notions, is vaguer, more primordial, and (I think) widespread. It is the thought that it is appropriate to evaluate expressions for carving at the joints only when they are “contentful”. *Predicates* are paradigmatically contentful. But logical expressions, on the other hand, are purely “formal”, so the thought goes. They do not describe features of the world, but rather are mere conventional devices. Since logical expressions are not “worldly”, it is inappropriate to speak of the world as containing structure corresponding to those expressions.

This is picture thinking. But behind the picture, I suspect, there lies an identifiable—and mistaken—philosophical doctrine: the doctrine of logical conventionalism.

Actually, what I really think is widespread is not so much an *acceptance* of logical conventionalism as a failure to fully repudiate it. The status in contemporary philosophy of logical conventionalism and the related doctrine of “truth by convention” is curious. On the one hand, few people self-identify as logical conventionalists. If pressed on why not, I suppose most would gesture at Quine’s famous critique in “Truth by Convention”. But on the other hand, the picture thinking described above really does have staying power, which would be explained by latent logical conventionalism. Moreover, the *language* of truth by convention persists: one still hears the phrase “true purely by virtue of meaning”, logical truths are still described as being “trivial” or “empty” (and are thus thought of as being epistemically unproblematic), and so on.<sup>23</sup>

Against logical conventionalism, I uphold Russell’s (1919, p. 169) diametrically opposed position: “logic is concerned with the real world just as truly as

---

<sup>23</sup>Ironically, one source of the lingering conventionalist strain may be a backlash against Quine’s critique of analyticity. Quine’s overarching critique contained, as a part, the empirical assertion that there are no facts of meaning (1951b). The manifest inadequacy of this view may have led to a failure to appreciate the most powerful part of the critique, namely, his attack on truth by convention (1936; 1960a). See also Boghossian (1997).

zoology, though with its more abstract and general features". Evaluating logical expressions for joint-carving is therefore not different in kind from evaluating any other expressions for joint-carving.

All I have to offer in support of Russellian realism about logic is a critique of conventionalism; discussion of intermediate positions remains a lacuna.

Logical conventionalism originated in the "linguistic theory of the a priori", popularized by A. J. Ayer in *Language, Truth and Logic* (1936):

Like Hume, I divide all genuine propositions into two classes: those which, in his terminology, concern "relations of ideas," and those which concern "matters of fact." The former class comprises the *a priori* propositions of logic and pure mathematics, and these I allow to be necessary and certain only because they are analytic. That is, I maintain that the reason why these propositions cannot be confuted in experience is that they do not make any assertion about the empirical world, but simply record our determination to use symbols in a certain fashion. (p. 31)

A proposition (sentence) is analytic, Ayer goes on to say, "when its validity depends solely on the definitions of the symbols it contains ...". Analytic propositions can be known a priori because they are "devoid of factual content" (p. 78), because they merely "record our determination to use words in a certain fashion". Something like this view was once widely held, by logical positivists, Wittgensteinians, and ordinary language philosophers.<sup>24</sup> The core of the view is that an analytic truth, for instance the truth that all horses are horses, is true purely by virtue of linguistic conventions. By adopting certain rules governing the use of logical words like 'all', language users somehow make 'all horses are horses' true.

Quine famously objected to the doctrine of truth by convention as follows. (B) is allegedly true by convention:

(B) An object is a bachelor iff it is an unmarried man.

Pretend that, as a matter of convention, 'bachelor' means the same as 'unmarried man'. Thus, (B) means the same as the following logical truth:

(A) An object is an unmarried man iff it is an unmarried man.

---

<sup>24</sup>Conventionalists include Ayer, Britton (1947), Carnap (1937, §69; 1950), and Malcolm (1940). Pap (1958, chapter 7) contains a thorough (critical) discussion of conventionalism. See also Lewy (1976, especially chapter 5), and Boghossian (1997).

The introduction of the convention governing ‘bachelor’ therefore makes (B) have the same truth-value as (A); but this on its own, Quine pointed out, does not make (B) true. The truth of (B) requires the “prior” truth of (A). As Quine says, “... definitions are available only for transforming truths, not for founding them” (1936, p. 81).

(B) *would* be rendered true by convention if the logical truth (A) itself were in some sense true by convention. But as Quine went on to argue, logical truths do not in any interesting sense owe their truth to conventions. In particular, he considers the thought that we can legislate a logical truth  $T$  by proclaiming “Let  $T$  be true.” He points out a problem (one that arises even before the pressing question of how such proclamations are supposed to do the trick). A would-be legislator of logical truth cannot individually legislate each logical truth, for there are infinitely many of them. He might think to make general legislations to the effect that every sentence of a certain form is to be true, such as:

- (I) Let every instance of the following schema be true: ‘If  $\phi$  then  $\phi$ ’.

But any such legislation results in truths of a single fixed logical form, whereas logical truths can take infinitely many forms. So, he might think, a second type of general legislation is required, which specifies that *if* certain statements are true, then others are to be true as well. Here is an example:

- (II) If a statement  $\phi$  and a statement ‘If  $\phi$  then  $\psi$ ’ are true, then let  $\psi$  be true as well.

The conventionalist’s hope, as Quine imagines, is to make true all the truths of propositional logic using legislations of types (I) and (II). For there exist complete axiomatizations of propositional logic with finitely many axiom schemas in which the only rule of inference is modus ponens; each axiom schema could be legislated in style (I), and modus ponens could be legislated into effect by proclaiming (II). One could then go on to legislate the truths of (first-order) predicate logic in a similar fashion.

According to Quine, the problem for conventionalism thus understood is that logic is needed to apply the conventions, and cannot therefore be grounded in the conventions. Suppose that statements  $\phi$  and ‘if  $\phi$  then  $\psi$ ’ have been legislated to be true by legislations of sort (I). (II) now says that *if* these sentences are true, *then*  $\psi$  is to be true as well. To derive from this the result that  $\psi$  is indeed true, we must perform modus ponens—we must use logic. But logic is exactly what the legislations are supposed to ground.

For various reasons, it seems to me that Quine's objection—that logic will be needed to legislate the infinity of logical truths—does not get to the heart of what is metaphysically problematic about conventionalism.<sup>25</sup> Imagine a finitary conventionalist, who tries to introduce conventional truth in a language whose set of well-formed formulas is finite. Or imagine a conventionalist with an infinitary mind, who can legislate each of the infinitely many logical truths individually. Logic would surely not be true by convention even in these cases, but in neither case does Quine's objection apply.

Moreover, the conventionalist might reply to Quine that legislations of form (II) are *conditional legislations* rather than *legislations of conditionals*. (Compare the distinction between conditional probability and the probability of a conditional, or between conditional obligation and an obligation to see to the truth of a conditional.) Quine's objection is actually pretty elusive, but one way of taking it is as follows:

After legislations of type (I) are made, for certain sentences  $\phi$  and  $\psi$ , both sentence  $\phi$  and the sentence ' $\text{if } \phi \text{ then } \psi$ ' are true by convention; and after legislation (II) is made, the following conditional sentence is also true by convention: ' $\text{If } \phi \text{ and } \text{if } \phi \text{ then } \psi \text{ are both true, then } \psi \text{ is also true}$ '. But we cannot pass from the fact that these three sentences are true by convention to the conclusion that  $\psi$  is true by convention unless we make the further assumption that the truths by convention are closed under modus ponens (and also under conjunction-introduction, truth-introduction, and truth-elimination). Since that further assumption is precisely what (II) was supposed to accomplish, (II) is ineffective.

This objection, notice, assumes that the effect of legislation (II) is that a certain conditional sentence (namely ' $\text{If } \phi \text{ and } \text{if } \phi \text{ then } \psi \text{ are both true, then } \psi \text{ is also true}$ ') is true by convention. But the conventionalist might reply that its effect is instead that it be the case that if  $\phi$  and ' $\text{If } \phi \text{ then } \psi$ ' are both true by convention, then  $\psi$  is also true by convention. The conventionalist might reply, that is, that his claim all along was that (II) results directly in the set of truths by convention being closed under modus ponens, and not that it results in a conditional sentence corresponding to modus ponens being true by convention.

---

<sup>25</sup>It does a better job of showing that conventionalism cannot epistemically justify logic: we already need to be justified in using logic before we can gain the justification that the conventionalist is trying to supply.

(II) is not an unconditional legislation that a certain sentence (a conditional sentence) be true by convention. It is rather an irreducibly conditional legislation, which results in its being the case that if certain sentences are true by convention, then so is a certain further sentence. The latter sentence is legislated conditionally on the former sentences being legislated. Quine in effect grants the conventionalist for the sake of argument that the words pronounced in (I), “Let any statement of the following form be true ...” have their desired effect. But the words are not magic: something about the pattern of beliefs and dispositions in the linguistic community that results from the pronouncement of those words is what allegedly does the trick: certain sentences individually become true by convention. If Quine is willing to grant that the words in (I) have this effect, why is he not also willing to grant that different words, the words in (II), which result in a different pattern of beliefs and dispositions, have a different effect, that of a conditional legislation?

Quine’s argument does not go far enough. An adequate critique must challenge the very idea of something’s being “true by convention”. Even an infinite mind, or a conventionalist with only finite aspirations, or a conditional legislator, could not make the logical truths, or any other sentence for that matter, true by convention (unless the sentence is *about* conventions). The components of this critique are not new, but are nevertheless worth repeating.

Part of the critique consists in pointing out that it is no easy task for the conventionalist to specify an appropriate sense in which logical truths are “true by convention”. There is a mundane sense in which *all* true sentences are partly true because of convention, since all sentences exhibit symbol-conventionality (section 4.3). Even a synthetic sentence like ‘Snow is white’ is true partly because of its meaning; it would not have been true if it had meant that grass is green. Of course, in order for ‘Snow is white’ to be true, the world must also cooperate: the world must really be as the sentence says. Snow must really be white. So, it might be thought, the conventionalist could claim that the requirement of worldly cooperation is not present for logical truths. But on the face of it, this is wrong. By convention we make it the case that the sentence ‘If it is raining then it is raining’ means that if it is raining then it is raining; but in order for the sentence to be true, the world must also cooperate; the world must really be as the sentence says. It must really be that if it is raining then it is raining. It is easy to overlook the requirement of cooperation in this case because it is so obvious that if it is raining then it is raining. But no sense has yet been given to the idea that the requirement is not present.

A conventionalist might reply that ‘If it is raining then it is raining’ “auto-

matically” becomes true upon being endowed with its meaning; that is the sense in which there is no further requirement that the world cooperate. But what does ‘automatically’ mean here? It could be understood in terms of necessity: it is necessarily true that if ‘if’ and ‘then’ mean what they do then ‘If it is raining then it is raining’ is true. Thus understood the claim is correct, but it does not secure a truth-making role for the convention. The mere fact that it is necessarily true that it is raining if it is raining ensures that ‘If it is raining then it is raining’ is “automatically”—in the current sense of ‘automatically’—true once it has been given its meaning. Conventionalism thus understood says little more than that logical and other analytic truths are necessary; nothing is left of the intuitive idea of their truth being grounded in conventions.

We are still in search of an appropriate sense in which logic is true by convention. Ayer’s claim that analytic truths “simply record our determination to use symbols in a certain fashion” is unhelpful. It suggests that analytic truths make statements *about* linguistic conventions. But this is a nonstarter; statements about linguistic conventions are contingent, whereas the statements made by typical analytic sentences are necessary (Broad (1936, p. 107), Lewy (1976, p. 9)). Ayer’s claim that analytic truths “say nothing about the empirical world” is similarly unhelpful: it is hard to attach any sense to it that advances his cause. ‘If it is raining then it is raining’ *seems* to say something about the empirical world: that the empirical world contains rain if it contains rain. Of course, the thing it says is a logical truth. We might define “about the empirical world” to exclude logical truths, but what would be the point?—the claim that logical truths “say nothing about the empirical world” could then play no role in explaining the epistemology or metaphysics of logical truth.

To further reinforce the difficulty of finding an appropriate sense of ‘true by convention’, consider that the phrase is intended to indicate an intimate sort of *dependence* of truth on convention. But what sort, exactly? The conventionalist will surely deny counterfactual or temporal dependence, at least of the sort that would imply absurd statements like the following:

Before we introduced our conventions, not all green things were green.

If we had introduced no conventions, not all green things would have been green.

Of course, metalinguistic counterfactual and temporal statements such as the following are unproblematic:

Before we introduced our conventions, the sentence ‘all green things are green’ was not used to express a truth.

If we had introduced no conventions, the sentence ‘all green things are green’ would not have been used to express a truth.

But *all* truths depend on conventions in this metalinguistic way; before we introduced our conventions, the sentence ‘Snow is white’ was not used to express a truth. It remains unclear just what sort of dependence of truth upon conventions is supposed to be distinctive of conventionalism.

Here are two further failed attempts to understand what the defender of truth by convention has in mind. Return to the would-be truth-legislator, who says “Let every sentence of the form ‘If  $\phi$  then  $\phi$ ’ be true.” What is this performance intended to accomplish? On the one hand, the legislator could be resolving to use the word ‘true’ in a new way; he could be listing the sentences to which this new term ‘true’ applies. But this obviously isn’t what the conventionalist wants. On the other hand, the legislator could be constraining the intended meaning for conditional constructions. He could be placing a necessary condition on the function from sentences to the propositions that they mean: this function must assign a true proposition to each sentence of the form ‘If  $\phi$  then  $\phi$ ’. Any function that violates this constraint, the legislator is saying, is not the means-in-English function. This, too, is not what the conventionalist wants, for it assumes an antecedent notion of propositional truth that has not been shown to depend in any way on convention.

This last point bears emphasis. We should all agree that one way to constrain the meaning of an expression,  $E$ , is to stipulate that  $E$  be interpreted so that certain sentences containing  $E$  turn out true, or that certain inferences involving  $E$  be truth-preserving. It can seem that such stipulations create truth, or truth-preservation, on their own. But this is not the case, as was illustrated by Arthur Prior (1960) in dramatic fashion. Prior imagined the introduction of a new sentential connective ‘tonk’, stipulated to obey a disjunction-like introduction rule “From  $\phi$  infer  $\phi$ -tonk- $\psi$ ”, as well as a conjunction-like elimination rule “From  $\phi$ -tonk- $\psi$  infer  $\psi$ .” The stipulations do not result in the two rules being truth-preserving, for the rules would allow us to infer any statement  $\psi$  from any other statement  $\phi$  (first infer  $\phi$ -tonk- $\psi$  from  $\phi$  using the introduction rule, and then infer  $\psi$  using the elimination rule). ‘Tonk’ is stipulated to stand for a meaning that obeys the two rules; but there simply is no such meaning; ‘tonk’ cannot be interpreted so as to obey the rules. Now, we do not believe the usual logical connectives to be like ‘tonk’. When we stipulate that conditional

sentences are to be so interpreted that every sentence of the form ‘If  $\phi$  then  $\phi$ ’ is true, or when we stipulate that ‘and’ and ‘or’ obey their usual introduction and elimination rules, we believe that these expressions *can* be understood so as to obey the stipulations. But the case of ‘tonk’ shows that the stipulations do not, on their own, create the truth, or truth-preservation.

The critique so far has not produced an argument against conventionalism; it has merely cleared away obstacles to understanding, by enumerating various things that conventionalists cannot mean by ‘true by convention’. Now, this sort of clarification can be effective. For some, conventionalism will lose whatever appeal it had, once the scales fall from their eyes. Nevertheless, direct arguments against conventionalism would be welcome.

It is difficult to argue against a doctrine that has not been clearly articulated. But what we can do is formulate the doctrine in schematic terms, and then argue that so long as those schematic terms behave in a certain way, the doctrine must be false. I will give two arguments of this form, assuming the following schematic understanding of the doctrine of truth by convention: “We can *legislate-true* the truths of logic.”

The first argument assumes that sentences that are *about* certain parts of the world cannot be legislated-true. These are the parts of the world that I cannot affect simply by wishing or pronouncing or legislating. I cannot, for example, make true the sentence ‘My computer monitor has been thrown out the window’ by wishing or pronouncing or legislating; I must defenestrate the monitor myself, or pay or incite someone else to do it. Indeed, given my lack of magical powers, the *only* statements that I can affect by mere wishing or pronouncing or legislating would seem to be sentences *about* conventions or related matters, such as which noises I make. We nonmagical humans can legislate-true such sentences because they are about *us*. The first argument, then, is this. Sentence (O) is not about us:<sup>26</sup>

(O) Either it is raining or it is not raining.

Since the only statements that we can legislate-true are those that are about us, we cannot legislate-true the logical truth (O).

Talk of “aboutness” is admittedly slippery. Now, all the first argument needs is that there is *a* sense of ‘about’ on which (O) is not about us, and on which

---

<sup>26</sup>This could be regarded as a stand-alone premise; or it could be supported thus: ‘It is raining’ is not about us; ‘It is not raining’ is not about us; the statements that are not about us are closed under disjunction. Supporting the premise in this way would draw this first argument closer to the second.

only sentences about us can be legislated-true. Still, an argument that makes no appeal at all to aboutness may be welcome.

The second argument fits the bill. What it assumes about the schematic notion of legislating-true is that i) I cannot legislate-true ‘It is raining’; and ii) I cannot legislate-true ‘It is not raining’; and iii) if I cannot legislate-true  $\phi$ , nor can I legislate-true  $\psi$ , then I cannot legislate-true the disjunction  $\lceil\phi \text{ or } \psi\rceil$ . In defense of iii): a disjunction states simply that one or the other of its disjuncts holds; to legislate-true a disjunction one would need to legislate-true one of its disjuncts. (To know (believe, promise, ...) a disjunction, one needn’t know (believe ...) one of its disjuncts; but this needn’t undermine iii), which is specific to legislating-true.) Given premises i)–iii), I cannot legislate-true (O).

It is open, of course, for the defender of truth by convention to supply a notion of legislating-true on which the argument’s premises are false. The challenge, though, is that the premises seem correct given an *intuitive* understanding of “legislate-true”.

# Chapter 7

## Questions

**F**RIENDS OF FUNDAMENTALITY face some abstract questions about its nature. My way of thinking about fundamentality—in terms of structure—is distinctive in large part because of how I answer the questions. My answers: the fundamental is complete, pure, subpropositional, absolute, determinate, and fundamental.

### 7.1 Complete?

It is natural to assume that the fundamental must be “complete”, that the fundamental must in some sense be responsible for everything.

Completeness seems definitive of fundamentality. It would be a nonstarter to say that the fundamental consists solely of one electron: thus conceived the fundamental could not account for the vast complexity of the world we experience.

A preliminary formulation of completeness might run as follows: *every non-fundamental truth holds in virtue of some fundamental truth.*<sup>1</sup> But the exact content of this formulation is far from clear. What do ‘in virtue of’ and ‘fundamental truth’ mean here? There are subtle issues about how to understand these notions in terms of my official notion of structure. I want to postpone discussion of those subtleties, however; so for now let us leave completeness stated in this intuitive way. A fundamental truth (or fact), intuitively, is a metaphysically basic or rock-bottom truth (fact). Facts about the positions of subatomic particles

---

<sup>1</sup>A refined principle would allow a nonfundamental truth to hold in virtue of multiple fundamental truths taken collectively.

would be, on most views, fundamental facts, whereas the fact that some people smile when they eat candy would presumably not be. ‘In virtue of’, intuitively, stands for the relationship whereby the fundamental facts underwrite or give rise to all other facts. The fact that some people smile when they eat candy holds in virtue, perhaps, of certain facts about the states of subatomic particles (or, given a less materialistic outlook: in virtue of these subatomic facts plus certain fundamental mental facts).

Though I will be leaving ‘in virtue of’ at an intuitive level for now, I should say up front that it is *not* to be understood in terms of modality, truthmaking, or fact-identity. Thus I reject these conceptions of completeness:

“All truths are necessitated by (or, supervene on) a fundamental description of the world.”

“Every truth has a fundamental truthmaker.”

“Every fact is identical to a fundamental fact.”

The modal gloss imposes no meaningful requirement of completeness for necessary truths, the truthmaking gloss requires a commitment to truthmaking, and the fact-identity gloss requires a commitment to facts, individuated in an appropriate way. (I’ll have more to say about truthmaking and facts in chapter 8.) A less objectionable way to cash out “in virtue of” would appeal to the—currently very popular—notion of ground: “all truths are grounded in fundamental truths”.<sup>2</sup> To say that the existence of a city is grounded in certain facts about subatomic particles is to say that the latter facts produce or account for or explain the existence of a city, in a distinctively metaphysical way. Although ground implies necessitation, necessitation is insufficient for ground. (Examples like the following are often given: snow’s being white does not ground its being the case that either grass is green or grass is not green, even though it’s necessary that if snow is white then grass is either green or not green; what grounds the truth of grass’s being either green or not green is its true disjunct: grass’s being green.) Thus the grounding approach lays down a meaningful requirement of completeness for necessary truths.<sup>3</sup> Further, properly understood, speaking in terms of ground requires no commitment

<sup>2</sup>On grounding see Fine (2001, 2010, 2011); Rosen (2010); Schaffer (2009a); Schnieder (2011); and the papers in Correia and Schnieder (2012).

<sup>3</sup>This is just one way in which ground improves on the coarser-grained notion of modality; see especially Fine (2001).

to truthmaking, or to facts, propositions, or any other abstract entities. I will indeed sometimes speak of the grounding of propositions or facts (and will likewise speak of in-virtue-of relations amongst facts and propositions), but such talk is dispensable: one can always construe ‘ground’ (and related locutions) as a sentence operator: “That  $\phi$  grounds its being the case that  $\psi$ ” (Fine, 2001). Later on I will criticize the ground-theoretic interpretation of completeness, and propose a different gloss of “in virtue of”. But the ground-theoretic gloss is much closer to my own than are those in terms of modality, truthmaking, or fact-identity, so if it is familiar, it can serve as a working heuristic: “Every nonfundamental truth is grounded by some fundamental truth.”

## 7.2 Pure?

There is a second assumption about structure that I think we ought to make—what I call “purity”: *fundamental truths involve only fundamental notions*. When God was creating the world, she was not required to think in terms of nonfundamental notions like city, smile, or candy.

As with completeness, there are subtleties about how exactly to understand purity in my preferred terms. “Fundamental notion” is easy (it means “carves at the joints”) but “fundamental truth” remains to be explained. Let us postpone discussion of these subtleties just a little longer. (But do notice that purity concerns two distinct concepts of fundamentality: the concept of a fundamental *notion*, and the concept of a fundamental *truth*.)

Suppose someone claimed that even though cityhood is a nonfundamental notion, in order to tell the complete story of the world there is no way to avoid bringing in the notion of a city—certain facts involving cityhood are rock-bottom. This is the sort of view that purity says we should reject. This might seem obvious and uncontroversial. But in fact, purity has some very striking consequences.

Here is a truth: there exists a city. Since the notion of a city is not fundamental, purity says that this truth is not fundamental. No surprises so far. Completeness then says that this truth holds in virtue of some fundamental truth  $T$ —perhaps some truth of microphysics. So we have:

- (i) There is a city in virtue of the fact that  $T$ .

Still no surprises.

But now consider (1) itself. Just like ‘There are cities’, (1) is a truth involving the notion of a city. And so, given purity, it cannot be a fundamental truth. And so, given completeness, it must itself hold in virtue of some fundamental truth.

Now, I accept this consequence (given the way I will eventually understand “in virtue of”). (1) is not itself fundamental (nor is any other in-virtue-of truth, in my view). So (1) must itself hold in virtue of other truths. But this is a nontrivial claim; and it is a claim that some people are going to want to resist. A certain sort of primitivist about in-virtue-of, for example, will refuse to explain truths like (1) in other terms. Purity stands in the way of this sort of primitivism; it requires facts about the *relationship* between the fundamental and the nonfundamental to be themselves nonfundamental. Thus purity brings a heavy explanatory burden: it requires there to be facts in virtue of which in-virtue-of-facts hold. But this is a burden we ought to shoulder. The rock-bottom story of the world ought not to mention cityhood at all, not even in sentences like (1). The primitivist about in-virtue-of who opposes this is in an awkward position. On the one hand she must surely acknowledge that *most* truths involving cityhood—truths such as “There is a city”, “Philadelphia is a city”, “Candy can be purchased in most cities”, and so on—are not fundamental; and she must surely feel the force of the thought that this is in some sense because such truths involve the nonfundamental notion of being a city. But then why the special exception for truths like (1)? Admitting that (1) is a fundamental truth would drag the notion of cityhood itself into the realm of the fundamental, since the admission concedes that the fundamental story of the world cannot be told without bringing in cityhood.

Let’s think a little more about purity, and in particular, how it relates to forms of primitivism. Consider the doctrine of modalism, which I understand as the claim that necessity is a fundamental notion (in my terms, carves at the joints). Now, many modalists would, I think, take this a step further, and say also that modal truths are fundamental truths. But given purity, it cannot be that *all* modal truths are fundamental. The modal truth that it is necessary that all cities are cities, for example, must be nonfundamental given purity, since it involves the nonfundamental notion of cityhood. But then given completeness, “Necessarily, every city is a city” must hold in virtue of some further fundamental truth  $N$ ; and  $N$  cannot, given purity, involve the notion of cityhood. Notice, though, that  $N$  *can* involve the notion of necessity, if modalism is true.  $N$  might, for example, have the form “Necessarily, all  $C$ s are  $C$ s”, where  $C$  involves only fundamental notions. (Think of  $C$  as a “metaphysical definition” of the notion of being a city.)

So what we have learned is this: even if the modal notion of necessity is fundamental, purity prohibits modal truths involving nonfundamental notions from being fundamental. The only fundamental modal truths are an array of “austere” or “pure” modal truths that give the necessary connections amongst fundamental notions. (The array will include necessities of logical truths that contain only fundamental notions—such as “Necessarily, all  $C$ s are  $C$ s”—but it will presumably include further truths; perhaps: “Necessarily, nothing is both negatively and positively charged”, “Necessarily, if  $x$  is more massive than  $y$  and  $y$  is more massive than  $z$  then  $x$  is more massive than  $z$ ”, and the like.)

I myself reject modalism (chapter 12); but I accept other sorts of primitivism for which purity has analogous consequences. Negation, on my view, is a fundamental notion; but since ‘eats’, ‘candy’, and ‘smile’ are nonfundamental notions, purity implies that the truth “It is not the case that something eats candy without smiling” is not a fundamental one. It holds, perhaps, in virtue of some fundamental truth of the form “It is not the case that something  $E$ s some  $Y$  but does not  $S$ ”, where  $E$ ,  $Y$ , and  $S$  are “metaphysical definitions” of ‘eats’, ‘candy’, and ‘smile’, respectively. Likewise, I think that quantifiers are fundamental (they carve at the joints); but given purity, such truths as “There exists a city” are nonfundamental, and hold in virtue of quantificational truths (perhaps of the form “There exists a  $C$ ”) that involve only fundamental notions. As with modality, even if negation and quantification are fundamental *notions*, the only fundamental *facts* involving those notions are pure—they involve those notions in combination only with other fundamental notions.

The issue of purity can be further explored by discussing a particular example—one that, I hope, an opponent of purity will agree is a sort of crucial experiment, on which the issue turns. First some setup. Let  $C_0$  be a predicate that describes New York City at the subatomic level in complete detail (relationally as well as intrinsically). The following is presumably true:

- (2) Necessarily, every  $C_0$  is a city.

(“City-zombies” are impossible.<sup>4</sup>) Finally, assume for the sake of argument (what I do not myself believe) that modalism is true—necessity is a fundamental notion.

Purity implies that (2) is not a fundamental truth, since it contains ‘city’. An opponent of purity, I think, will take this as a good place to draw a line in the sand. She will say that (2) *is* a fundamental truth; the modal connection

---

<sup>4</sup>Cf. Bennett (2006).

between  $C_0$  and cityhood is incapable of further explanation. (This is a good place to draw the line in the sand because the modal connection between macro-predicates like ‘city’ and their micro-realizations is particularly resistant to reduction.<sup>5</sup>) I, on the other hand, accept purity’s consequence that (2) is not fundamental (even under the pretense of modalism). My argument has been simply that the fundamental story of the world ought not to mention cityhood at all, not even within sentences like (1) and (2). (“When God created the world, she did not need to use ‘city’.”) But this can be brought out more vividly.

Think of a sentential operator like ‘necessarily’ as a machine. You feed it a statement (proposition, interpreted sentence), and it spits out a truth-value. The output *true* means that the input statement is necessarily true; the output *false* means that the input statement could have been false. Think of the fundamental facts of the form “Necessarily,  $\phi$ ” and “It’s not the case that: necessarily,  $\phi$ ” as corresponding to the inputs  $\phi$  that the machine is equipped to handle—the inputs for which the machine “knows what to do”. If it’s a fundamental fact that it’s necessary that all electrons are electrons, the machine “knows what to do” with the input ‘all electrons are electrons’; it spits back *true*. The thought in favor of purity is then this: the machine should *not* know what to do with the input ‘Every  $C_0$  is a city’. If the machine did know what to do with this input, it would “know” how to inspect the notion of a city, and figure out what its microphysical modally sufficient conditions are. And this is an inappropriate capacity for the machine to have. When God created the fundamental notion of necessity, she needed to endow it with the fundamental capacity to interact with other fundamental notions (perhaps: all, and, not, electron, and the like), but *not* with notions like city, smile, and candy. (After all, fundamentally speaking there are no such notions.) This is not to deny that (2) is true; it is only to deny that (2) is fundamentally true. Demanding that (2) be fundamentally true demands more of a fundamental notion of necessity than it has to give.

### 7.3 Purity and connection

In the previous two sections I formulated completeness and purity in intuitive terms—“fundamental fact”, “in-virtue-of”—rather than in my own official

---

<sup>5</sup>See Lewis (1986b, pp. 150–7) and section 12.9 of this book.

terms. This was not only for ease of digestion: it was also because the issues are not particular to my own metaphysics of fundamentality. The question of whether facts about the relation between the fundamental and the nonfundamental (like (1)) are themselves fundamental, for example, confronts everyone who takes the notion of fundamentality seriously. But ultimately I want a formulation of completeness and purity in my own terms. The next few sections head in that direction.

What is the relationship between the fundamental and the nonfundamental? In what sense do all fundamental matters “boil down to” or “derive from” or “hold in virtue of” fundamental matters, as completeness says they do? How does the nonfundamental *connect* to the fundamental?

We learned something important about the nature of the connection in section 7.2. As we saw, purity implies that claims like (1) and (2), which describe in-virtue-of and modal connections, respectively, express merely nonfundamental truths. But purity lets us draw a more general conclusion. On *any* conception of the nature of the connection, connecting truths—true statements expressing the distinctive connection between fundamental and nonfundamental matters—are going to, like (1) and (2), involve nonfundamental notions. (This is so by definition; connecting statements relate *nonfundamental* notions to fundamental ones.) So on any conception of the nature of the connection, the connecting truths are going to have to be nonfundamental.

This in turn tells us something about how to attack the question of the relationship between fundamental and nonfundamental: we should *not* attack it using the method of metaphysical posit—by positing a metaphysically fundamental notion (such as in-virtue-of, necessity, or truthmaking) for the connection. The point of using the method of posit would be to claim that the connecting truths involving the posited notion are fundamental truths, thereby obviating the need to say in virtue of what such connecting truths hold. But purity implies that the connecting truths could not be fundamental.

(Although purity tells us that it’s pointless to posit fundamental notions like in-virtue-of, necessity, or truthmaking *in order to attack the question of the connection*, it does not prohibit such posits per se.<sup>6</sup> One could, for example, posit a fundamental notion of in-virtue-of, and claim that in-virtue-of truths involving only fundamental notions—for example, “There exist things that are either negatively or positively charged in virtue of the fact that there exist things that are negatively charged”—are fundamental truths. But such truths are not

---

<sup>6</sup>Thanks to Bruno Whittle for this point.

connecting truths, and do not help with the problem of how the fundamental relates to the nonfundamental. And absent some independent motivation, there would be no reason for the posit.)

We have been focusing on the relationship between the fundamental and the nonfundamental, worrying about whether facts about this relationship are fundamental, and so on. It's a little tempting (though only a little) to dismiss such concerns by saying something like this: "Given the fundamental facts, nonfundamental facts follow *automatically*; they're not *extra* facts, somehow *over and above* the fundamental facts; so there's no need for facts connecting the fundamental to the nonfundamental." But the problem with this thought is pretty transparent. What do the italicized phrases mean here? It's natural to construe them all in terms of in-virtue-of.<sup>7</sup> To say that nonfundamental facts "follow automatically" from, and are neither "extra" relative to nor "above and beyond" the fundamental facts is simply to say that they hold in virtue of fundamental facts. But then it becomes clear that the concerns have not been dismissed at all, only relabeled. The fact that there being cities follows automatically from, etc., a certain fundamental fact is exactly the sort of fact whose status we have been questioning.

A related attempt to dismiss the concerns is a little more tempting: "The fact that there is a city *just is* the fact that there is a C; so there is no need for any fact connecting the 'two'. The relationship between so-called fundamental facts and nonfundamental facts is simply identity (so really we shouldn't call some of them fundamental and others nonfundamental); and the status of identity facts is unproblematic." It's worth spending a little time discussing this, since it illustrates a way of hiding metaphysical commitments that will be important later (section 8.5). The objection's crucial claim is that the relationship between fundamental and nonfundamental is unproblematic because it consists simply of identities like these:

(\*) The fact that there is a city = the fact that there is a C.

Now, I agree that identities are in a sense unproblematic. But notice that the singular terms flanking '=' in (\*) are not names, but are rather complex singular terms, formed using the locution 'the fact that'. This is important. Grammatically, this locution combines with a sentence  $\phi$  to form a singular term, 'the fact that  $\phi$ '. Metaphysically, the locution functions as a connector between whatever ideology is allowed to occur within  $\phi$  and certain entities

---

<sup>7</sup>Similar remarks apply if we construe them in terms of necessitation or truthmaking, etc.

(the facts). The use of this locution marks a serious metaphysical commitment; intuitively, it is this locution, not the identity sign, that both does the work in connecting the fundamental to the nonfundamental and also creates problems with purity. This is clearest if we break (\*) into the following three components (where  $a$  and  $b$  are proper names of facts):

- (i) the fact that there exists a city =  $a$ ;
- (ii) the fact that there exists a  $C = b$ ;
- (iii)  $a = b$ .

When the objector says that identities between facts raise no problems, this is partly right: purity does not conflict with (iii)'s being a fundamental truth.<sup>8</sup> Nor does it conflict with (ii)'s being fundamental (provided our fan of facts regards 'the fact that' as a fundamental notion). But purity does rule out (i)'s being a fundamental truth, since (i) contains 'city'—even if 'the fact that' is taken to be a fundamental notion. Return to the metaphor of the machine: 'the fact that  $\phi$ ' is a machine that takes  $\phi$  as input and picks out the corresponding fact. If (i) expresses a fundamental fact, then the machine knows how to pick out the appropriate fact when  $\phi$  contains 'city'; this is an ability that one of reality's basic building blocks should not have.

## 7.4 Metaphysical semantics

As we saw in the previous section, we cannot attack the problem of the relationship between fundamental and nonfundamental by the method of posit—by positing a fundamental notion for the connection. We should instead take facts about the connection to be nonfundamental facts, which may ultimately be explained in terms of fundamental facts that do not involve a fundamental notion of connection.

This is not to say that we should require ourselves to actually specify what those fundamental facts are, in terms of which facts about connection may be explained, before we are willing to speak of the connection. That would require us to possess a metaphysical reduction of the connection; and metaphysical reductions are quite generally very hard to come by. (More on this

---

<sup>8</sup>So long as facts are taken to exist in the fundamental sense, anyway; otherwise it might be objected that names like  $a$  somehow bring in nonfundamental notions.

methodological issue in section 7.6.) Instead, we should introduce a suitable undefined but nonfundamental notion for the connection. We should explain and clarify the suitable notion as best we can, and there should be reason to believe that it could *in principle* be metaphysically reduced; but we need not ourselves possess a reduction.

What is this suitable notion to be? (“In virtue of” has been a mere placeholder.) Earlier I mentioned reasons not to construe it in terms of necessity, truthtaking, or fact-identity. Might the suitable notion be that of ground, regarded as a nonfundamental notion? This is a close relative of my own proposal; but for reasons to be given in section 7.9, I prefer a slightly different, linguistic approach. The suitable notion, in my view, is that of a *metaphysical semantics*—a certain sort of semantic theory. As we will see, completeness may then be understood roughly as the claim that every language has a metaphysical semantics.

A metaphysical semantics is a semantic theory with two distinctive features. First, meanings are to be given in purely joint-carving terms. For example, if the semantic theory takes the form of a truth-theory, then the truth-conditions must be stated in perfectly joint-carving terms. More on this below.

Second, the explanatory goals differ from those of linguistic semantics. In one way they are more ambitious, and in another, more modest. Metaphysical semantics is more ambitious in that by giving meanings in fundamental terms, it seeks to achieve something not sought by linguistic semantics: to show how what we say fits into fundamental reality. Metaphysical semantics is more modest in that it tries to explain a narrower range of phenomena. The semantic theories of philosophers of language and linguists attempt to explain a broad range of phenomena, psychological and social, as well as narrowly linguistic. According to a traditional conception (largely associated with Frege), meaning plays a broad theoretical role: the meaning of a sentence is conventionally encoded by that sentence, grasped by anyone who understands the sentence, is communicated when the sentence is used; sentence-meanings are the objects of thought and other propositional attitudes, and so on. Successor theories have in some cases bifurcated this role, but in other cases have expanded it, by taking semantics to interface with adjoining theories of cognitive science, both linguistic (such as syntax) and psychological. The metaphysical semanticist seeks to explain *some* of the same phenomena as does the linguistic semanticist. For example, just like the linguistic semanticist, she wants to help explain why English speakers will point to the salient horse, rather than the salient car, when they hear the sounds “Point to the horse!”; and like the linguistic

semanticist, she will invoke concepts like truth and reference to do so. But she is not concerned to integrate her semantics with other linguistic or psychological theories. Thus she is not trying to integrate her semantics with syntactic theory, for example. And she is free to assign semantic values that competent speakers would be incapable of recognizing as such, for she is not trying to explain what a competent speaker knows when she understands her language. She might, for example, assign to an ordinary sentence about ordinary macroscopic objects a meaning that makes reference to the fundamental physical states of subatomic particles. And she might simply ignore Frege's (1952/1892) puzzle of the cognitive nonequivalence of co-referring proper names, since she is not trying to integrate her semantics with theories of action and rationality.

Returning to the first distinctive feature of metaphysical semantics: suppose a metaphysical semantics for a language  $L$  takes the form of a truth theory—a theory issuing in theorems of the form:<sup>9</sup>

Sentence  $S$  of  $L$  is true in  $L$  iff  $\phi$ .

The requirement that meanings be “given” in purely joint-carving terms amounts to the requirement that  $\phi$  be phrased in purely joint-carving terms. (‘Sentence of  $L$ ’ and ‘is true in  $L$ ’ are, of course, not purely joint-carving. Remember that the notion of metaphysical semantics is not intended to be a fundamental one.)

A truth theory is just one form that a metaphysical semantics might take. A metaphysical semanticist might conclude that, rather than assigning truth-conditions, a more explanatory approach would be to assign expressivist assertion-conditions to normative discourse, say, or proof-theoretic assertion-conditions to mathematical sentences. In the former case, the metaphysical semantics would issue in theorems of the form:

Normative sentence  $S$  of  $L$ , as uttered by speaker  $x$ , is expressively appropriate for  $x$  in  $L$  iff  $\phi(x)$

where  $\phi(x)$  is a condition on speaker  $x$ 's attitudes. Here the requirement that meanings be given in purely joint-carving terms requires  $\phi(x)$  to be stated in purely joint-carving terms. If psychological language is fundamental then  $\phi(x)$  may contain such language; otherwise (the more likely case, to my mind)  $\phi(x)$

---

<sup>9</sup>There is no need for ‘iff’ to have a sense that is somehow distinctive of metaphysical reduction. It can have the same sense that it has in any explanatory theory—the material biconditional, say.

will need to be phrased in the terms one would use for giving metaphysical truth-conditions for factualist discourse about psychology.

Expressivism thus construed is immune to a common dilemma. Suppose we extend the expressivist semantics to encompass a disquotational truth predicate (and related vocabulary). Expressivism is supposed to be a form of nonfactualism; but the claims one can assert in  $L$  (such as “The sentence ‘It is wrong to kill’ is true”) now sound just like the claims that factualists make. What happened to the nonfactualism?<sup>10</sup>

The answer is that the metaphysical semantics for factual discourse (say, discourse about chemical or biological phenomena) has a different shape from the metaphysical semantics for discourse about value. To a speaker’s sentence, the former assigns a condition on the world—a closed sentence—whereas the latter assigns a condition on the speaker—an open formula applied to a variable to which the speaker is assigned; and this condition, moreover, concerns the speaker’s mental state.<sup>11</sup> True, this difference disappears if one states a semantics for  $L$  from within  $L$  (provided  $L$  contains, or is enhanced to contain, the appropriate semantic vocabulary). Such a (nonmetaphysical) semantics might then take the form of a truth theory, issuing in theorems like:

‘It is wrong to kill’ is true in  $L$  iff it is wrong to kill.

But the difference remains at the level of metaphysical semantics. A metaphysical semantics could not take this form, assuming that ‘wrong’ and ‘kill’ fail to carve at the joints. And the expressivist who is opposed to reductive naturalism will argue that no metaphysical semantics issuing in truth-conditions of the form:

‘It is wrong to kill’ is true in  $L$  iff  $\phi$

where  $\phi$  is a condition on the world (rather than the speaker) phrased in fundamental terms, is adequate to our use of normative language.

Returning to metaphysical semantics of the truth-theoretic form: what exactly is required of a fundamental truth-condition  $\phi$  for a sentence  $S$  in

---

<sup>10</sup>See Dreier (2004); Fine (2001). My solution to the problem is in the vicinity of Dreier’s and Fine’s: we all agree that the reason expressivism is nonfactualist (Dreier says “irrealist”) has something to do with expressivism’s implications for how value relates to the fundamental. I prefer my approach to Fine’s for the reasons given in section 7.9 below.

<sup>11</sup>There are other differences between truth-conditional and expressivist semantics, especially when they are integrated with a broader theory of the mind. For example, truth-conditions and expressive appropriateness conditions play different roles in communication and deliberation.

such a theory? We know from the literature on Donald Davidson's (1967*b*) approach to semantics that sentences  $\phi$  and  $\phi'$  can have the same truth-value, even necessarily so, despite the fact that  $\phi$  is an appropriate truth-condition for  $S$  while  $\phi'$  is not.<sup>12</sup> ‘Snow is white’ is an appropriate truth-condition in a (nonmetaphysical) semantics for ‘Snow is white’; neither ‘Grass is green’ nor ‘Snow is white and  $2 + 2 = 4$ ’ is appropriate. A metaphysical semantics must successfully explain the linguistic behavior of the population in question, and a truth-theory with necessarily true conjuncts tacked onto each of its truth-conditions is presumably not explanatory, though it's a hard question why not. I have no particular answer to the question, though I suspect that the approach of section 3.2 is applicable. And if the question proves intractable, the metaphysical semanticist could abandon the Davidsonian approach. I have chosen that approach largely because it's simple, not because metaphysical semantics is wedded to it.

What I have said about metaphysical semantics falls far short of a full characterization, and it could surely be improved in various ways. However, it is unlikely that there is any single best way to improve it. The notion of a metaphysical semantics for a language, after all, is a high-level notion—a notion posited in the course of explaining a high-level phenomenon, the phenomenon of language-use by flesh and blood people; and there are often different, equally good ways to explain such phenomena, as well as equally good ways to carve the world into groups of phenomena to be explained. Given our discussion of purity, this is a virtue. We should take a nonfundamental, no doubt vague, and perhaps not even uniquely correct, approach to the question of the relationship between fundamental and nonfundamental.

## 7.5 Completeness and purity reformulated

We can now, at last, take up the question of how to formulate completeness and purity in my preferred terms. Those theses were, recall, initially formulated as follows:

**Completeness** Every nonfundamental truth holds in virtue of some fundamental truth.

**Purity** Fundamental truths involve only fundamental notions.

---

<sup>12</sup>See Soames (1992) against the Davidsonian approach.

The theses involve three crucial notions: fundamental truth, in-virtue-of, and fundamental notion. A fundamental notion is just one that carves at the joints; and in-virtue-of I am going to explain in terms of metaphysical semantics. But how to understand the notion of a fundamental truth?

Here we encounter something interesting. There are two natural definitions of fundamental truth in terms of the other two notions; what is interesting is that on each definition, one thesis comes out trivial and the other comes out nontrivial. The first natural definition of a fundamental truth is that of a truth involving only fundamental notions. On this definition, purity comes out trivial while completeness comes out nontrivial. The other natural definition of a fundamental truth is that of a truth that does not hold in virtue of any truth. Now purity becomes nontrivial while completeness becomes—somewhat—trivial. I say “somewhat” because completeness does not follow solely from the definition; but it does follow from the definition together with a natural assumption about how in-virtue-of behaves, namely that it is transitive and “well-founded” in the sense that if a truth holds in virtue of any truth at all, then it is connected by an in-virtue-of chain to some truth that does not hold in virtue of any truth (“no unbounded descending chains of in-virtue-of”).<sup>13</sup> So it would seem that completeness and purity are in some sense not wholly independent, at least when the notion of a fundamental truth is regarded as being definable in terms of the other two notions.

My—somewhat arbitrary<sup>14</sup>—decision for how to define ‘fundamental truth’ will be the first: a fundamental truth is a truth involving only fundamental terms. Thus understood, purity becomes trivial. But notice that its upshot has not disappeared. Completeness now says the following: “Every truth that involves at least one nonfundamental notion holds in virtue of some truth that involves only fundamental notions”; and this has implications that are analogous to the distinctive implications of purity discussed in section 7.2. Completeness, as it’s now understood, implies that even if a notion is fundamental, any fact involving that notion together with some nonfundamental notions must hold in virtue of pure or austere facts about that notion, which do not involve any

---

<sup>13</sup>Thanks to Karen Bennett here. For more on well-foundedness, see section 7.11.2.

<sup>14</sup>I could instead adopt the second definition of ‘fundamental truth’, rendering completeness (somewhat) trivial. The corresponding version of purity would then be: truths that do not hold in virtue of any truth involve only fundamental notions. That is: any truth that involves a nonfundamental notion must hold in virtue of some other truth. Modulo the assumptions of transitivity and well-foundedness of in-virtue-of, this is equivalent to the result of the other decision.

nonfundamental notions. (It implies, for example, that even if necessity is a fundamental notion, the truth that every  $C_0$  is a city must hold in virtue of truths that involve only fundamental notions.)

So: completeness, as we're now understanding it, says that every truth that involves at least one nonfundamental notion holds in virtue of some truth that involves only fundamental notions. It remains to eliminate "in virtue of" in favor of metaphysical semantics. Actually, I wish simultaneously to generalize completeness a little, to handle the case of expressivist and other such language:

**Completeness (new version)** Every sentence that contains expressions that do not carve at the joints has a metaphysical semantics.

By a metaphysical semantics for a sentence, I mean either a truth-condition, an expression-condition, a proof-condition, or perhaps some other sort of semantic condition, that is assigned to that sentence by some metaphysical semantics for its language. This principle constrains the notion of carving at the joints in accordance with the issues we have been discussing in this chapter. There is no need for a further principle of purity, since the distinctive implications of that principle discussed in section 7.2 are implied by this new version of completeness. (But when I wish to emphasize these implications, I will speak of "purity".)

## 7.6 Metaphysics after conceptual analysis

How much should we expect from ourselves, if we attempt to actually give metaphysical truth-conditions? Not much, I think.

We certainly should not expect to be able to give truth-conditions that are intuitively correct in every possible world or conceivable circumstance. Judging from the history of conceptual analysis, that is unattainable. In retrospect, we should never have expected that project to succeed. Why should there be any simple definitions, preserving intuitive or cognitive significance, of any of our words in any other terms? Words aren't generally introduced as definitional equivalents of pre-existing phrases, and even then they subsequently take on semantic lives of their own. Current meaning derives from a long, complex history of use, which would seem unlikely to result in neat equivalences. Our failure to come up with counterexample-free definitions of 'cause', 'knows', and 'good' is not due to the philosophical depth of these concepts. We'd have no

better luck with ‘city’, ‘smile’, or ‘candy’. Words just aren’t neatly equivalent to other words, and there’s no reason to expect them to be.

And it only gets harder if the truth-conditions must be stated in absolutely fundamental terms, as metaphysical truth-conditions must be. Our ignorance of the facts of physics then compounds our ignorance of the facts of meaning. We have no chance of actually giving a metaphysical semantics for any significant fragment of a natural language.

A more reasonable goal is the construction of “toy” metaphysical truth-conditions. These will be toy in at least two ways. First, they needn’t match with intuitively correct usage in absolutely all possible worlds or conceivable circumstances. The mesh need only be approximate (the more mesh, the better). Second, they needn’t be stated in perfectly fundamental terms. Rather, they must be stated in terms that are fundamental enough for the purpose at hand (and again, the more fundamental the better).

What is the point of toy metaphysical truth-conditions? One point is to convince us that real, non-toy metaphysical truth-conditions exist.<sup>15</sup> Consider, for example, the controversy over whether causation is fundamental. To help resolve this controversy, we might try to produce reductive (i.e., not involving ‘cause’) toy metaphysical truth-conditions for ‘cause’. If all attempts fail, then the case for fundamental causation will receive a nice boost, especially if we discern in-principle reasons for the failures. For if reductive metaphysical truth-conditions for ‘cause’ do not exist (and if ‘cause’ lacks an expressivist or otherwise non-truth-conditional metaphysical semantics), then completeness forces us to say that causation is fundamental. But if a toy semantics *can* be produced, we needn’t recognize fundamental causation (not because of completeness, anyway). The toy needn’t be perfect. For its purpose is not to *be* a real metaphysical semantics, but rather to convince us that *there is* a real metaphysical semantics, even if that metaphysical semantics is too complex for us to discover. The purpose of the toy is to convince us that the unfathomable workings of history and usage can do their thing with ‘cause’, as well as with ‘city’, ‘smile’, and ‘candy’. The reason philosophers obsess over the definition of ‘cause’ is not that they think that a word must be defined before it’s legit (otherwise no philosophers would speak of cities, smiles, or candy). It’s rather that it’s a live issue whether causation is part of the fundamental furniture of

---

<sup>15</sup>Another is to study how our concepts relate to one another. Even a simplistic semantics might illuminate the overall shape of a network of concepts consisting of wrongdoing, blame, guilt, shame, and the like.

the universe. We don't obsess over the definition of 'candy', not because we could easily define it if we wanted to (we couldn't), but rather because no one seriously contemplates fundamental candy.

Conceptual analysis is out of fashion in metaphysics, but there is uncertainty about what its replacement should be. Reduction? Supervenience? Realization? The proposed replacements have tended to be either inadequate or unilluminating (or both). The recent trend is to think in terms of a kind of purely metaphysical analysis. There's a bad idea in here mixed in with a good one. The bad idea is that we should posit a fundamental gizmo for the relationship between analysans and analysandum (truthmaking, necessity, fact identity, ground). The good idea is Armstrongian: metaphysical analysis is not linguistic analysis.<sup>16</sup> But that leaves a big gap: if not linguistic analysis then what? I say: metaphysical truth-conditions—toy models of them, at any rate.

## 7.7 Metaphysical semantics for quantifiers

Metaphysical semantics are not required by definition to take any particular form. They must presumably be compositional in some sense (since they must be explanatory and hence cast in reasonably joint-carving terms, and must contend with infinitely many sentences). But this still allows considerable variation.

In particular, the form of metaphysical truth-conditions can depend on what is fundamental. After all, metaphysical truth-conditions must be stated in perfectly fundamental terms; thus which terms are in fact fundamental, and what is true at the fundamental level, will affect what metaphysical truth-conditions can look like.

An illustration comes from ontology. If fundamental ontology is abundant, then the metaphysical truth-conditions for existential claims in nonfundamental languages can be existential in form; but if fundamental ontology is sparse, then these metaphysical truth-conditions need to take some other form.<sup>17</sup> To bring this out I will consider an example at length. I will give (toy) metaphysical truth-conditions for certain statements of chemistry, first assuming classical mereology, and then assuming mereological nihilism.

Assume classical mereology. More fully, assume for the sake of argument that the joint-carving notions are those of logic:  $\exists$ ,  $\forall$ ,  $\sim$ ,  $\wedge$ ,  $=$ , etc.; a pred-

---

<sup>16</sup>See, for example, Armstrong (1978a,b).

<sup>17</sup>This corresponds to Fine's (2003) distinction between proxy and non-proxy reductions.

icate  $<$  for mereological parthood; and the following physical predicates:  $E$  (“is an electron”),  $P$  (“is a proton”),  $N$  (“is a neutron”),  $R$  (“orbits”), and  $U$  (“is a nucleus”). ( $U$  is to apply to fusions of protons and neutrons that are bound together into a nucleus by the strong nuclear force.<sup>18</sup>) Our fundamental language is thus one in which only these notions are primitive. One can then define, in the fundamental language, various mereological notions using  $<$  in well-known ways. For example, one can define  $Oxy$  (“ $x$  overlaps  $y$ ”) as meaning  $\exists z(z < x \wedge z < y)$ . And for any fixed positive integer  $n$ , one can define an  $n + 1$ -place fusion predicate  $xFu^n y_1 \dots y_n$  (“ $x$  is a fusion of  $y_1 \dots y_n$ ”) thus:

$$xFu^n y_1 \dots y_n =_{\text{df}} y_1 < x \wedge \dots \wedge y_n < x \wedge \forall z(z < x \rightarrow (Ozy_1 \vee \dots \vee Ozy_n))$$

And assume further that classical mereology’s principle of “unrestricted composition” is true. For present purposes we may take this as the assumption that for each  $n$ , ‘ $\forall y_1 \dots \forall y_n \exists x xFu^n y_1 \dots y_n$ ’ is true.

Now consider a very simple language of chemistry, which is just like the fundamental language except for containing new predicates:  $H$  (“is an atom of hydrogen”) and  $L$  (“is an atom of helium”).<sup>19</sup> Thus, this language contains sentences like:

$$(1) \exists x Hx$$

“There exists an atom of hydrogen.”

$$(2) \exists x Lx$$

“There exists an atom of helium.”

We can give metaphysical truth-conditions for these sentences by, in essence, defining atoms as fusions of their subatomic particles. Say that a sentence  $\phi$  in the fundamental language “translates” a sentence  $\chi$  in the language of chemistry iff  $\phi$  results from  $\chi$  by replacing occurrences of  $H$  and  $L$  according to the following definitions:

$$Hx =_{\text{df}} \exists y \exists z(Ey \wedge Pz \wedge Ryz \wedge xFu^2yz)$$

“ $x$  is a hydrogen atom iff  $x$  is a fusion of an electron and proton where the electron orbits the proton.”

---

<sup>18</sup>On a more plausible view,  $U$  would nonfundamental, and defined in terms of mereology and a fundamental predicate for the strong nuclear force. The approach in the text is for simplicity.

<sup>19</sup>I have in mind hydrogen-1 and helium-4, respectively.

$$Lx =_{\text{df}} \exists y_1 \exists y_2 \exists z_1 \exists z_2 \exists w_1 \exists w_2 \exists v (Ey_1 \wedge Ey_2 \wedge y_1 \neq y_2 \wedge Pz_1 \wedge Pz_2 \wedge z_1 \neq z_2 \wedge Nw_1 \wedge Nw_2 \wedge w_1 \neq w_2 \wedge Uv \wedge Ry_1v \wedge Ry_2v \wedge vFu^4 z_1 z_2 w_1 w_2 \wedge xFu^3 y_1 y_2 v)$$

“ $x$  is a helium atom iff  $x$  is a fusion of two electrons and a nucleus, where the nucleus is a fusion of two protons and two neutrons, and the electrons orbit the nucleus.”

We could say, then, that the metaphysical truth-condition of any sentence in the language of chemistry is its translation. Thus, the metaphysical truth-conditions of (1) and (2) are, respectively:

$$\exists x \exists y \exists z (Ey \wedge Pz \wedge Ryz \wedge xFu^2yz)$$

$$\exists x \exists y_1 \exists y_2 \exists z_1 \exists z_2 \exists w_1 \exists w_2 \exists v (Ey_1 \wedge Ey_2 \wedge y_1 \neq y_2 \wedge Pz_1 \wedge Pz_2 \wedge z_1 \neq z_2 \wedge Nw_1 \wedge Nw_2 \wedge w_1 \neq w_2 \wedge Uv \wedge Ry_1v \wedge Ry_2v \wedge vFu^4 z_1 z_2 w_1 w_2 \wedge xFu^3 y_1 y_2 v)$$

This was a particularly simple example of metaphysical truth-conditions. Given classical mereology, our fundamental ontology already contained helium and hydrogen atoms, so to speak; all we lacked was primitive predicates classifying them as such. So the definition of translation was quite easy; all we needed to do was define the predicates  $L$  and  $H$  in fundamental terms. Things get more complex if a sparser fundamental ontology is true, as we'll now see.

Assume next that mereological nihilism is true—no mereologically composite things exist. More fully, assume that  $\forall x (Ex \vee Px \vee Nx)$  is true in our fundamental language—nothing exists other than electrons, protons, and neutrons (these lack proper parts, let us pretend). Drop  $<$  from the fundamental language (it is unneeded since nothing has proper parts). The fundamental language must no longer speak of nuclei (there aren't any); let us replace, therefore, the predicate  $U$  with a two-place predicate  $B$  (“bonded”) holding between the protons and neutrons that we formerly called the parts of nuclei. Similarly, instead of speaking of electrons orbiting nuclei, let us now speak of electrons as orbiting protons and neutrons. Let us assume that  $B$  is transitive and symmetric, and reflexive over protons and neutrons, and that if  $Rxy$  and  $Byz$  then  $Rxz$ —if an electron orbits one subatomic particle “in a nucleus” then it orbits each subatomic particle “in that nucleus”.

What form must metaphysical truth-conditions for sentences of chemistry now take? We can no longer translate claims of chemistry into fundamental claims simply by giving definitions, in fundamental terms, of the predicates  $H$  and  $L$ . If we proceeded that way, then the translation of sentence (1), i.e.,

$\exists x Hx$ , would have the form  $\exists x \psi(x)$ , where  $\psi(x)$  is the proposed definition, in fundamental terms, of  $Hx$ . Since only subatomic particles exist, and since we want (1) to come out true in the language of chemistry,  $\psi(x)$  must apply to subatomic particles. But  $\psi(x)$  was supposed to be the definition of ‘ $x$  is a hydrogen atom’.<sup>20</sup>

We need a different strategy of translation. A natural approach is to translate (1) and (2) into:

$$(1_N) \exists x \exists y (Ex \wedge Py \wedge Rx y)$$

$$(2_N) \exists x_1 \exists x_2 \exists y_1 \exists y_2 \exists z_1 \exists z_2 (Ex_1 \wedge Ex_2 \wedge x_1 \neq x_2 \wedge Py_1 \wedge Py_2 \wedge y_1 \neq y_2 \wedge Nz_1 \wedge Nz_2 \wedge z_1 \neq z_2 \wedge By_1 y_2 \wedge By_1 z_1 \wedge By_1 z_2 \wedge Rx_1 y_1 \wedge Rx_2 y_1)$$

On this approach, the translation of ‘There exists a hydrogen atom’ is: ‘There exist an electron and a proton, the first of which orbits the second’. The translation omits reference to the hydrogen atom itself; it states the nihilistic basis for the entire sentence ‘There exists a hydrogen atom’.<sup>21</sup>

---

<sup>20</sup>Well, one *could* (at the price of artificiality) specify a tricky translation scheme in which  $\psi(x)$  does indeed apply to subatomic particles, namely, those subatomic particles that are “part of” hydrogen atoms:  $\psi(x) =_{\text{df}} \exists y \exists z (Ey \wedge Pz \wedge Ry z \wedge (x = y \vee x = z))$ . One would need to make adjustments elsewhere. For example, the language of chemistry’s predicate  $P$  could not be translated as the fundamental language’s predicate  $P$ ; otherwise  $\exists x (Px \wedge Hx)$  (“something is both a proton and a hydrogen atom”) would be translated as a truth.

<sup>21</sup>Here is a general translation scheme of the desired sort. Let  $F$  be the set of sentences in the fundamental language that express the assumptions we are making about our fundamental predicates.  $F$  thus includes the claim that  $B$  is transitive, symmetric, and reflexive-over-protons-and-neutrons. So in any model,  $M$ , of  $F$ ,  $M(P) \cup M(N)$  is, if nonempty, partitioned into equivalence classes under the relation  $M(B)$  (I use “ $M(\pi)$ ” for the extension of predicate  $\pi$  in  $M$ ). Think of these equivalence classes as “nuclei” (they obviously aren’t really nuclei;  $M$  is only a model). For each such equivalence class,  $c$ , call the ordered pair  $\langle c, b \rangle$  an “atom”, where  $e \in b$  iff:  $e \in M(E)$  and  $\langle e, o \rangle \in M(R)$  for some  $o \in c$ . Each “atom” is an ordered pair of a “nucleus”—a set of “protons” and “neutrons”—and the set of “electrons” that “orbit” the members of the “nucleus”. Next we construct an augmented model,  $M'$ , by adding the “atoms” to  $M$ ’s domain. (The added “atoms” are to be new; so if any are already present in  $M$ ’s domain, first pair them with some arbitrarily chosen object not in the transitive closure of  $M$ ’s domain.) Let the extensions of  $E$ ,  $P$ ,  $N$ ,  $B$ , and  $R$  in  $M'$  be as they were in  $M$ . And let  $M'$  also interpret the extra predicates of the language of chemistry: let the extension in  $M'$  of  $<$  include all pairs  $\langle d, \langle c, b \rangle \rangle$  where  $\langle c, b \rangle$  is an “atom”, and either  $d \in c$  or  $d \in b$ , plus further pairs so that  $<$  satisfies the axioms of mereology; and assign extensions to  $H$  and  $L$  in the obvious way. (Place an “atom”  $\langle c, b \rangle$  in  $M'(H)$ , for example, iff  $b$  has exactly one member, and  $c$  has exactly one member, which is a member of  $M'(P)$ .) Finally, say that  $\phi$  in the fundamental language translates  $\chi$  in the language of chemistry iff for every model  $M$  of  $F$  in which  $\phi$  is true,  $\chi$  is

In such a metaphysical semantics, the truth-conditions for sentences of the form  $\exists x Fx$  do not have the form  $\exists x \psi(x)$ , with  $\psi$  a translation of the predicate  $F$ . Quantification over  $F$ s disappears when we move from the chemical sentence to its metaphysical truth-condition. Given this, it is natural to say that the existential quantifier in the language of chemistry does not mean what it means in the fundamental language used to give metaphysical truth-conditions. Quantification in the language of chemistry is nonfundamental quantification. We might make this explicit by using ‘there is’ in the language of chemistry, reserving ‘ $\exists$ ’ for fundamental quantification.<sup>22</sup>

Quantifiers in many languages—for example, ordinary languages in which we quantify over tables and chairs—might in this way express nonfundamental quantification, if fundamental ontology is sparse. Granted, the metaphysical semantics for a more complex language will need to be more complex than the toy semantics just mentioned. And particularly austere views about fundamental ontology or ideology might make it impossible to give metaphysical truth-conditions for some high-level language—which might be a reason for abandoning such austere views. The point here is just to demonstrate some of the resources available for giving metaphysical truth-conditions, and to show how a sentence’s metaphysical truth-conditions might look quite unlike that sentence, as with the truth-conditions  $(1_N)$  and  $(2_N)$ .

## 7.8 Metaphysics and the study of language

Suppose that mereological nihilism is true. Should linguistic semanticists—real live semanticists in linguistics departments, not metaphysical semanticists—then follow the lead of last section’s metaphysical semanticist, and assign truth-conditions like  $(1_N)$  and  $(2_N)$  to sentences of the language of chemistry? It might be thought that they should; otherwise, they would count the chemist’s sentences as being false.

There is an alternative. If the language of chemistry can have a metaphysical semantics that allows its sentences to be true despite mereological nihilism,

---

true in the corresponding augmented model  $M'$ . Notice that a given sentence might now have more than one translation. (Notice also the use of set theory to specify the translations, even though the fundamental theory in question was nominalistic. There’s no immediate conflict since only metaphysical truth-conditions themselves, not the description of how to arrive at them, must be stated in purely fundamental terms.)

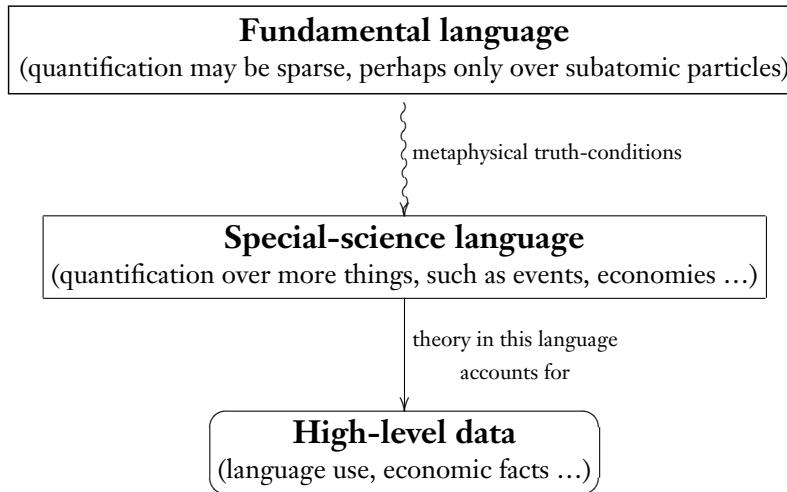
<sup>22</sup>See also section 9.3.

then why not the metalanguage used by linguists? Linguists could then use sentences quantifying over atoms of hydrogen, helium, and the like, rather than sentences like  $(1_N)$  and  $(2_N)$ , to give truth-conditions for the sentences of chemistry. The chemist's 'There exists a hydrogen atom' would then have, as a *linguistic* truth-condition, the linguist's homophonic sentence 'There exists a hydrogen atom.' This linguistic truth-condition must be distinguished from the *metaphysical* truth-condition shared by both sentences: ' $\exists x \exists y (Ex \wedge Py \wedge Rxy)$ ' (i.e.,  $(1_N)$ ). Similarly, linguists might assign to ordinary sentences about tables and chairs, truth-conditions that themselves quantify over tables and chairs, provided the sentences of their metalanguage have appropriate metaphysical truth-conditions.

There is a tradition in the philosophy of language according to which linguistic and metaphysical inquiry should tightly constrain each other. Davidson (1977) is a representative example. According to Davidson, a good semantic theory must count ordinary sentences as being, for the most part, true. Suppose the best semantic theory for a discourse assigns truth-conditions to ordinary sentences that quantify over entities of a certain sort. Then, since the ordinary sentences must be counted as true, the assigned truth-conditions must be true; and so, the entities in question must exist. For example, Davidson argued, the best linguistic theory of adverbial modification assigns truth-conditions quantifying over events (1967a); thus we must embrace an ontology of events. This is an example of linguistic theory constraining metaphysics, but there is no reason in principle, given this tradition, to resist the reverse direction of influence: that of metaphysics constraining linguistic theory. Powerful metaphysical arguments that events do not exist, for example, would give us reason to reject Davidson's approach to adverbial modification.

An advantage of the metaphysical semantics approach, as against the Davidsonian tradition, is that it allows for a looser relationship between metaphysics and linguistics. We can agree with Davidson that linguistic semantics ought to count ordinary sentences as being mostly true, without needing to embrace an ontology of events, tables and chairs, or atoms of chemistry, because ontology concerns fundamental existence, whereas linguistic semantics is given in the metalanguage of linguistics, whose quantifiers need not be fundamental. The linguist's sentences about events, atoms, and tables and chairs can be true even if such entities do not fundamentally exist, given an appropriate metaphysical semantics for those sentences.

The approach allows, more generally, for a looser relationship between metaphysics and the special sciences. That relationship may be pictured thus:



Linguistics, psychology, economics, and other special sciences may be carried out in their own languages—largely natural languages, enhanced here and there with special-purpose vocabulary. Sentences of special-science languages have metaphysical truth-conditions, but these are of no more concern to the special scientist than the underpinnings of her discipline in fundamental physics. Explanations of high-level data are given in the language of these special sciences, not in the underpinning languages of physics or metaphysics.

The advantage of this approach is that it allows linguists, psychologists, and economists to be guided by considerations internal to linguistics, psychology, and economics. It would be inappropriate to complain to an economist that economies don't really exist, or to insist that an engineer rewrite her book on repairing potholes to reflect the fact that holes do not really exist. Likewise, it would be inappropriate to require linguists to warp semantics around metaphysical scruples about molecules of helium, or tables and chairs, or events.

Conversely, it allows metaphysicians to be guided by considerations internal to metaphysics. It has always seemed odd that insight into the fundamental workings of the universe could be gained by reflection on how we think and speak. Of course, such reflection can provide *some* constraint on metaphysics. Human thought and speech are real phenomena, and so must fit somehow into any adequate metaphysics. But this is a far cry from reading off one's fundamental metaphysics directly from the structure of thought and talk.

Thus we have a limited Carnapian (1950) spirit of tolerance. Special sciences can conduct their business without interference from metaphysics, if their

languages can be given a metaphysical semantics. (The same point from another angle: metaphysics can be relatively free from interference from the special sciences. Note how structure opens up breathing room for metaphysics.) The tolerance is limited for two reasons. First, the metaphysical semantics must not be too complex; otherwise one might question whether the science is genuinely explanatory. Whether, and if so when, an alleged science is unexplanatory for this sort of reason is a difficult question about special-science explanation.<sup>23</sup> Second, it cannot be just assumed that a metaphysical semantics can be given. In particular, Carnap's insouciance about the ontological status of mathematical entities is *not* justified, since it is particularly hard to see how mathematical language can be given a metaphysical semantics if, fundamentally speaking, mathematical entities do not exist. It is comparatively easy to see how chemistry and biology could rest on top of physics, since the world of physics is fine-grained enough to supply sufficiently many facts to underly chemistry and biology. But the infinitary nature of mathematics presents a special challenge. If one's conception of the fundamental is overly sparse—if it contains for example, neither mathematical entities nor a correspondingly rich structure of modal or higher-order logical facts—then there may simply be no way to give a metaphysical semantics for mathematical language.

## 7.9 Nonfundamental ground

I have given a linguistic account of connecting facts—facts connecting nonfundamental to fundamental. It might be urged, against this, that the matter of how chemistry, biology, economics, and so forth, relate to the fundamental does not *seem* to concern language, and hence that we ought to regard connecting facts as involving some nonlinguistic notion such as ground. As we saw, purity requires connecting facts to be nonfundamental; but a friend of ground might embrace this.

The friend of ground that I have in mind shares my main approach to fundamentality: his basic notion is that of structure, and he embraces purity; it's just that he accounts for the connection between fundamental and nonfundamental in terms of ground, not metaphysical semantics.<sup>24</sup> Given his embrace of purity, he must say that ground-theoretic connecting facts—facts like: there being

---

<sup>23</sup>As pursued, for example, by Fodor (1974) and Kim (1992).

<sup>24</sup>In chapter 8 I consider views that *replace* the notion of structure with ground and/or related notions.

$C_0$ s grounds there being cities—are themselves grounded in facts that do not involve nonfundamental notions; and he presumably cannot produce reductive definitions of ‘ground’, ‘city’, and other relevant terms that demonstrate compliance with purity. However, I am in the same boat: I am committed to the existence of a metaphysical semantics for sentences containing ‘metaphysical semantics’, but I cannot produce that metaphysical semantics.

It would be easy to overstate the difference between this approach and my own. After all, a metaphysical semantics is supposed to explain linguistic phenomena in purely fundamental terms, and the sort of explanation required is distinctively metaphysical in nature since the meanings must be given in fundamental terms. Still, there is a reason to prefer my approach: it handles nonfactual discourse more smoothly.

Given my approach, there is a simple and natural way to distinguish factual from nonfactual discourse: the difference is one of the “shape” of the metaphysical semantics. The shape is truth-conditional in the former case, and some other shape in the latter. For example, an expressivist metaphysical semantics for evaluative discourse might take the form of an assignment of expressively appropriate assertion conditions; and a formalist metaphysical semantics for mathematical discourse might take the form of assertion conditions that are sensitive to the proofs that the speaker possesses.

How will the friend of ground distinguish factual from nonfactual discourse? I see two main possibilities. The first would be to exclude nonfactual discourse from the scope of grounding. One way to implement this would be to say that grounding only concerns *facts*; and since nonfactual sentences do not express facts, grounding does not apply to them. But the notion of fact thus invoked is in need of explanation (it cannot be disquotational), and surely the explanation ought to have something to do with grounding. So let us understand talk of facts disquotationally from now on, and implement this first possibility differently: nonfactual sentences do express facts, alright—call these “nonfactual facts” (sorry)—it’s just that those facts are ungrounded. Moreover, nonfactual facts are not fundamental (that is, they are not cast purely in joint-carving terms), which is what distinguishes them from, say, physical facts, which are also ungrounded.

The problem here, however, is that the approach has nothing to say about *how* nonfactual facts relate to the fundamental. Various sorts of nonfactual facts (moral and mathematical, perhaps) are all lumped together as being ungrounded and nonfundamental. My approach, in contrast, makes distinctions within the class of nonfactual sentences, depending on the “shape” of their metaphysical

semantics.

To be fair, the friend of grounding can make distinctions within the class of nonfactual facts by an indirect method, following Fine (2001). Unlike mathematical facts themselves, facts about our beliefs about mathematics, and metalinguistic facts about mathematical language, are grounded in the fundamental, and may be thus grounded in a different fashion from how facts about our beliefs in value, or metalinguistic facts about evaluative language, are grounded in the fundamental. Perhaps the former facts are grounded in facts involving proof whereas the latter are grounded in facts involving our attitudes. So the friend of grounding can draw the distinctions that need to be drawn. But they must be drawn so indirectly, and so differently from how analogous distinctions are drawn in the case of factual discourse.

A second possibility for treating nonfactual discourse would be to allow nonfactual facts to be grounded. This is Fine's approach (2001). Some nonfactual facts are grounded in others: the fact that either murder or snorkeling is wrong is grounded in the fact that murder is wrong. But not all nonfactual facts are grounded: the fact that murder is wrong is perhaps an example. This fact is ungrounded, but is distinguished from the physical facts in that it is not fundamental.<sup>25</sup> This second possibility is the more attractive one, I believe, but it faces the same problem as the first. It has nothing to say about *how* ungrounded nonfactual facts are related to the fundamental; it lumps all ungrounded nonfactual facts—moral, mathematical, say—together. As with the first possibility, distinctions can indeed be made, by examining how moral beliefs and metalinguistic facts about moral language are grounded; but these distinctions would be drawn indirectly.

My objection, then, is not that the grounding approach cannot draw the distinctions that need to be drawn. It is that it does not draw them in the most perspicuous way.

The root of the problem is the connection between the grounding approach and the disquotational conception of fact. Think of the grounding approach as follows. We begin in natural language, a language that is highly heterogeneous in that it contains both factual and nonfactual discourse. We then introduce a disquotational notion of fact, which applies to all asserted sentences, whether factual or no. Finally, we apply the notion of grounding to facts thus under-

---

<sup>25</sup>To say that this fact is grounded in our attitudes, for example, would turn moral expressivism into a form of descriptivism.

stood.<sup>26</sup> Now, disquotational notions of fact, truth, property, and the like are prized because of their ability to obliterate metaphysical differences. When I lack information about what my neighbor has said—including information about its status as factual or nonfactual—it helps to have catch-all notions of fact, truth, and so on, by which I may express agreement, disagreement, or otherwise make cognitive contact with my neighbor. But however valuable this metaphysical neutrality of disquotational notions is for ordinary purposes, it is a liability in the present, metaphysical, context, for here we are trying to highlight differences in how our thought and talk connect with fundamental reality. In the present context, it's best not to adopt such a catch-all conception of fact, for doing so already obscures many of the differences we wish to capture—even if those differences can be accounted for, down the line, in some indirect way.

There are other—related—reasons to prefer metaphysical semantics over grounding. I'll mention two. First, consider the approach to the liar paradox according to which both the liar sentence and its negation must be rejected (where rejection is not the same thing as assertion of the negation). On a grounding approach, we seem not to be able to say anything at all here. Since neither the liar sentence nor its negation is assertable, we can say nothing of the form “ $\phi$  grounds  $L$ ”, where  $L$  is either the liar sentence or its negation (assuming ‘grounds’ to be factive). But a natural approach can be taken if we speak in terms of metaphysical semantics rather than ground; see section 10.6.

Second, consider nonfundamental natural language quantification. Suppose that in the fundamental sense of ‘there is’, there are no such things as statues or lumps of clay, but that natural language is governed by a metaphysical semantics specifying that if some clay is appropriately shaped, then the following sentence is true in English: “There exists a lump made from that clay with modal properties  $m_L$ , and there also exists a distinct statue made from that clay, which has modal properties  $m_S$ .<sup>27</sup>” If we do not semantically ascend, and ask simply after the grounds of facts construed disquotationally, we will be led to an awkward place, as follows. Since there exist a lump and statue as described, there must exist a pair of singular facts, the fact that  $L$  has  $m_L$  and the fact that  $S$  has  $m_S$ , where  $L$  and  $S$  are the lump and statue in question. Further, if these are distinct facts, there must surely be some fundamental ground of one that is not a ground of the other—something fundamental that grounds their

---

<sup>26</sup>Fine (2001) ultimately prefers to regard ‘ground’ as a sentence operator rather than as a predicate of facts (or propositions); but similar remarks apply since the sentences to which this sentence operator may be applied are heterogeneous; they may be either factual or nonfactual.

distinctness. But no such differential ground can be located.<sup>27</sup> Now, it's not as if the friend of ground has no response. He might claim that although no ground differentiates the pair of facts, the complex fact that  $L$  and  $S$  are distinct and instantiate  $m_L$  and  $m_S$  respectively, does have a ground: namely, the fact that the clay exists and is appropriately shaped.<sup>28</sup> But it remains awkward that the facts in question lack differential grounds. Intuitively, one wants to say, there really is no such fact as that  $L$  has  $m_L$ , or that  $S$  has  $m_S$ , because there really are no such things as  $L$  or  $S$ . A more satisfying picture of the situation is achieved by semantic ascent. The metaphysical semantics for English provides metaphysical truth-conditions for various statements about statues and lumps of clay, but it does not do so by associating referents to singular terms like ' $S$ ' and ' $L$ '. It rather does so by associating complex truth-conditions for whole sentences containing quantifiers over, or singular terms for, both statues and lumps. These truth-conditions render 'Lump  $L$  has  $m_L$  and the distinct statue  $S$  has  $m_S$ ' true, despite the absence of distinct metaphysical truth-conditions for the sentences ' $L$  has  $m_L$ ' and ' $S$  has  $m_S$ '.<sup>29</sup>

These examples—of evaluative discourse, the liar paradox, and nonfundamental quantification—illustrate how adopting the disquotational notion of fact and asking after the grounds of facts thus understood can obscure what is important about the metaphysics of nonfundamental matters. We *can* speak of facts and ground in these contexts, but a clearer view is attained if we semantically ascend and describe how our discourse about values, truth, and nonfundamental entities relates to fundamental reality.

## 7.10 Subpropositional?

Conceptions of fundamentality may be *propositional* or *subpropositional*—they may be notions that apply to entire propositions, or to constituents of propositions. To avoid reifying propositions and their constituents, we can put it linguistically: a locution for talking about fundamentality might be *sentential*—applying to entire sentences—or *subsentential*—applying to parts of sentences.

Lewisian naturalness is subpropositional: it is properties and relations, rather than entire propositions, that are evaluated for naturalness. The notion of a fundamental truth is propositional: truths are entire propositions, not

---

<sup>27</sup> Compare Sider (2008b). deRosset (2010) raises related issues about grounding.

<sup>28</sup> Compare Dasgupta (2010) on plural ground and Fine (1994b) on reciprocal essence.

<sup>29</sup> Depending on details, this might call for a refinement of the completeness principle.

proposition-parts. The notion of ground is propositional: entire propositions ground one another.

Structure is subpropositional. In my official regimentation, judgments of structure take the form  $\mathcal{S}(\alpha)$ , where  $\alpha$  may be a subsentential expression (such as ‘is an electron’ or ‘and’). Thus the ultimate locus of fundamentality is for me subpropositional. (I have no objection to propositional notions of fundamentality—such as various notions of fundamental truth—so long as they are defined in terms of the subpropositional notion of structure.)

There are both systematic and intuitive reasons for taking structure to be subpropositional. The systematic reasons will emerge in section 8.3: a subpropositional notion is explanatorily more powerful. The intuitive reason is that subpropositionality is tied to the following attractive picture: there are some fundamental “building blocks”—the “ultimate constituents of reality”—and the nature of reality is given by the arrangement of those building blocks.

## 7.11 Absolute?

Conceptions of fundamentality may be *comparative* or *absolute*. Lewisian perfect naturalness, for example, is absolute: one says of a property or relation that it is perfectly natural (or not) simpliciter. Lewis also spoke of properties and relations being more or less natural; this is an example of comparative fundamentality. (Lewis defined comparative naturalness in terms of perfect naturalness and length of definitions; but an alternate approach would be to take the former as basic and define the latter in terms of it.<sup>30</sup>)

Structure is absolute: I say ‘is structural’ rather than ‘is more structural’. (In my official regimentation, the structure operator attaches to a single expression rather than to a pair of expressions: “ $\mathcal{S}$ (is an electron)”, “ $\mathcal{S}$ (and)”, and so on.) (I have no objection to comparative notions of fundamentality so long as they are defined in terms of the absolute notion of structure. More on this below.)

The main reason for taking structure to be absolute is that facts about structure (in interesting cases, anyway) are fundamental, whereas facts about comparative fundamentality are nonfundamental. Why regard facts about structure as being fundamental? Because structure is itself structural (section 7.13). While this allows *some* facts about structure to be nonfundamental, in interesting cases, claims about structure cannot be further explained. Why regard

---

<sup>30</sup>I discuss this variant in Sider (1993a, chapter 3).

facts about comparative fundamentality as being nonfundamental? Because of purity. The point of a comparative conception of fundamentality would largely be to connect fundamental to nonfundamental matters; but given purity, such comparisons could not be fundamental facts.

There is a further reason in favor of absolutism. As before, in interesting cases the facts of structure are fundamental; and fundamental facts are always determinate (section 7.12) and objective (chapter 4). But in many interesting cases, it's hard to believe that the facts of comparative fundamentality are determinate and objective—consider, for example, the question of whether geological notions are more fundamental than biological ones.

### 7.11.1 Absolutism and comparative structure

I have not been practicing what the previous section preaches. Throughout this book I have spoken of comparative structure: of carving “reasonably well” at the joints, carving “equally well” (though not perfectly) at the joints, carving “badly” at the joints, and so on. Examples: a nonsubstantive question was characterized as one whose answer depends on which of various candidate meanings we adopt, where the candidates are *equally* joint-carving, and where no other candidate is *more* joint-carving (section 4.2). The doctrine of reference magnetism appealed to imperfect joint-carving (section 3.2). Explanations were required to be cast in joint-carving terms, and in the case of special-science explanation, “joint-carving” cannot mean *perfectly* joint-carving (section 3.1). Relatedly, think of the traditional “levels” picture of the sciences, with physics at the bottom, chemistry next, and the other sciences arranged in some order or other on top. This ordering can be thought of as corresponding to the comparative fundamentality of the notions of those sciences.

Thus we need a comparative notion of structure in many of the applications. This may be reconciled with absolutism by distinguishing the fundamental notion of structure from the notion of structure in those applications; it is only the fundamental notion which is absolute. (Or we may rephrase thus, to avoid putting weight on notion-identity: the fundamental facts about structure involve only absolute structure; all facts about comparative structure are nonfundamental.) Talk of comparative structure must have metaphysical truth-conditions in terms of absolute structure (and other fundamental notions).

How to give such metaphysical truth-conditions? How to define comparative structure? I do not know. But I can suggest several elements to employ in a definition.

One element comes from Lewis. Lewis's notion of a perfectly natural property or relation is absolute; but he went on to define a comparative notion. A property or relation is more or less natural, he said, depending how short a definition it can be given in a perfectly natural language—a language in which all predicates stand for perfectly natural properties and relations (1986b, p. 61). This approach could be generalized in the case of structure: one notion is “definitionally more structural” than another, let us say, iff it has a shorter definition in a fundamental language—a language in which all expressions are structural.

Lewis's approach has been thought to face serious challenges (Hawthorne, 2007; Sider, 1995; Williams, 2007), chief among which is that it counts every two properties that require an infinite definition (of the same cardinality) as being equally natural.<sup>31</sup> But this is a limitation only when properties require infinite definitions, and it is far from clear that properties of interest do. It might appear otherwise because such properties have infinitely many realizations. A complete specification of a certain kangaroo, in perfectly natural terms, down to the last microphysical detail, is just one realization of the property of being a kangaroo; and there are infinitely many such realizations, since (for example) kangaroos can vary continuously in length. So, let us grant, the property of being a kangaroo has an infinite definition in a perfectly natural language: the disjunction of its realization-predicates. But this needn't be the shortest definition. The property of being a kangaroo might also have a finite functional definition:

$$x \text{ is a kangaroo} =_{\text{df}} x \text{ has some realization (property) or other that plays role } R$$

if the perfectly natural language allows quantification over properties, and if role  $R$  is finitely definable in that language.

The class of finitely definable properties is particularly rich if quantification over arbitrarily high-order properties and relations is available. Begin with an initial class of finitely definable properties. Then give finite definitions of further, higher-order, properties, that make reference to the initial properties. (For example, one might define a space  $S$  of the initial properties, as well as a geometry on  $S$ , and then define properties of the form: *having some member of*

---

<sup>31</sup>There is a parallel worry about a Lewisian account of special-science laws: Lewisian laws must be simple; unnatural properties detract from simplicity; special-science properties are equivalent to infinite disjunctions of physical properties and hence highly unnatural.

*region R of space S*, where *R* is finitely definable. Also one might make use of causal notions—viz., *having some member of region R which plays such-and-such causal role*—provided causal notions can themselves be finitely defined.) Next give finite definitions of still higher-order properties, which can make reference to the previously defined properties. And so on. Given higher-order resources, there is reason to be optimistic that properties of interest are finitely definable in perfectly natural and fundamental languages.

(The procedure requires quantification over properties, but such quantification needn’t be fundamental. Suppose, for example, that the language in which the definitions are cast includes no quantification over properties, but does include quantification over sets. Then we could simulate quantification over properties by identifying fundamental properties and relations with their extensions, and other properties with quasi-linguistic set-theoretic constructions out of fundamental properties. The conjunction of *p* and *q*, for example, could be identified with the triple  $\langle \wedge, p, q \rangle$ , where  $\wedge$  is a symbol standing for conjunction.<sup>32</sup>)

A further worry about the Lewisian approach is that mere length of definitions is an inadequate measure. Shouldn’t the degree to which a definition is “disjunctive” render the defined notion less natural? But there are strategies for refinement available here. One might, for instance, require all definitions to be in some standard form (prenex disjunctive normal form, say, if the language is a first-order predicate calculus); and one might evaluate definitions, as given in this standard form, by a more complex measure that takes more into account than length (the number of disjuncts might, for instance, be taken to count against a definition more than the average number of conjuncts per disjunct).

The first element, then, is the notion of being definitionally-more-structural-than. A second element is “lawlikeness”: the degree to which a notion figures into simple and strong generalizations. Green is more lawlike than grue because green figures in simpler and stronger generalizations than does grue. (As with Lewis’s account of laws, the relevant notions of simplicity and strength must be spelled out.)

Lawlikeness has little utility on its own, since one can cook up simple and powerful generalizations with even highly non-joint-carving notions—recall Lewis’s (1983b, p. 367) “law”  $\forall x Fx$ , which implies everything true because *F* is a predicate true only of things in the actual world. But suppose

---

<sup>32</sup>There is obviously some arbitrariness in this construction. For a defense of such arbitrariness in another context, see Sider (2006, section 2).

we restrict our attention to somewhat definitionally structural notions—green, grue, kangaroo, being-a-snail-or-a-kangaroo, and the like. With respect to such notions, lawlikeness is an interesting measure—green, for example, is more lawlike than grue. It is only when applied to *extremely* definitionally unstructural notions, such as  $F$  in Lewis’s “law”, that lawlikeness is uninteresting. (It doesn’t matter how exactly we sharpen the idea of a “somewhat” definitionally structural notion; on any reasonable sharpening, lawlikeness will generate an interesting ordering.)

The third element is really a class of elements. Philosophers of science have done much subtle and detailed work on explanation in the special sciences. Lawlikeness is one concept they use to characterize how the special sciences are explanatory, but there are many others: probability, unification, cause, and so forth.<sup>33</sup> These concepts can be understood without distinctive fundamental metaphysics; indeed, much of this work is designed to show how the special sciences can be explanatory in an ultimately physical world. So any of these concepts could play a role in a definition of a comparative notion of structure. The degree to which a notion plays a role in causal statements, for example, could be appealed to in such a definition, provided the notion of causation is itself given a metaphysical semantics.

So far the elements have been more or less “objective” (modulo some arbitrariness around the edges). A fourth element is more subjective. Perhaps economics deserves its place in the hierarchy of “levels” partly because of its value to us as a tool for prediction and control of things we care about. Disinterested Martians who for this reason do not develop economic theory would not be *missing out* in the way that they would if they did not develop chemistry or physics. In general, the comparative structuralness of a notion might be in part a function of how important it is to us. Section 2.5 argued against subjectivism about structure. But the opponent there was a *general* subjectivist about structure, including absolute structure. It would be really bad if there were no objective privilege at all to the physical conceptual scheme—recall “knee-jerk realism”. But all I am allowing here is a partial subjectivism about one facet of comparative structure. This is comparatively benign, and does not conflict with knee-jerk realism, at least of the sort I find in myself. Absolute structure provides a bedrock of objectivity, on top of which a modicum of subjectivity in comparative structure may be overlaid.

I can supply, then, these elements. A nice project would be to actually

---

<sup>33</sup>See Strevens (2006) for an overview.

produce a definition of comparative structure from them. I'd like to do this, but I don't know how. To anyone trying this at home: if you get discouraged, remember three things. First, remember that what you are attempting is a definition of a complex high-level notion in perfectly fundamental terms. No one can give a fully adequate definition of this sort for *any* complex high-level notion. A reasonable goal in such endeavors is persuading ourselves that *some* metaphysical semantics exists, by producing toy accounts that work well in a decent number of cases and have no in-principle defects. (Producing the elements of the definition is a start.) Second, don't be alarmed if a certain amount of vagueness or arbitrariness creeps into your attempts. These can be tolerated since comparative structure is not intended to itself be structural. Third, remember that there is no need to settle on a single definition once and for all; perhaps different applications call for different definitions of comparative structure. Granted, my case for primitivism about structure has been that positing a single notion of structure illuminates multiple domains (chapter 2, and see also section 7.13). But that single posited notion is absolute structure, which can be used to define multiple sorts of comparative structure.

Comparative structure, as defined from these elements, will not itself carve perfectly at the joints. But it can nevertheless carve reasonably well at the joints, provided the definition from the elements is a reasonable one. Thus its explanatory value needn't be compromised. While good explanations must be cast in reasonably joint-carving terms, they needn't be cast in perfectly joint-carving terms (there can be good explanation outside of physics, after all). Some vagueness or arbitrariness may enter the definition, which introduces the potential for nonsubstantive questions. But it does not follow that *all* questions about comparative structure are nonsubstantive. Only the questions whose answers turn on the vagueness or arbitrariness are rendered nonsubstantive. This is as it should be; some questions about comparative structure *are* nonsubstantive. For example, it may not be a substantive question whether geological notions are more structural than biological ones. It is compatible with this that the usual hierarchy of "levels" of special-science notions is by-and-large objective; the nonsubstantivity is only "around the edges".

### 7.11.2 Absolutism and infinite descent

It might be objected that absolutism requires a "ground floor". For the alternative to a ground floor is some sort of infinite descent of ever more fundamental facts or notions, and making sense of infinite descent seems to require a com-

parative notion of fundamentality.

Distinguish three sorts of “ground floor”: ideological, mereological, and propositional. Only the first is demanded by my absolutism; but the first is unobjectionable.

The first sort of ground floor is ideological. The alternative is infinite *ideological* descent—a chain of ever more fundamental notions, with no notions more fundamental than every member of the chain. My account does indeed require an ideological ground floor. For me, the facts about fundamentality are in the first instance facts of the form “Notion  $\alpha$  is structural” ( $\mathcal{S}(\alpha)$ ); and it’s hard to see how to construct a metaphysical semantics for talk of comparative notion-fundamentality in terms of such facts that would allow infinite ideological descent. (Certainly the account of definitional structure in section 7.11.1 does not allow it.) But infinite ideological descent is a seriously weird hypothesis, and seems unproblematic to deny. The hypothesis denies the existence of a book of the world, a complete perfectly fundamental description of reality, since for any chosen concepts, more fundamental concepts could be chosen.

Infinite *mereological* descent—*a.k.a.* the metaphysician’s beloved hypothesis of “gunk”—is comparatively mundane. An object is gunky if each of its parts has further proper parts; thus gunk involves infinite descent in the part–whole relation. The corresponding sense of ‘ground floor’ is mereological: atomism. My theory of fundamentality does *not* require atomism, because gunk can be described using perfectly structural ideology.<sup>34</sup> Suppose, to take a toy example, that a certain gunky patch varies continuously in color from one wavelength to another. As Frank Arntzenius and John Hawthorne (2005, section V.2) show, the color facts about this patch could be taken to emerge from the totality of facts of the form *the average wavelength of part  $x$  of the patch is  $\lambda$* . And there are various bits of fundamental ideology one could introduce to characterize these facts about average wavelength (such as ‘ $x$  has a higher average wavelength than  $y$ ’ and ‘ $x$ ’s average wavelength together with  $y$ ’s average wavelength equals  $z$ ’s average wavelength’—the details will depend on one’s general approach to quantities.) For a more physically realistic example, Jeffrey Sanford Russell (2010) shows how to characterize the topology, mereology, and measure-theoretic facts about gunky space using a finite list of primitive notions; his primitive notions could be taken to carve at the joints.<sup>35</sup>

---

<sup>34</sup>I myself reject gunk (for independent reasons); see Sider (2011).

<sup>35</sup>See also Arntzenius (2008).

The final sort of ground floor is propositional, or factual: a ground floor of facts. A propositional ground floor is opposed to infinite propositional descent—an infinite descending chain of the in-virtue-of relation over facts or propositions. Now, as we've seen, "in-virtue-of" is not my preferred way of talking about fundamentality. Still, we can ask: Can I accommodate—in terms of structure—the kinds of scenarios that would be described by the friend of in-virtue-of as involving infinite propositional descent? The answer depends on which of two kinds of scenario is alleged. In the first sort, each level in the chain is a new "sort" of proposition—a proposition involving new ideology that is more fundamental than ideology involved in propositions higher up in the chain. I cannot accommodate this scenario, since it would require infinite ideological descent. But I can accommodate a tamer scenario, in which the propositions throughout the descending chain are all of the same sort—they all involve the same ideology. For example, it might be claimed that the proposition that a certain gunky object has a certain mass is true in virtue of a proposition about the masses of its parts under some finite decomposition, that the latter proposition is true in virtue of propositions about the masses of still smaller parts under a more fine-grained but still finite decomposition, and so on infinitely:

$a$  is 1 g mass in virtue of ...

...  $b$  being 0.5 g and  $c$  being 0.5 g (where  $a = b + c$ ), which holds in virtue of ...

...  $d$  being 0.25 g and  $e$  being 0.25 g and  $f$  being 0.25 g and  $g$  being 0.25 g (where  $b = d + e$  and  $c = f + g$ ), which holds in virtue of ...

("+" signifies fusion). Even though this scenario would be described by the friend of in-virtue-of as one of infinite propositional descent, my description of it does not require infinite ideological descent. I can simply state all the propositions involved in the chain:

$a$  is 1 g

$b$  is 0.5 g

$c$  is 0.5 g

$a = b + c$

etc.

leaving out the in-virtue-of claims (since I renounce them), and add the only relevant claim about fundamentality, which is that the ideology common to all the propositions—namely, mereological and mass-theoretic—is absolutely structural. (The exact nature of the mass-theoretic ideology common to all the propositions depends on one’s views about the nature of quantity. One view would be that the ideology is comparative: “ $x$  is more massive than  $y$ ”, or “ $x$ ’s mass together with  $y$ ’s mass equals  $z$ ’s mass”.)

So: the only limitation stemming from the absoluteness of structure is that I cannot accommodate infinite ideological descent. This is no real limitation, I say, because there is no reason to suppose that this weird scenario obtains.

“Objection: there *is* reason, namely empirical inductive reason, to believe in infinite ideological descent. Physicists once thought that everything depended on the features of molecules. But molecules gave way to atoms, which gave way to protons, neutrons and electrons, which have given way to quarks, leptons, and gauge bosons. Each time a new type of particle was discovered, physicists posited new features of the newly discovered particles, whose distribution accounted for, but could not be accounted for in terms of, the distribution of the distinctive features of the older particles. This historical progression of theories will probably continue forever, so there are no ultimate features on which everything depends.”—This is a bad argument, for a few reasons. First, it is an induction from only four cases. Second, by moving from “finite” observations to an “infinite” conclusion, the argument makes a big leap. Compare it to the argument that there must be infinitely many people, since for each person we’ve observed, there exists a taller person.<sup>36</sup> Third, the argument fails by drawing a conclusion that is drastically dissimilar from the initially observed pattern. The initially observed pattern is an historical progression of physical theories:

Theory 1: The fundamental features are those of molecules.

Theory 2: The fundamental features are not those of molecules, but are rather those of atoms.

Theory 3: The fundamental features are not those of atoms, but are rather those of protons, neutrons, and electrons.

---

<sup>36</sup>Thanks to Cian Dorr here.

Theory 4: The fundamental features are not those of protons, neutrons, and electrons, but are rather those of quarks, leptons, and gauge bosons.

The conclusion drawn is that there are no fundamental features, since for any physical feature had by any particle, there are further features (had by smaller particles) that do not depend on the first feature. But this conclusion isn't inductively suggested by the initial pattern. The conclusion has the superficial appearance of a kind of limit point of the initial pattern, if that pattern were infinitely extended. By moving through Theories 1–4, so the idea goes, scientists have been moving closer and closer to the conclusion. But this impression vanishes upon closer inspection. Each Theory in the progression does not *add* a new layer of fundamental features, but rather *ditches* the previous Theory's layer (since it regards the previous layer as just depending on the newly hypothesized layer). Extending the pattern indefinitely results in a series that simply has no intuitive limit. For comparison, imagine a countably infinite series of chairs:  $c_1, c_2, \dots$ . Suppose first that in scenario 1,  $c_1$  is filled; in scenario 2, chairs  $c_1$  and  $c_2$  are each filled; in scenario 3, chairs  $c_1, c_2$ , and  $c_3$  are each filled; and so on. I suppose there's some sense in which the limit of this series is a scenario in which all the chairs are filled. But consider a second series in which only  $c_1$  is filled in scenario 1, *only*  $c_2$  is filled in scenario 2, *only*  $c_3$  is filled in scenario 3, and so on. This series has no intuitive infinite limit. The imagined infinite extension of the progression through Theories 1–4 is like the second series.

This third criticism of the inductive argument depends on the fact that I construed Theories 1–4 in terms of "fundamental features", by which I meant absolutely fundamental features. Suppose they were construed instead in terms of comparative fundamentality:

Theory 1a: Molecules have certain distinctive features.

Theory 2a: Atoms have certain distinctive features, which are *more fundamental than* those of molecules.

Theory 3a: Protons, neutrons, and electrons have certain distinctive features, which are more fundamental than those of atoms.

Theory 4a: Quarks, leptons, and gauge bosons have certain distinctive features, which are more fundamental than those of protons, neutrons, and electrons.

If continued infinitely, this progression *does* seem to have an infinite limit (it's like the first chairs series): that for every feature, there are more fundamental features—*infinite ideological descent*. But this is not dialectically effective against absolute fundamentality, since comparative fundamentality was assumed from the start, in the characterizations of Theories 1a–4a.

“Even if there is in fact no infinite ideological descent, infinite ideological descent is *epistemically possible*, and hence should be allowed by any good theory of fundamentality.”—If infinite ideological descent is epistemically possible then my theory of fundamentality is not epistemically necessary. That's ok! My theory is intended to be an educated guess about the nature of the world, not as some sort of a priori deduction that must hold with certainty. Neutrality on “first-order” questions like that of infinite ideological descent is not a reasonable constraint on the metaphysics of fundamentality. A metaphysics of fundamentality is supposed to give the truth about the nature of fundamentality, not provide a dialectically neutral framework in which to conduct first-order debates.

“Even if infinite ideological descent is not actual, it is nevertheless *metaphysically possible*, so a good theory of fundamentality should permit it.”—As we will see in section 12.5, the Humean theory of modality to be defended undermines such arguments from possibility. In brief, the impossibility of infinite descent amounts to little more than its nonactuality, so there is no distinctively modal way to support it.

## 7.12 Determinate?

In addition to being pure, complete, subpropositional, and absolute, I hold that the fundamental is also determinate.

“The fundamental is determinate” is not particularly clear, and improving on the situation is difficult because there are so many different ways to understand what “determinacy” amounts to, but perhaps we can put it thus. First, no special-purpose vocabulary that is distinctive of indeterminacy—such as a determinacy operator or a predicate for supertruth—carves at the joints. Second, fundamental languages obey classical logic. The combination of these two claims is perhaps the best way to cash out the elusive dogma that vagueness and other forms of indeterminacy are not “in the world”.

This is not to deny the value of determinacy-theoretic vocabulary, or supervaluationism, or nonclassical logic; it is just to deny them a place at the

fundamental level. They might yet play a role in explaining vagueness in nonfundamental languages (see section 10.6).

## 7.13 Fundamental?

Last question: is fundamentality fundamental?

There are two questions here. First, is fundamentality a fundamental *notion*? And second, are facts about fundamentality fundamental *facts*? Given my subpropositional approach to fundamentality, the first question is primary. In my terms it is the question of whether structure is itself structural—of whether carving at the joints carves at the joints. In the official regimentation: is it the case that  $\mathcal{S}(\mathcal{S})$ ?<sup>37</sup> My answer is yes.

My answer to the second question is: not all of them. In my terms, the question is whether facts about structure are fundamental facts. Facts about structure that involve only structural notions are indeed fundamental facts; but facts about structure that involve nonfundamental notions cannot be fundamental facts, given purity. For example, the fact that grue does not carve at the joints involves the nonfundamental notion grue, and so cannot be a fundamental fact. Thus it must, given completeness, have a metaphysical truth-condition stated in purely fundamental terms. But that truth-condition can mention structure, since structure is a fundamental notion. The truth-condition might have the form “ $G$  does not carve at the joints” (officially: “not  $\mathcal{S}(G)$ ”), where  $G$  is a “metaphysical definition” of ‘grue’.

Back to the first question. My reasons for saying that structure is structural emerge from considering an opposing viewpoint.

A vivid test for whether a given expression,  $E$ , carves at the joints is this: did God need to think in  $E$ -terms when creating the world? Clearly, she needed to think in terms of quantification, mass, distance, and so on; accordingly, those notions carve at the joints. But did she need also to think in terms of structure? It is natural, I must admit, to say no. All she needed to do was decide which objects to create, how massive to make them, and where to put them; she didn’t need also to consider whether quantification, mass, distance, and the rest were

---

<sup>37</sup>On some regimentations of talk of structure, one might raise a worry about the intelligibility of the question: “Being structural is a property of semantic values; the question would be about the proposition that the property of being structural instantiates itself; there is no such proposition.” I’m not sure if I accept any part of this objection, but at any rate, my preferred regimentation for talking about structure avoids the worry because  $\mathcal{S}$  is an operator.

structural.

If structure is not structural, then completeness requires all statements about structure to have metaphysical truth-conditions. Since those truth-conditions must be stated in perfectly structural terms, they cannot contain ‘structure’; but they can contain terms that are structural—terms of physics, perhaps. What might such truth-conditions look like? I will sketch an answer that I will call “Melianism”, since it contains elements of an intriguing view due to Joseph Melia.<sup>38</sup>

Let’s simplify by discussing Lewisian (perfect) naturalness, rather than the more general notion of structure. We are after a definition of ‘natural’ in terms of natural properties and relations (by hypothesis these do not include naturalness itself). Where  $N_1, N_2 \dots$  are the natural properties and relations, the Melian definition is:

$$P \text{ is natural} =_{\text{df}} P = N_1 \text{ or } P = N_2 \text{ or } \dots$$

This is a highly disjunctive definition. The Melian embraces this. There is no need for a nondisjunctive, explanatory notion of naturalness, he says, because naturalness itself is never invoked in explanations. Whenever Lewis would cite naturalness in an explanation, the Melian cites particular natural properties. Why are these two electrons exactly alike? Because they have exactly the same natural properties, Lewis says. The Melian says instead: because they both have charge  $c$ , mass  $m$ , and spin  $s$ .

Objection: “We may not know which natural properties the electrons possess.”—The Melian reply is that this is an epistemic limitation of ours, not a deficiency in the proffered explanation. The best explanation of the electrons being duplicates cites the particular natural properties they share, even if we do not possess that explanation.<sup>39</sup>

Objection: “The Melian’s definition is circular since ‘natural’ was used to pick out the list  $N_1 \dots$ ”—There would be objectionable circularity only if ‘natural’ occurred on the right hand side of the definition. But it does not. ‘Natural’ was used to pick out the list by description; and perhaps our only access to the list—and thus to the definition—is via this description. Nevertheless, the definition itself contains only the list, not the description. Further, the Melian could avoid appealing to ‘natural’ in even this indirect way if he accepted the epistemology of section 2.3. He could then pick out the list by the description

---

<sup>38</sup>Melia put this forward in conversation, so please don’t blame him for what follows!

<sup>39</sup>Compare Melia (1995, 2000).

‘properties that figure in our best theory’. (This would not be an identification of naturalness with the property of figuring in our best theory. Naturalness would remain identified with a disjunctive property of the form *being N<sub>1</sub> or N<sub>2</sub> or ...*; the description would represent our best guess as to the identity of the disjuncts.)

“Objection: Melianism diminishes the significance of duplication, intrinsicality, lawhood, and other notions defined in terms of naturalness.” Here we approach a more telling objection. Since Melian naturalness is highly disjunctive, so will be the defined notions. So although the Melian agrees with Lewis on the first-order questions—on which objects are duplicates of which, which properties are intrinsic, what the laws are, and so on—he must regard Lewis’s focus on *these* first-order questions—questions of duplication, intrinsicality, law, and so forth—as being arbitrary, because based on highly disjunctive notions. This is a strange predicament. The Melian is trying to achieve Lewis’s aims on the cheap, but his theory implies that these aims are metaphysically arbitrary and not particularly worth pursuing.

And the predicament is worse than strange; Melianism undermines all of the applications of naturalness. The Melian admits that a notion so disjunctive as Melian-naturalness cannot be explanatory, and tries to get around this by claiming that explanations can always cite particular natural properties rather than naturalness itself. We can explain why two electrons are exactly alike, he said, by pointing out that each has charge *c*, mass *m*, and spin *s*. But that is no explanation; explanations must cite generalizations. Explanations of similarity-facts require generalizations about similarity; explanations of meaning-facts require generalizations about meaning; explanations of substantivity-facts require generalizations about substantivity; and so on. Moreover, explanations must cite generalizations of sufficient scope. It would be no good to cite the generalization that any two electrons sharing charge *c*, mass *m*, and spin *s* are similar; the generalization is too specific. The generalizations must cite naturalness—or better, structure. But then structure cannot have a Melian definition, if the generalizations are to be explanatory.

The argument of this book is that the explanatory power of our overall theory is enhanced by positing structure. Must the posited notion be itself (perfectly) structural, or could it have a definition in structural terms? What we have seen so far is that it could not have a Melian, disjunctive definition, for that would undermine its applications.

Could it have a non-Melian, nondisjunctive definition? Then even though structure would not be structural, it might still be capable of figuring in expla-

nations, just like notions of the special sciences.

But it is hard to see how a definition of structure could avoid being disjunctive. Chemical kinds have nondisjunctive (metaphysical) definitions because instances of chemical kinds are reasonably physically alike, and physical notions are structural. Biological kinds have nondisjunctive definitions because instances of biological kinds are functionally, if not physically, alike; and functional notions have reasonably nondisjunctive definitions in structural terms. But consider the instances of the notion of structure. These include notions of mass, charge, spatiotemporal distance, set membership, conjunction, disjunction, and universal quantification, let us suppose. This class of notions is neither physically nor functionally unified. Nor does it seem to be unified in any other way that would allow a nondisjunctive definition.<sup>40</sup> Other than the fact that all its members are structural, the class is highly heterogeneous.

Relatedly, consider the notions to be defined in terms of structure—similarity, intrinsicality, laws and explanation, meaning, induction, physical geometry, and substantivity. The instances of any one of these notions are heterogeneous in physical and functional and other “first-order” ways. When  $x$  is similar to  $y$  and  $x'$  is similar to  $y'$ , there needn’t be any physical or functional or other first-order commonality between the pairs  $\langle x, y \rangle$  and  $\langle x', y' \rangle$ ; there needn’t be any physical or functional or other first-order commonality between any two laws of nature; and so on. What unifies all the pairs of similar objects, and all the laws, is just the fact that they involve structural notions in certain ways; and the structural notions themselves are neither physically nor functionally unified.

The “first-order heterogeneity” of structure, and of structure-involving notions, is an in-principle obstacle to a nondisjunctive definition of structure. Thus the choice is stark: either adopt extreme realism about structure—holding that structure is itself structural—or else give up altogether on explanations that invoke structure, which is tantamount to giving up on structure itself. My choice is for the former.

The status of metaphysics itself hangs on this choice. In their loftiest moments, metaphysicians think of themselves as engaged in a profoundly important and foundational intellectual enterprise. But if fundamentality is highly disjunctive, the field of metaphysics itself—which is delineated by its focus on fundamental questions—would be an arbitrarily demarcated one.

---

<sup>40</sup>Or better: that would allow a nondisjunctive and objective definition. The class does seem unified by the fact that its members are all indispensable in our best theories.

Although it offers explanatory power (and a pleasing self-conception for metaphysicians), extreme realism about structure raises some difficult questions. I'll mention three.

First, the argument for saying that structure is structural was that this is needed to insure that structure can take part in genuine explanations. But look at the notions other than structure involved in the putative explanations: simplicity, correctness of interpretation, candidate meaning, and so on. These do not seem structural either. Doesn't this already undermine the genuineness of the explanations? And if it doesn't—if the genuineness of the explanations is compatible with their involving nonstructural notions—then why can't the explanations remain genuine if structure isn't structural?

This can be answered. Genuineness of explanation does not require *perfectly* structural notions, as we see from the special sciences. It is enough that simplicity, correctness of interpretation, and the rest, are somewhat structural. The reason for thinking that structure cannot be merely somewhat structural is its first-order heterogeneity—if structure is not perfectly structural then it is disjunctive and therefore highly nonstructural.

The second question I find more challenging. In section 6.3 we considered two possibilities for regimenting talk of structure, each involving the operator  $\mathcal{S}$ . On one,  $\mathcal{S}$  had a very flexible grammar, and on the other, the language needed to be supplemented with dummy variables. Neither smacks of fundamentality. Each seems to require our fundamental languages to be much more complex than they would otherwise need to be. The complexity could be avoided by taking talk of semantic values as fundamental;  $\mathcal{S}$  could then be a predicate of semantic values. But semantics does not smack of fundamentality.

The third question is also challenging. Realism about structure requires a fundamental posit, and such posits are generally to be avoided. The concern is particularly pressing given my own preference, expressed many times in this book, for simplicity. I have been disdaining, and will continue to disdain, primitive modality, law, cause, tense, logical consequence, higher-order quantification, and other such luxuries. But when it comes to my own pet concept, structure—it might be alleged—my scruples go out the window.

It's not that I have no answer to this charge. My answer is that structure is no luxury, since it cannot be reduced without loss. Still, it smells fishy, doesn't it? This is a serious challenge facing the audacious doctrine of realism about structure.