

Inferentialism and Some of Its Challenges

ROBERT BRANDOM
University of Pittsburgh

I. Some methodological preliminaries

Our ordinary ways of talking and thinking about the contentfulness of linguistic utterances and mental states distinguish between what is said or thought and what we are thereby talking or thinking *about*.¹ Intentionality, we may say, comes in two flavors: ‘that’ intentionality and ‘of’ intentionality. A central semantic task in the area where the philosophy of language and the philosophy of mind overlap is accordingly to offer an account of how these two dimensions of sapience are related to one another.

One prominent explanatory strategy begins by explaining what it is for something to represent something else: paradigmatically what it is for a singular term to pick out an object, and a predicate to pick out a property of (in a different sense of ‘of’) or relation among objects. One then explains what it is for sentential constellations of those representing elements to be *true* in terms of set-theoretic inclusion relations among the various represented items. Finally, one explains the propositional content expressed by sententially shaped or labeled representings by modalizing the truth assignments of the previous step: taking those conditions to be the ones under which those representings *would* be true. This familiar sort of approach results from combining two prioritizing commitments. On the one hand, with respect to the two semantic dimensions just pointed out, it moves from a story about what is represented to one about what is expressed—roughly, from a theory of reference to a theory of sense. On the other hand, it proceeds categorially

¹ They show up in the technical idiom of Heidegger’s *Sein und Zeit* as ‘das Beredete’ and ‘das Geredete’. See, e.g., J. Macquarrie and E. Robinson (trans.) *Being and Time* (New York: Harper and Row, 1962), p. 206.

in a bottom-up direction: from the contents of subsentential expressions to those of sentential ones.

Of course, one might go about things differently. A semantically and categorially converse strategy, to pick one, starts with a notion of the propositions expressed by whole sentences. It might, for instance, identify those contents—what can be said or thought—with sets of possible worlds. At the next stage, it seeks to understand the contributions made to the specification of such contents by the subsentential expressions deployed in the sentences that express them. For example, singular terms might be associated with functions whose arguments are possible worlds and whose values are objects in those worlds. Then, if an additional step is needed to get there, an account would be offered of what is picked out, referred to, or represented by those subsentential contents. Thus in each world, one might identify what is represented as the object assigned to that world by the function associated with the representing singular term. Other sorts of approach are also possible.²

It is possible to assemble considerations that speak, *ceteris paribus*, for or against the various methodological approaches botanized by this two-by-two semantic/categorial characterization. Thus the pragmatic priority of sentences—the thought that leads Kant to treat the judgment as the smallest unit of awareness, Frege to prioritize what can be true since it is the smallest unit to which pragmatic force can be attached, and Wittgenstein to pick out sentences as the smallest unit whose use counts as making a move in a language game—offers some reasons in favor of categorially top-down orders of explanation. The need to explain our capacity to produce and understand indefinitely many novel thoughts and the sentences that express them offers a countervailing reason to prefer a compositional, categorially bottom-up strategy. Broadly cognitive issues of what it is to grasp a content or understand a meaning, as well as functional ones concerning what it is to use an expression as a representation, or for it to function in some system as such a representation speak for an approach that focuses on the first, sense-like sort of semantic content. But the need to explain why the liability of some sayables or thinkables to assessment of their truth is essential to their having the contents they do—or indeed, any contents at all—offers grounds for treating the representational or referential semantic dimension as central.

I doubt that decisive arguments for or against any of these orders of semantic explanation are available at this level of methodological abstraction. The considerations on offer, while genuine and important,

² Stalnaker, for instance, in *Inquiry* (Cambridge, MA: MIT Press, 1984) offers a categorially top-down story of roughly this shape in which the top-level semantic notion is of the *representation* of a set of possible worlds.

offer at best probative reasons, not dispositive ones. I suppose that understanding will best be served by patient working-out of more specific semantic theories pursuing the various kinds of explanatory strategy, and their rigorous assessment according to the criteria of adequacy that provide the motivations that have been forwarded for preferring one sort of approach over another in advance of looking at such detailed proposals. It is in that spirit that I want to consider, in this paper, a number of objections that suggest themselves to the kind of semantic theory I have called ‘inferentialist’.

II. What Is inferentialism?

Logical empiricism revived classical empiricism by appealing to the new quantificational predicate logic Russell had developed out of Frege. Its motivating thought was that this way of understanding the inferential articulation of the immediate deliverances of sense provided powerful new expressive tools to put in the place of traditional appeals to processes such as association and abstraction, which had proven themselves woefully inadequate to rendering the contents of interesting empirical concepts—never mind mathematical ones. It is a striking fact about the contemporary scene that two broad classes of theories of concepts (I’ll discuss a third in a moment) correspond to these two dimensions—sensuous and logical—into which the logical empiricists sought to factor conceptual content.

One popular strategy looks to the *observational* use of concepts as the key to conceptual content. Here one thinks of the use of ‘red’ or ‘square’ as non-inferentially elicited as a response, typically, to red or square things. The focus is accordingly on the reliable differential responsive dispositions linking, say, tokenings of ‘horse’ to horses. Fodor’s and Dretske’s semantic theories are principal examples of this class of approach. These deserve to be seen as contemporary descendants of classical *empiricist* theories of content. Another strategy is to look to the contents of *logical* concepts as providing the key to understanding conceptual content generally. Here the idea is to generalize Gentzen-style specifications of the meanings of logical connectives by pairs of introduction and elimination rules to notions of the circumstances and consequences of application of an expression. Dummett is the principal figure in this tradition, in which he is followed by others such as Peacocke, and myself in *Making It Explicit*. These deserve to be seen as offering specifically *logical* versions of traditional *rationalist* theories of content. Each is a self-consciously one-sided approach, by contrast to the even-handed appeal to both observation and logical inference in virtue of which Carnap’s neo-Kantian roots become evident.

The genus of which semantic inferentialism is a species then has these features. It is:

- a) categorially sententialist or top-down,
- b) expressive, or sense-based, and
- c) rationalist in its choice of conceptual paradigm,

in contrast to theories that are either categorially bottom-up, representational or reference-based, and empiricist in their choice of conceptual paradigm. Any theory of this genus faces three structural demands. Along the categorial dimension, it must show how to assign contents to subsentential elements. Along the semantic dimension, it must show how to underwrite a notion of reference or representation. And it must show how to model conceptual content generally, including especially the content of observational concepts, on that of the logical concepts it treats as paradigmatic.

What distinguishes inferentialist semantics within this genus is the concept it proposes to explore as a candidate for the notion of sentential sense treated as central by accounts of that genus and motivated by the case of logical concepts: inferential role. The idea is to understand propositional contents as what can both serve as and stand in need of *reasons*, where the notion of a reason is understood in terms of inference. So propositional contentfulness is taken to be a matter of being able to play the role both of premise and of conclusion in inferences. Once the notion of introduction and elimination rules as exhaustively constitutive of the content of logical concepts has been generalized to take in the circumstances and consequences of application of non-logical concepts, the step to inferentialism is taken when one understands their content as exhaustively constituted by the material, non-logical *inferential* connection between those circumstances and consequences. The content of a concept such as temperature is, on this view, captured by the constellation of inferential commitments one undertakes in applying it: commitment, namely, to the propriety of all the inferences from any of its circumstances of appropriate application to any of its appropriate consequences of application.

Inferentialism is not the only way to try to develop an account that takes the contents of logical concepts as paradigmatic. (Peacocke, for instance, takes a somewhat different tack.³) But one might well ask what motivation there is for adopting this paradigm at all. Granted

³ In *A Study of Concepts* (Cambridge, MA: MIT Press, A Bradford Book, 1995).

(though there is, to be sure, no unanimity even on this point) that Gentzen-style definitions offer us impressively clear and demonstrably useful specifications of the meanings of logical connectives, why should we think that that model can helpfully be generalized from logical to non-logical, paradigmatically empirical concepts? The drunk's reason for looking for his lost keys under the streetlamp—that the light is better there—notoriously does not provide a reason for thinking the keys are likely to be found there. Why doesn't a corresponding criticism apply to contemporary logical-rationalist approaches in semantics?

One reason is provided by looking at the motivations for a third contemporary candidate for a privileged subclass of conceptual contents on which to model the rest: *modal* concepts. For logical empiricism, like its classical Early Modern ancestors, also signally failed to render the contents of ordinary concepts. One retrospectively obvious reason is its lack of expressive resources sufficient to render the distinction between accidental generalizations and lawlike regularities—the very incapacity Kant had diagnosed already as fatal to the classical version of empiricism. The advent of a technically adequate semantics for modal concepts put philosophers in a position to create a third wave of empiricism. It had many of the same basic motivations and aspirations as the logical empiricism of the middle third of the twentieth century, but could use more powerful modal logical apparatus in place of the extensional logic previously appealed to as providing the logical cement binding together and articulating the sensuous content provided by perception. David Lewis may be taken as an index figure of this movement.⁴

What I want to focus on for the moment is the appreciation of the significance of what might be called 'non-Tractarian' concepts—paradigmatically modal, probabilistic, and normative ones—for our understanding of conceptual content generally. In an autobiographical sketch, Sellars says the central idea motivating his work was one that occurred to him already in the 1930's: that the centrality of modal concepts in formulating empirical ones—the insight he put into the lucid title of one of his most impenetrable essays "Concepts as Involving Laws, and Inconceivable Without Them"—meant that

What was needed was a functional theory of concepts which would make their role in reasoning, rather than their supposed origin in experience, their primary feature.⁵

⁴ At the level of generality in play in this telling of the story. Of course there is *lots* more going on in Lewis than this—but Fodor is not *just* an empiricist either.

⁵ In H.-N. Castañeda (ed.) *Action, Knowledge, and Reality* (Indianapolis: Bobbs-Merrill, 1975), p. 285.

The expressive role distinctive of modal vocabulary is to make explicit the distinction between counterfactually robust inferences and those that are not—a distinction without which, Sellars reminds us, following Kant, we must fail to understand not only the content of such theoretical empirical concepts as rigidity or mass, but also such observational ones as red or horse. It is not at all clear how these modal features of empirical concept use might be understood in terms of chains of reliably covarying events linking horse-stimuli to ‘horse’-responses. In this connection it might be noticed that although the liberal use of modal vocabulary in the metalanguage in which they conduct their explanatory enterprise is of the very essence of Dretske’s and Fodor’s semantics for the observational concepts they treat as paradigmatic of concepts in general, neither one so much as attempts to reconstruct within their theory the contents of those modal concepts. And the recalcitrance to empiricist reduction of what is expressed by modal vocabulary applies, *ceteris paribus*, equally to probabilistic and normative vocabulary. It provides the motivation for Sellars’s inferentialism.

III. Varieties of Inferentialism

Inferential approaches to what is expressed by sentences can take various forms. What may be called ‘*weak*’ inferentialism claims only that the inferential connections among sentences are *necessary* for them to have the content that they do, in the sense that unless at least some of those inferential involvements were as they are, the sentence would mean something different. By ‘inference’ in such a formulation is meant *correct* inference: the ones people *ought* to make, rather than the ones they are actually disposed to make, if those two notions fall apart. The claim of weak inferentialism, so understood, ought not, I think, to be controversial. For even those who understand what is said by sentences in terms ultimately of what is represented by their subsentential components take it that their representational content *determines* the proprieties of inferences they are involved in. So at least some of the inferences are such that if they were not correct, the sentence would mean something else. For only if they had different truth conditions—for instance—could the sentences be understood as playing different roles as premises and conclusions of good inferences.

At the other end of the spectrum is what might be called ‘*hyperinferentialism*’. It is the claim that the inferential connections among sentences, *narrowly construed*, are *sufficient* to determine the contents they express. By ‘narrowly construed’ in this formulation is meant that only correct inferences in which the sentence plays the role of premise or conclusion are to be considered. Hyperinferentialism is—Gentzen

claims, and I think we ought to agree—true for *logical* vocabulary. The introduction rules for logical connectives appeal only to inferential grounds for applying them (as principal connectives) that are sentences in which the connective being defined or introduced does not occur (as principal connectives). And their elimination rules appeal only to inferential consequences of applying them (as principal connective) in which the connective in question does not occur (as principal connective). But hyperinferentialism is extremely implausible as applied to other sorts of vocabulary, especially to vocabulary that has *observational* uses that are essential to its meaning. The meaning of ‘red’ is not just a matter of what other concepts its applicability is properly inferrable from (e.g., scarlet), and the applicability of what other concepts is properly inferable from its applicability (e.g., colored), nor of the applicability of what other concepts preclude its applicability (e.g., green or multiple of 3). Its *noninferential* applicability to *red things* is also an essential propriety of the use of the concept red: one that must accordingly be underwritten by any adequate account of the meaning or content expressed by the use of the word ‘red’. If taking the meanings expressed by logical vocabulary as a model for the semantics of other sorts of expressions could take the form only of commitment to hyperinferentialism, then there would be little reason to take this explanatory strategy seriously.

But there is a *via media* between the uncontroversial but unhelpful truism of weak semantic inferentialism and the powerful and interesting but unsustainable (outside the realm of logical and perhaps some mathematical discourse) thesis of semantic hyperinferentialism. What may be called ‘*strong*’ inferentialism claims that the inferential articulation of concepts, *broadly* construed, is *sufficient* to determine their contents. By ‘broadly construed’ in this formulation is meant three things. First, the inferences in question must be understood to extend beyond *logically* or *formally* good ones—those whose correctness is settled just by the logical form of the sentences involved. They must include also those that are *materially* correct—that is, those that intuitively articulate the contents of the *nonlogical* concepts involved. Sellars offers as examples the inference from “A is to the East of B,” to “B is to the West of A,” and from “Lightning is seen now,” to “Thunder will be heard soon.” Second, besides material inferential relations among sentences in the sense of their proper role as premises and conclusions, material *incompatibilities* among sentences, which underwrite inferential relations in a narrower sense, are included. Thus the fact that the correct applicability of square precludes the correct applicability of triangular, so that the inference from square to not-triangular is a good one, is also to be considered.

Third, and most important for understanding the difference between the hyperinferentialist and strong inferentialist theses in semantics, is that inferential relations between *noninferential* circumstances of appropriate application and *noninferential* appropriate consequences of application are also taken into account. The way the Gentzen hyperinferentialist model for the semantics of logical concepts is to be extended is by taking seriously the thought that in using *any* expression, applying *any* concept, one is undertaking a commitment to the correctness of the (in general, material) inference from the *circumstances* in which it is correctly applied to the correct *consequences* of such application. And this is so even where some of those circumstances or consequences of application are *noninferential*. Thus the visible presence of red things warrants the applicability of the concept red—not as the conclusion of an *inference*, but *observationally*. And the point is that the *connection* between those circumstances of application and whatever consequences of application the concept may have can be understood to be *inferential* in a broad sense, even when the items connected are not themselves sentential. In a culture in which **white** is the color of death, and things associated with death are to be shunned or avoided—a culture, to be sure, that would mean something somewhat different than we do by their word corresponding to our ‘white’—the connection between the visible presence of white things and the practical response of shunning or avoiding, which their practitioners endorse by using the concept in question, is an inferential one in the broad sense in question here.

It is strong semantic inferentialism that is articulated and endorsed in *Making It Explicit*. The two key moves in extending the inferential approach beyond its paradigmatic application to logical concepts are first, looking to *material* inferential and incompatibility relations, and second taking into account the *inferential* relations linking circumstances and consequences of application, even where these are *noninferential* circumstances or consequences.⁶ The generic

⁶ In fact, it is also an essential move to replace the undifferentiated notion of propriety with the more articulated distinction of two flavors of deontic status: commitment and entitlement. Doing that permits one to define incompatibility (two claims are incompatible if commitment to one precludes entitlement to the other), and three sorts of inferential relations: commitment-preserving inferences (a generalization to the case of material inferences of *deductively* good inferences), entitlement-preserving inferences (a generalization to the case of material inferences of *inductively* good inferences), and incompatibility entailments (a generalization to the case of material inferences of *modally robust*, i.e., counterfactual-supporting, inferences). See Chapter 6 of *Articulating Reasons* (Cambridge, MA: Harvard University Press, 2000) for a rationale.

location of this approach in the botanization presented in the first two sections of this paper—as categorially top-down, and semantically expressive rather than representational—then dictate the principal constructive obligations of inferentialist approaches in semantics. It must be extendable somehow to *subsential* expressions, such as singular terms and predicates, quantifiers and so on, which cannot play the directly inferential roles of premise or conclusion in inferences. And it must somehow underwrite assessments of the *representational* content of expressions of all categories: the truth of sentences and the reference or denotation of terms, predicates, and so on. In *Making It Explicit*, the first of these tasks is addressed by using Frege’s methodology of marking *inferential* invariances under *substitution* of one expression for another. Roughly, two subsential expressions play the same *indirectly* inferential role just in case substituting one for the other never turns a good inference (in the broad sense) into a bad one. The second task, offering an account in inferential terms of the representational dimension of content, is addressed (in Chapter Eight of *Making It Explicit* and Chapter Five of *Articulating Reasons*) by appealing to the different *social perspectives* corresponding to the difference between the *practical deontic attitudes* of *undertaking* a commitment (oneself), and *attributing* it (to another), which are made explicit in the difference between *de re* and *de dicto* ascriptions of propositional attitude.

This constructive task of semantic theories of inferentialist shape—roughly, underwriting representational semantic characterizations in terms of inferential ones, or understanding reference in terms of inference—points to a further division. For another important way in which it is useful to characterize varieties of semantic inferentialist programs concerns the nature of their methodological aspirations. The big distinction in this vicinity is that between *reductive* versions of inferentialism and *expressive* ones. As I would understand it, reductive inferentialism would claim first that there can *be* expressions or intentional states standing in inferential relations and playing inferential roles, hence having conceptual content, without yet standing in representational ones, and second, that one can then *build* representational relations and roles, and so content, out of those inferential ones. Compare, for the case of the converse order of explanation, a story where representational relations are defined in terms of nomological relations between representings and representeds, and a story is told only *much* later about how they must interact to yield representations of states of affairs that are truth-evaluable (hence believable), and so inferentially related.

Expressive inferentialism, by contrast, is a claim about *understanding* inferential and representational relations. It is at the level of the *senses* of the concepts inference and representation, rather than at the level of their *referents*. The expressive inferentialist acknowledges that nothing can stand in genuinely *inferential* relations unless it also has *representational* content. There is no inference without reference. But it is claimed nonetheless that one can specify sufficient conditions for expressions to be used so as to possess conceptual content (of *both* sorts) in a purely *inferential* metalanguage. Seeing what it is about this inferential articulation that amounts to the possession of representational content is then explaining what is expressed by the representational semantic metalanguage (which turns out always already to have been in principle applicable) in the terms of the inferential semantic metalanguage. So one of the most important concluding moves developing the expressive inferentialism in *Making It Explicit* (reprinted in Chapter 5 of *Articulating Reasons*) is an account of the broadly inferential role locutions must play in order for them to mean what ‘of’ and ‘about’ do when expressing the representational dimension of intentionality—the uses of ‘of’ and ‘about’ epitomized by “thinking of a horse” and “talking about colors,” rather than “the pen of my aunt,” and “weighing about five pounds.”

IV. Which Inferences?

It is obvious that a key question for strong expressive inferentialism—as for any species of this genus of semantic explanatory strategy—is (as Fodor and Lepore put it in a recent article⁷) “Which inferences are meaning-constitutive?” I want to discuss briefly two kinds of response to this question (without meaning to deny that other avenues are also open). A natural way into the issue is provided by one of Quine’s central challenges in “Two Dogmas of Empiricism.” Transposed from the idiom of analytic *truths*—sentences whose truth is underwritten by the meanings of the nonlogical expressions occurring in them—to one of, as it were, analytic *inferences*, i.e., those whose correctness is underwritten by the meanings of the nonlogical expressions occurring in them, the challenge is this. Semantic theories associate contents (or other semantic interpretants) with expressions, performances, or states in order to explain or at least codify proprieties of the *use* of those expressions, the practical *significance* of those performances, or the proper *functioning* of those states. Specifications of the distinction between correct and incorrect uses accordingly stand

⁷ “Brandom’s Burdens: Compositionality and Inferentialism,” *Philosophy and Phenomenological Research* Vol LXIII, No. 2, September 2001, pp. 465-483.

to attributions of content very much as statements in an observation language stand to statements in a theoretical language in ordinary empirical science: the *point* of the theory is to explain what is observable, and the point of semantics is to explain practice. So if one of the principal moves proposed by a semantic theory is to distinguish two flavors of inference (or truth), namely those that are constitutive or expressive of meaning, and those that are good rather in virtue of something else, we are entitled to ask what feature of the *use* of expressions with those meanings or contents reflects this semantic difference. (One need not be committed to semantic instrumentalism to ask this question, any more than one need be committed to instrumentalism more generally in order to ask what observable phenomena are explained by or manifest an hypothesized, unobservable, theoretical state of affairs.)

Quine, of course, surveys a number of candidates such as unrevisability, and finds all of them wanting. He famously concludes that postulating meanings, and thereby committing oneself to some inferences (truths) having a semantic privilege that not all do, can do no explanatory work. Sellars offers a direct response to this challenge. The practical difference that Quine rightly demands is, according to Sellars, just the difference between counterfactually robust inferences and those that are not. The inference from something's being copper to its melting at 1083.4°C is partly constitutive of the concept copper because *if* the coin in my pocket *were* copper, it *would* melt at that temperature. The inference from the coin's being in my pocket to its being copper is *not* partly constitutive of the concept copper because *if* this nickel *were* in my pocket, it would *not* follow that it was copper. This is a directly responsive answer to Quine's challenge, because we do in our ordinary linguistic practice distinguish between inferences based on their modal status as counterfactual-supporting or not—between those that would explicitly be licensed by law-like regularities such as connect atomic structure and melting point, and those that would be licensed only by accidental regularities, such as those connecting atomic structure and location in space. For Sellars, conceptual connections are just the lawful ones. (That is why “concepts involve laws and are inconceivable without them.”) His was the first explicitly *modal* theory of meaning.

To be sure, this approach has some radical consequences. If we are wrong about the laws of nature, then not only have we gotten the facts wrong, we are using incorrect *concepts*. Investigating the world is an attempt at once to rectify our claims and our concepts. Conceptual change is part and parcel of scientific change, because every new law we discover and every old one we are forced to give up brings

with it a change in our concepts. Semantics is not a discipline that can be pursued independently of our empirical inquiries into the rest of the world. These consequences may not sit well with our pre-theoretic talk about meaning, but so what? Our ordinary talk of hot and cold runs together differences in temperature and differences in specific heat (“Stone floors are colder than wood floors,”⁸) but we wouldn’t want to hold our theoretical physics hostage to our casual, unreflective practice. Why do so in semantics?

Another approach is Quine’s own: put *all* the inferences in the same basket, by rejecting the distinction between those good in virtue of meaning and those good in virtue of matters of fact. Such an approach can still accommodate a notion of sentences as contentful—as, we can say if we like, expressing contents. For sentences can still be understood as playing roles as premises and conclusions of inferences. But since for the most part the inferences in question are *multipremise* inferences, there is a sense in which no sentence plays its inferential role all on its own. For what a claim is evidence for or against and what is evidence for or against it depends on what collateral commitments are available to serve as auxiliary hypotheses in extracting inferential consequences. The inferential significance of endorsing any one sentence depends on what else one is committed to. This is the line of thought that leads Quine to think that “the unit of meaning is the whole theory,” rather than individual sentences. This holist consequence goes against the grain of the semantic atomism that has been at the center of the empiricist tradition—in both its traditional and its distinctive twentieth century logical and modal forms. But that incompatibility doesn’t by itself show that the holist response is incoherent.

I think that both these strategies are open to contemporary semantic inferentialists. Sellars’s approach seems to be wholly viable, though it has not, as far as I know, yet been pursued by other theorists. In *Making It Explicit*, I adopt a version of the Quinean strategy. This does in a certain sense involve giving up the notion of content—at least, it means giving it up in any sense that would be recognizable to a semantic atomist. But the suggestion is that we can do without talk of contents and meanings as *things* associated with sentences one by one in favor of talk of inferential connections among sentences that are contentful precisely in virtue of standing in those relations. The question of whether “all of them” is, like Sellars’s different one, a responsive and viable answer to the question: “Which inferences matter for

⁸ Joe Camp’s instructive example in *Confusion* (Cambridge, MA: Harvard University Press, 2002), p. 14.

content?" then becomes the question of whether one can make sense of semantic holism.⁹ So let us look at that issue.

V. Holism

The biggest challenge for holistic semantic theories has always been accounting for the possibility of *communication* or of interpersonal *understanding*. If the inferential significance of a claim depends on what else one is committed to, then any difference between the collateral commitments of speaker and audience can mean that a remark has a different significance in the one's mouth than it does in the other's ear. How is it then possible to make sense of the idea that they understand one another, so as to be able to agree or disagree? If the contents expressed by sentences must be individuated as finely as the theories they are embedded in, the intelligibility of communication across theories—the very notion of conveying information—is threatened. And the issue arises as urgently for diachronic communication as for synchronic, face-to-face cases. If, because of his very different collateral commitments, Rutherford meant something quite different by 'electron' than I do, it seems I can't disagree with him about whether electrons have fixed positions and orbits, since I can't either say or think anything with the content he would have expressed by saying "electrons orbit the nucleus."¹⁰ How, then, are we to understand so much as the possibility of cognitive progress in science?

Quine, of course, fully appreciated the force of these challenges. It is precisely such considerations of semantic incommensurability of

⁹ I take it that this is how Fodor and Lepore see things, though their discussion is not as clear on this point as one would like. They says "Like everybody else who thinks that content comes from inferential role, Brandom needs, but doesn't have, a story about which inferences are the ones that content comes from." ("Brandom's Burdens," p. 471) As I point out in the text, Sellars does have such a story. But everyone needs to make the distinction only if it is incoherent to appeal to all the inferences, and so to refuse to make the supposedly required distinction between "what the world contributes to the reliability of inferences and what the (putative) rules of language contribute." Acknowledging in a footnote that I am "explicitly suspicious" of the distinction, the authors say: "So be it; but then he needs some argument that he can make sense of the notion of content without employing it. To say that he has no such argument on offer would be to put the case very mildly." Working out a theory of that shape, as I do in Part Two of *Making It Explicit* (Cambridge, MA: Harvard University Press, 1994), ought to be the very best such argument. But in fact Fodor and Lepore seem to have in mind their worries about holism, which I address in the next section. Insofar as the holism is the source of the problem with not privileging some inferences, there is no *separate* issue about "which inferences." I address the holism question in the next section, where I sketch some of the arguments that are given in more detail in *Making It Explicit*.

¹⁰ As Israel Scheffler and Paul Feyerabend had pointed out already in the middle 1960s.

meanings that led him to insist that concepts such as communication, information, and cognitive progress—most importantly as they matter for the serious business of the conduct of science, but also in our more informal transactions with each other—must be understood as arising at a different semantic level: that of *reference*, rather than *meaning*. Our starting-point in this essay was the commonsensical observation that besides what is said or thought there is also what is talked or thought *about*. Though Rutherford and I may be saying different things—expressing different contents—when we use the phrase “the location of the electron,” we can both be talking about or referring to the same things: electrons. And we may be classifying them as falling within the same extension when we characterize them as “fast-moving.” Given his collateral beliefs concerning its divinity, the Zoroastrian sun-worshipper surely means something different by the term ‘the Sun’ than I do, but we may still both believe *of* the Sun and *of* visible things that the first is to be counted among the second. Or, if we don’t already share it, he can convey that information to me, in spite of our other semantic differences.

Quine’s way out, then, is to neutralize the otherwise corrosive effects of holism at the level of what is said or thought by appeal to the representational dimension of what is talked or thought about. Coreference or coextensionality is an equivalence relation that maps the disparate meanings expressed by sentences animated by distinct theories onto one another in just the way needed to underwrite the possibility of communication or information, reciprocal understanding, and so on. Again as noted at the outset of this essay, any semantic theory that begins with ‘that’ intentionality must eventually explain its relation to ‘of’ intentionality in any case—must proceed from an account of *senses expressed* to one of *objects* (and sets of objects) *denoted*. So it is open to any theory that adopts this order of semantic explanation to adopt Quine’s strategy of appealing to what is talked or thought about to secure an account of the nature of communication.

One might, it is true, at this point ask what work the holist theory of meaning (the expressive dimension of intentionality) would be doing, if it is immediately abandoned in favor of the representational dimension of intentionality when hard questions are asked about communication and shared understanding. But the answer on which the adoption of the order of explanation being considered is premised is clear: it is at that level that we are to understand the use of expressions (the functioning of states or performances) in virtue of which we can understand them as having representational properties and relations. The account in *Making It Explicit* is of this sort. The explanation of what is expressed by declarative sentences, and so by the ‘that’ clauses of

de dicto ascriptions of propositional attitudes, is offered in inferentialist terms. Then substitution-inferential commitments and their anaphoric inheritance are shown to be sufficient to explain what is made explicit in the *de re* portions of such ascriptions: the part where it is specified what one is speaking *of* or thinking *about*. Thus the implicit representational dimension that is expressed explicitly by the use of terms such as ‘of’ and ‘about’, which is invoked when we say that both Rutherford and I were thinking of electrons, and that the Zoroastrian was talking about the Sun, is explained in ultimately inferential terms. Inferentialists do not, and should not want to, deny the existence or the importance of the representational dimension of intentionality. Rather, they are committed to an order of explanation that seeks to *understand* ‘of’ intentionality in terms of ‘that’ intentionality. Coreference—a kind of intersubstitutability of expressions—then provides a respect of similarity across inferential roles, grouping them into extensional equivalence classes, which interlocutors *can* share.

In his writings on holism, Fodor has argued that where meanings-as-inferential-roles are individuated so finely that there is little hope of different interlocutors (or even one interlocutor at different times) having *identical* ones, it is of no use for the holist semantic theorist to attempt to retreat to the idea of mutual understanding consisting in the grasp of at least *similar* meanings, or meanings that in some sense “overlap.” One might initially think that, for example, Rutherford and I at least share *some* of the inferential consequences of application of our uses of ‘electron’. We both agree, for instance, that it follows from something’s being an electron that it is negatively charged, has a mass that is orders of magnitude smaller than that of the proton, that its movement creates a magnetic field in the direction specified by the right-hand rule, and so on. But that overlooks the fact that ‘charge’, ‘proton’, and ‘magnetic field’ all by hypothesis *also* mean something different in his mouth than in mine. Once we realize that these are all in exactly the same boat as ‘electron’, we see that we’ve just put the issue off, rather than solving it by defining a sense of ‘similarity’ that consists in having *some*, but not all, inferential antecedents and consequences in common. We actually have *nothing* in common.

But this conclusion is too strong. For we do share the *words*, at least in the sense of noise- or sign-design types. When Rutherford sees lightning, he, like me, is committed to the correctness of applying ‘electron’; and when he does apply it, he, like me, is committed to the correctness of the application of ‘charge’, ‘magnetic field’ and so on. One might respond at this point on Fodor’s behalf that although *some* kind of similarity metric is induced by counting the *noises* that express the conclusions two interlocutors would draw from, or the promissory claims

they would count as evidence for, claims expressed using ‘electron’, still that is only because we have restricted ourselves exclusively to *nonsemantic* properties of their utterances. So nothing like shared *meaning* is thereby underwritten. But once again, this is too hasty a conclusion. Here one might think of Davidson’s account of communication as interpretation. Davidsonian interpretation is explicitly understood as consisting in mapping the *noises* made by the interpretive target onto the noises made by the interpreter. In understanding another, I am to use sentences in *my* mouth to attach as labels to his sentences, and thereby serving to ‘measure’ them. My interpretation is a useful one—I have understood what the other says—just insofar as the inferential moves I am committed to endorse with the noises I produce mirror (perhaps with a qualifying commentary couched in my own idiom) his. The ideal interpretation is a homomorphism, a structure-preserving mapping, from his noises onto mine, preserving the consequential structures. Where not all the moves between his noises that he is committed to the correctness of are matched by similar commitments on my part regarding my counterpart (according to the interpretation mapping) noises, only similarity, and not identity of consequential role—and so only partial understanding—is achieved. And recall that inferential role in the broad sense includes the inferential connections between circumstances of appropriate application and appropriate consequences of application quite generally—even when the circumstances or consequences of application themselves are *noninferential*. Thus Rutherford and I are both disposed to respond to a bolt of lightning by applying the term ‘electron’, and to respond to applying the expression ‘high voltage, high amperage electron flow’ to a bare piece of metal by avoiding contact with it. These language entry and language exit moves, no less than the language-language ones, also give us something important in common, even when described at a so-far-subsemantic level, that is, in a nonsemantic vocabulary. I do not see why the structures so-described do not underwrite a perfectly intelligible notion of *partially* shared, or merely *similar* inferential roles.¹¹

A more radical response than Quine’s, one that obviates the necessity of considering similarities rather than identities of meaning as what is shared, is one that rejects the model of sharing on which the worry about the intelligibility of communication in the face of the holistic character of meaning is based. The motivating picture is at root a Lockean one: the speaker has an idea in his head, and his

¹¹ I do think there are some deep issues about holism in the vicinity. I discuss some of them in Chapter 6 of *Tales of the Mighty Dead: Historical Essays in the Metaphysics of Intentionality* (Cambridge, MA: Harvard University Press, 2002).

uttering the words he does succeed in communicating that idea to me if the idea hearing them uttered arouses in me is the same (repeatable) as the one he has. The meaning or content is, as it were, to be transported from his head to mind, or reproduced in mine. But one could think of understanding rather on the model of a cooperative practice or activity. In particular, I can be said to understand your remark insofar as I can compute its inferential significance both for you and for me, and navigate successfully back and forth across the two perspectives on its content constituted by the background of auxiliary hypotheses drawn from *your* collateral commitments and the ones drawn from *mine*.

To do that, I need to be able to determine what you would take to be the consequences of your claim, and what would be evidence for it and against it, given *your* theory of things. This will matter if I want to predict what else you'll go on to say or believe, or what you will try to do in a particular situation. (If you believe that animal over there is a deer and desire to shoot a deer, then you may try to shoot the animal.) I also need to be able to determine how *I* should draw inferences, using *your* utterance as a premise. That includes mapping your noises onto mine, and then extracting the inferential consequences from the claim that would be expressed by *my* corresponding sentence utterance, with the help of auxiliary hypotheses drawn from other claims *I* endorse. This will matter if I want to extract *information* from your remark, that is, to draw conclusions from it about what is *true*, and if I want to predict what you are likely to *succeed* in doing. (If you *do* try to shoot the animal, I know, though you do not, that what you are likely to succeed in shooting is in fact a horse.)

The capacity to understand each other is the practical ability to navigate across the gulf between doxastic perspectives created by the effect of differing collateral commitments on the inferential significance of one noise in the mouth that utters it and the ear that hears it. When that implicit skill is made explicit in the form of claims (thinkables, believable, sayables), it takes the form of *de dicto* and *de re* ascriptions of propositional attitudes. "The speaker believes *that* that animal is a deer," "The speaker believes *of* or *about* a horse that it is a deer." The expressive role of explicitly representational terms such as 'of' or 'about' in this usage is to mark that the identity claim (intersubstitutional commitment) connecting 'that animal' and 'a horse' is one that the ascriber *undertakes*, and is using to express the content of the propositional commitment *attributed* to the speaker. When what is at issue is what *I*, the ascriber, should conclude from your remarks, I am going to map your 'deer' talk onto my 'horse' talk. I *say* that (rather than just *doing* it), precisely by using *de re* forms of ascription: talking about

what you are talking *about*, what you are speaking *of*, what you *represent* as what. Those *de re* ascriptions are the home language games of such explicitly representational locutions—the ones that make explicit the representational dimension of intentionality, the ‘of’ flavor, rather than the *de dicto* ‘that’ flavor. That is, it is their use in such contexts that *makes* them explicitly representational locutions. And that expressive role can be understood in terms of the *inferentially* articulated and specified distinction of perspective between premises attributed and those endorsed. In that way adopting the navigation-across-perspectives model of communication can also be seen as a way of developing the Quinean retreat-to-reference response to the challenge to make communication intelligible within a holistic semantics of what sentences express.

Practitioners who understand each other in this practical sense—who can successfully make use of each other’s remarks in their own reasoning, both about what the other has reason to do (given his or her beliefs) and about what one has oneself reason to do (given one’s own beliefs)—do indeed ‘share’ something. But what they share is like the dance that Fred and Ginger are doing together—one and the same dance, even though individually they are doing different things (him going forward, her backward; she dipping, he holding; she twirling, he leaping...). It is not like the cadence that the soldiers marching in step share: something visible already in what *each* is doing individually, simply repeated across them all. We can think of conceptual (that is, inferentially articulated) content as like Fred and Ginger’s dance: something that is *essentially perspectival*, in that grasping it (like engaging in the dance) requires doing *different* things from each individual participating (in the conceptual case, depending on the background constituted by their other commitments). This is a different model of understanding and communication from the Lockean repetition or reproduction model—the soldiers marching model. An account of this shape will count as leaving communication unintelligible only to those who insist upon the Lockean model as the *only* way to understand understanding each other.

Besides the Quinean retreat-to-reference response and the practical-navigation-across-perspectives response, there is a third way in which inferentialism can (and the theory expounded in *Making It Explicit* does) respond to worries about the effects on the intelligibility of communication of the relativity of inferential significance to collateral commitments. For another great division among theories of concepts and their contents is that between broadly Cartesian and broadly Kantian approaches. Cartesians think of concepts as something like mental particulars. The principal question about them concerns the

thinker's grasp of them: how well do we really understand them. Kantians think of concepts rather as norms or rules that bind those who apply them, determining what would count as a *correct* judgement in which they are applied, or a *successful* intention in which they were applied. For them the principal question does not concern our grip on the concepts, but rather their grip on us. So for Kant, issues of the *bindingness* of concepts, the way in which they become *valid* (gültig) for a thinker and agent, is the central philosophical issue. Transposed into a linguistic key, the question becomes what I must do in order to count as having applied, say, the concept copper, in thought or assertion—to have subjected myself to assessments of the truth of my claim or the success of my action accordingly as what I am talking about or acting on is or is not *copper* (rather than, say, just some reddish metal).

A good model of the second kind is playing a counter in a game. Once I count as a player in the game, I can play a counter that *has* a certain significance—obliges me to make some further moves under various circumstances, precludes me from making others, entitles but does not oblige me to make others, and so on. And the facts about the normative significance of that move may significantly outrun what I *understand* that significance to be. I may not realize all, or even very many, of the aspects of the normative significance of my performance, for it nonetheless to have that significance. I do not, for instance, need to *know* that the melting point of copper is 1083.4°C in order to call something copper, and thereby to have committed myself to its not melting at 1083°, but melting at 1084°C—in the sense of having said something that is true if and only if that condition obtains. Thus my remark or thought is subject to assessment according to that norm, even though I may not be aware of that fact.

On a Kantian picture, then, you and I can share a concept even though our *dispositions* to call something 'copper' are quite different—perhaps because of our different collateral commitments. In spite of such differences, we may be understood as binding ourselves by the very same complex norm for assessments of truth and success by our use of the word 'copper'. For what matters for such assessments is what auxiliary hypotheses (e.g., about the melting point of copper) are *true*. The fact that you and I have different views about which these are, and so are disposed to draw different conclusions from something's being copper, does not alter what *really* follows from it. *De dicto* specifications of the content of another's thought depend on the inferences she is disposed to draw from it: what she *thinks* she is committing and entitling herself to thereby. They articulate her *conception*. *De re* specifications of the content of another's thought, by contrast, depend on

what inferences to and from it are *in fact* (including counterfactually) correct. They articulate the *concept* she has applied, even insofar as she is ignorant of or mistaken about its content. The practical navigational capacities that are made explicit in *de re* specifications of the contents of ascribed propositional commitments express the standing commitment each of us has to their being *one* set of inferential roles that bind *all* interlocutors: those, namely, determined by multipremise inferences in which the collateral commitments supplying auxiliary hypotheses are *true*.

I conclude that the sensitivity and relativity of the inferential significance of a sentence to collateral commitments poses a threat to the intelligibility of communication *only* for a theorist whose own collateral commitments at the metalevel include *all* of:

- a) commitment to communication's having to take place at the level of meaning, rather than of reference;
- b) commitment to a Lockean reproductive model of the sort of sharing that communication consists in; and
- c) commitment to a Cartesian, rather than a Kantian model of our relation to concepts.

The inferentialism of *Making It Explicit* explicitly rejects all of these assumptions. (If you think the couch cannot possibly go on the wall next to the fireplace, that may be because you are not thinking about moving enough of the rest of the furniture. Almost certainly you are not thinking about knocking out one of the interior walls.)

VI. Compositionality

A related worry about inferentialism concerns the *productivity* of language and thought: the fact that competent speakers and thinkers are able to produce and understand an indefinite number of sentences that express *novel* contents—not just novel in the sense that that speaker, hearer, or thinker has never entertained them before, but in the much stronger sense that *no-one* has ever done so. This striking observation was first made by Chomsky almost fifty years ago, and it has been adequately confirmed since in many ways—ranging from statistical analyses of empirical corpora of actual utterances to theoretical analyses based on the sentences of given lengths generated by particular partial grammars given fixed lexicons. About the only idea anyone has as to

how to explain this striking fact is to treat language as *compositional*, at least in the weak sense that semantic interpretations of unfamiliar sentences are understood as generated by operations on semantic interpretations of their familiar parts.

The compositionality challenge to inferentialism arises because it is essentially a categorially top-down order of semantic explanation. It begins, not with the contents of subsentential components, but with what is expressed by whole declarative sentences. The thought is that (as a recent paper puts it): “productivity demands compositionality, and compositionality implies the priority of subsentential semantics to sentential semantics.”¹² The first of these claims ought to be granted (at least for a suitably broad understanding of ‘compositionality’). But the second is surely too strong. Compositionality does not imply semantic atomism, but at most what Dummett calls ‘molecularism’. A molecularist order of semantic explanation starts with sentences, and so contrasts both with fully holist theories, which start with whole idioms or theories, from above, on the one hand, and atomist theories, which start with subsentential expressions such as singular terms and predicates, from below. That is, it takes seriously the idea that the sentence is the minimal unit for which one can take *responsibility*, or which can express the undertaking of a *commitment* (Kant), the minimal unit to which pragmatic *force* can attach (Frege), and the minimal unit with which one can make a move in a language game (Wittgenstein). The notion of semantic content is accordingly introduced to begin with as a way of codifying or explaining proprieties of the use of expressions of this category: sentences. (One might or might not look to *inferential* properties of the sentences at this point in the story, as inferentialists do.) One might do this for a relatively small finite corpus of sentences—say, those which a given interlocutor has used either as speaker or hearer, during his initiation into the language. One then looks to codify or explain features of the use of those sentences by noting similarities among them, paradigmatically that they contain occurrences of the same subsentential expression. Looking at the contribution the occurrence of that subsentential expression makes to the features of the use of the whole sentences in virtue of which they are intelligible as semantically contentful then allows one to attribute semantic content, in a derivative, indirect sense, also to those subsentential expressions. That in turn can allow one to generate semantic interpretants for a much larger class of sentences, compounded in familiar ways out of the familiar parts. Compare: offering a structural analysis of the behavior of

¹² “Brandom’s Burdens,” *op. cit.*, p. 480.

ordinary solids, liquids, and gases in terms of their molecular composition, and only later seeking to explain the behavior of those molecules, and many others one has *not* observed, on the basis of their atomic structure.

Like Dummett, I think that Frege pursued such an explanatory strategy of *decomposition* and *recomposition*. He starts with sentences, as the bearers of truth values, and assigns Bedeutung to singular terms only so as to be able to analyze the sentences as applications of functions to arguments. His strategy for specifying the truth conditions (and inferential roles) of arbitrarily iterated quantifications depends on one being able to form complex predicates by substitutional variation of whatever sentences are already on board, and then forming new sentences from those predicates by clamping quantifiers onto them. This way of proceeding does not even depend on the original sentences having literal lexical-syntactic parts; it is enough if they can be sorted as similar to one another in respects that act enough like the equivalence classes generated by orthodox substitutional variants.¹³ There may be various reasons not to want to adopt the method of decomposition and recomposition. But it is surely a coherent strategy for achieving compositionality. So achieving the effect of compositionality does not *imply* the explanatory priority of sentences to subsentential structures.

The sort of inferentialism pursued in *Making It Explicit* explains our capacity to understand novel sentences in two different ways. First, it includes an account of the introduction of *logical* vocabulary in a way that is straightforwardly and traditionally compositional. Logical vocabulary is demarcated by its playing a distinctive expressive role: making *explicit* in propositional (claimable, believable, thinkable) form broadly *inferential* commitments that otherwise remain implicit in practical assessments of practice. The paradigm of logical vocabulary in this sense is the conditional, which lets one *assert* an inferential connection between (what is expressed by) the sentences that appear as its antecedent and consequent. In the same way, negation can be understood as codifying commitment to the *incompatibility* of claims. But on this account, *normative* vocabulary also qualifies as logical, in that it serves to codify commitment to the correctness of various patterns of *practical* reasoning. And other bits of vocabulary, from identity and quantificational locutions to the ‘of’ and ‘about’ that express representational intentional directedness

¹³ See my “Singular Terms and Sentential Sign Designs,” *Philosophical Topics*, XV, No. 1, Spring 1987, pp. 125-167, and Bas Van Fraassen’s “Quantification as an Act of Mind,” *Journal of Philosophical Logic*, 11, No. 2, August 1982, pp. 343-369.

(like normative locutions in that they are not usually thought of as distinctively *logical* vocabulary) show up as playing expressive roles of these kinds. For each, it is shown how they can be introduced into languages that do not already contain them, in such a way that the inferential roles of sentences in which they appear are settled by the inferential roles of non-logical sentences formulated in the prior vocabulary. It is in these terms, for instance, that the inferential roles of arbitrarily iterated mixtures of *de dicto* and *de re* ascriptions (sentences such as “Tom believes of Benjamin Franklin that Henry Adams believed that he was not the inventor of the lightning rod,”) are computed. For the special class of logically compound sentences (including those formed by using *subsential* logical locutions), this account is compositional in the classical bottom-up sense, albeit conducted within an inferentialist framework.

The inferences (and so inferential roles) generated for compounds formed by applying vocabulary that is *logical* in this special, semantically expressive sense are not restricted to those that are good in virtue of logical *form* (even for classical logical particles like conditionals and negation), because of the special expressive role discerned between the use of *nonlogical* sentences and the logical compounds that make explicit aspects of that use. They are, however, all inferences that are good in virtue of the (broadly inferential) *meaning* or *content* expressed by the use of the logical vocabulary in question. But not all vocabulary is logical vocabulary even in this extended semantically expressive sense, and (so) not all material inferences manifest the contents of logical concepts. What about the rest?

Here the account in *Making It Explicit* adopts a version of the decomposition-recomposition strategy. The material inferential behavior of nonlogical sentences is dissected with a substitutional scalpel by using Frege’s basic methodological idea that what it is for two subsential expressions to play the same indirectly inferential role (‘indirect’ since they don’t express reasons in the sense that they cannot play the directly inferential roles of premise and conclusion of inferences) is for substitution of one for the other to preserve (important features of) the inferential roles of the sentences they occur in. Equivalence classes of subsential expressions generated by these relations can then be associated with what are called (in Chapter Six of *Making It Explicit* and Chapter Four of *Articulating Reasons*) “simple material substitution inferential commitments.” Those subsential, indirectly inferential contents then determine the correctness of all the substitution inferences involving compound

expressions in which they occur.¹⁴ This is how compositionality works for material inferences: via substitutions. It can be shown that the formal power of this substitutional decomposition and recomposition is equivalent to that of the standard functional-categorical grammars that David Lewis considers in “General Semantics.”¹⁵ That is, the result is as compositional as the most powerful approaches that assign semantic interpretants to, say, singular terms and sentences (let us say, objects and sets of possible worlds), and then generate semantic interpretants for derived categories functionally: predicates as functions from (tuples of) objects to sets of possible worlds, adverbs (such as ‘slowly’) as functions from functions from objects to sets of possible worlds to functions from objects to sets of possible worlds, and so on.

¹⁴ Of course, not everyone agrees that this construction is successful. See Alex Oliver’s *TLS* review of *Articulating Reasons*, and Fodor and Lepore *op. cit.* They are quite right to point out that the characterizations of singular term and predicate will work for natural languages only if they are first somewhat regimented or syntactically pre-processed. It is not hard to show how to do that so as to deal with the cases they forward as counterexamples. (So, for instance, Oliver’s case of the non-substitutability of ‘the first Postmaster General’ for ‘Ben Franklin’ in the phrase ‘good old Ben Franklin’ requires that appositives be made explicit, in a form such as ‘Ben Franklin, who is good and old’, in which case the substitution is syntactically allowed and semantically correct. Cf. *Making It Explicit*, p. 388, where this construction is discussed.) Transformations need to be applied before categorial analysis, and after categorial synthesis. This is a point, again, that is explicitly acknowledged in *Making It Explicit*, even though it is not addressing that phase of the process.

Another sort of criticism both sets of critics make, however, is based on a simple misunderstanding of the view they are criticizing. Thus, to pick a representative example from Fodor and Lepore (p. 476), they object to the denial that there can be systematically asymmetric substitution relations among singular terms, as there can be among predicates, that “Father was at Magdalen,” entails “Father was at Oxford,” but not vice versa. But this looks like a counterexample only if one drops the crucial initial quantifier from the claim. For there to be an asymmetric relation of the kind asserted, it must be the case that the inference from P(Magdalen) to P(Oxford), but not the converse, holds for *every* predicate P, not just for some specific one. And this is obviously not the case: consider P = Peter has never set foot in.... (Of course one’s next thought would be that, so understood, the corresponding asymmetry does not hold for the predicates ...walks and ...moves. But that is not right. That thought overlooks the fact that the predicates in question are not *simple* predicates, but *complex* ones—in Dummett’s vocabulary. That is, they are sentence frames identifiable with equivalence classes of sentences that are substitutional variants of one another.) All this is explicitly discussed, along with other putative counterexamples, in the extended section VI of Chapter Six of *Making It Explicit*, “Objections and Replies,” to which the reader of *Articulating Reasons* was referred for the details of the argument.

¹⁵ And is, in the Appendix I to Chapter 6 of *Making It Explicit*. The Lewis article is in G. Harman and D. Davidson, *Semantics for Natural Languages* (Dordrecht: Reidel, 1972).

That is not the end of the story, however. For once again, not *all* the material inferences sentences are involved in can be computed on the basis of substitutions or the functions they generate. There will be a significant residue of multipremise material inferences whose correctness is *not* settled in this way by the contribution made to the inference potential of the various individual components that occur in the premises and conclusion, even when they are all summed. New things happen. That the apple is red does not follow either from its being ripe or from its being a Macintosh, but does follow from both premises together. The inference that would be explicitly licensed by the conditional “If an apple is *both* ripe and a Macintosh, then it is red,” is not a substitution inference. If that claim is true, then it does articulate part of the commitment that is implicit in applying the terms ‘Macintosh’, ‘ripe’, and ‘red’ and in that sense is a feature of the contents they express. But it is not a consequence of features of their use that are manifested one by one. In this sense (though, given the rejection of a distinction between inferences made good by meaning and inferences made good by how the world is in favor of seeing them all as having both sources—for very different reasons), I agree with Fodor and Lepore that “in general, the inferential role of a sentence/thought is not determined by the inferential roles of its constituents.”¹⁶ But I think this fact evidences not a particular defect of inferentialism, but simply a fact about languages and (so) concepts. It is important not to treat languages as more compositional than they are. They are compositional with respect to their substitution inferences, but not with respect to the rest.

Some material proprieties of multipremise inferences cannot be discovered simply by inspecting the use of their component expressions. The substitution inferences give us a good handle on the proprieties of use governing novel compound sentences. They ensure that we always have a place to start in sorting the inferences involving the novel sentences into good ones and bad ones. But since what inferences are good—and so, on this line, what our words mean and what content our thoughts have—depends on how the world actually is (for instance, on what color ripe Macintosh apples are), we may have to go out into the world to find out what follows from or is evidence for or against novel claims. The idea that this sort of failure of compositionality is a *problem* for semantic theories is a product of commitment to semantic theories having to be both categorially bottom-up, and a Cartesian transparency thesis about the epistemological availability of the

¹⁶ Fodor and Lepore, *op. cit.*, p. 472.

contents we grasp in the sense of being able to deploy them in thought. But these commitments are optional, and are rejected by inferentialists.

VII. Conclusion

In the first three sections of this paper, I sought to place inferentialism in a more general botanization of semantic theories, and to distinguish various more specific shapes versions of that approach can take. In the second half of the paper, I considered inferentialist responses to worries about its capacity to distinguish the inferences constitutive of conceptual content, about the intelligibility of communication and reciprocal understanding given its attendant holism, and about its ability to deal with the phenomena of productivity and (so) compositionality. Those interested in more detailed workings-out of those responses will find some in *Making It Explicit*. But I don't take it that the arguments on offer there settle the ultimate viability of the specific version of inferentialism they articulate. I've told the story in the way I have here in part as a reminder of how misleading it can be to assess the claims of a systematic theory piecemeal. Very often one move makes sense only in the context of others—as for instance the sort of holism inferentialism involves has many more resources available for explaining what it is for two interlocutors to be applying the *same* concept when concepts are understood in the Kantian way, in terms of their normative grip on us, rather than in the Cartesian way, in terms of our epistemic grip on them. One thread that runs through much of the tapestry I've been displaying is the way in which an inferentialist semantic metalanguage gives us the resources to explain what is expressed by the locutions that make explicit specifically *representational* relations. How one understands the relation between the expressive and the representational dimensions of intentionality—between 'that' intentionality and 'of' intentionality—must be at the core of any theory of content or intentionality.