Against Disquotation

Richard Kimberly Heck

Department of Philosophy, Brown University

Abstract

Disquotationalism is the view that the only notion of truth we really need is one that can be wholly explained in terms of such trivialities as: "Snow is white" is true iff snow is white. The 'Classical Disquotational Strategy' attempts to establish this view case by case, by showing that each extant appeal to truth, in philosophical or scientific explanations, can be unmasked as an appeal only to disquotational truth. I argue here that the Classical Strategy fails in at least two cases: attributions of truth to context-dependent utterances and uses of falsity in psychological explanations of behavioral failure.

Deflationary theories of truth come in two basic forms. Views of the first sort concern the truth of propositions. The rough idea is that there is nothing more to its being true that snow is white than snow itself's being white. An especially strong form of the view would be that the sentences

(1)   Snow is white.

and

(2)   It is true that snow is white.

express the very same proposition, and similarly for other such pairs.[1] A weaker form of the view would be that (1) and (2) are analytically or conceptually equivalent. An intermediate view would be, to borrow a term from Hartry Field, that (1) and (2) are 'fully cognitively equivalent', meaning, to first approximation, that believing (desiring, etc) the one

---

[1] We'll see in section 2 there is potentially a confusion implicit in this way of putting the point, but it will do for now.

has the same significance for an agent's behavior (including their merely mental behavior) as does believing the other.

The other sort of view, which is known as *disquotationalism*, is a view about the truth of sentences. This time, the rough idea is that nothing more is required for the sentence "Snow is white" to be true than snow's being white. The strongest possible form of this view would be that (1) and

(3)   "Snow is white" is true.

have the very same meaning. A weaker form of the view would be that (1) and (3) are analytically or conceptually equivalent. And an intermediate form, which is what Field (1994, pp. 250–1) initially endorses, is that (1) and (3) are 'fully cognitively equivalent'.

Both the view that I am calling 'disquotationalism' and the name by which I am calling it have their roots in the work of W. V. O. Quine (1970b, p. 12), who famously wrote: "By calling the sentence ['Snow is white'] true, we call snow white. The truth predicate is a device of disquotation." Elsewhere, however, Quine (1956, p. 187) explicitly denies that (1) is "analytically equivalent to" (3), or even necessarily equivalent to it, insisting that "agreement in truth value can be claimed, and no more": His view, that is to say, is just that (1) and (3) are *materially* equivalent, and almost no-one would deny that claim. Of course, Quine is famously skeptical about analyticity, but the reason for his caution is not to be found there but instead in an objection that Alonzo Church (1950) had brought against an analysis of belief-sentences due to Rudolf Carnap (1947, pp. 61–2).[2] If (1) and (3) are synonymous, then their translations into German

(4)   Der Schnee ist weiss.

(5)   "Snow is white" ist wahr.

should also be synonymous, which they plainly are not. But there is now a large literature on Church's objection. Perhaps some of the moves made there could be adapted to this case.

---

[2] Carnap's proposal was that "John believes that snow is white", for example, would relate John to the *sentence* "snow is white", rather than to a proposition or something of that sort. Church credits the idea of using translation as a test for synonymy to C. H. Langford (1937).

A similar, but to my mind more worrying, objection was voiced many years before by G. E. Moore (1953, p. 276).[3] In the case of propositional deflationism, the objection is that the truth of (2) seems to require the existence of the proposition that snow is white, whereas the truth of (1) does not. The obvious response is that the objection tendentiously assumes that the complement clause "that snow is white" refers to the proposition that snow is white. But all the objection really requires is that there is *something* to which the complement clause refers to which no constituent of (1) refers.[4] One could perhaps deny even that claim, but a similar response is plainly unavailable to the disquotationalist, who faces a similar objection: The truth of (3) requires that there be such a sentence as "Snow is white", whereas the truth of (1) does not. It would be utterly implausible to deny that the quotation-name that occurs in (3) is a name of a sentence, or that the truth of (3) requires the existence of that sentence.[5]

Marian David (2005, §IV), from whom I learned of Moore's case for precedence, discusses several responses that a deflationist might explore, arguing that none of them work. Any of these could be adapted by a disquotationalist, but none of them seem to me to be any more helpful in that form. There is, however, a strategy that David does not consider, namely, the one that is actually adopted by Field (1994, pp. 250–1), who simply concedes Moore's objection and weakens the view to one on which a belief expressible by (3) is 'fully cognitively equivalent' to one expressible by (1) only *modulo* the belief that the sentence "Snow is white" exists. Presumably something similar would need to be said about such pairs as

(6)   "Snow is white" is false.

(7)   Snow is not white.

and also about embeddings of such sentences in conditionals. Whether all that can be made to work is not clear, and David is no doubt correct that "it is surprising that the *prima facie* difficulty [that Moore's objection]

<hr>

[3] Moore's lectures were originally given in 1910–11 though not published until much later.

[4] Thus, insisting upon a 'paratactic' treatment of complement clauses (Davidson, 1968) would not answer Moore's objection.

[5] Pro-sentential theories (e.g. Grover et al., 1975) may deny exactly that. But consider such generalizations as: Every true Σ_1 sentence is provable in Robinson arithmetic. The quantifier must range over sentences, and there is no indication that "is true" and "is Σ_1" apply to such sentences in fundamentally different ways.

poses... is rarely mentioned" (2005, p. 387). But it is hard to see this objection deciding the issue, either.

What will most matter for our purposes is that disquotationalism is a *much* more radical view than deflationism is. To get a sense for this, note that a disquotationalist would regard the following statement as true:

(8)   "Snow is white" would have been true even if "snow is white" had meant that pigs fly.

This may seem surprising,[6] but it is simply a consequence of the extremely strong relationship between (1) and (3) that the disquotationalist requires there to be. As disquotationalists understand (8), it is equivalent to:

(9)   Snow would have been white even if "snow is white" had meant that pigs fly.

And that is because, as Field (1994, p. 266) remarks, echoing Quine, "... to call 'Snow is white' disquotationally true is simply to call snow white...", and that is true whether or not "is true" occurs inside a modal context. If so, then, at least arguably, such claims as (8) are ones disquotationalists must accept. Suffice it for now to say that Field (1994, p. 275) not only accepts such claims but emphasizes them.[7]

Propositional deflationists, by contrast, have no such commitment. They may happily regard (8) as false. Sentential truth can be explained in terms of propositional truth thus:

(10) S is true iff $\exists p[S$ expresses $p \wedge p$ is true]

That is: A sentence is true iff it expresses a proposition that is true. The crucial question is then how one understands the expression relation. One option is to understand it too in a deflationary way. In that case, one may well find oneself committed to (8). And some propositional deflationists, such as Paul Horwich (1998), do hold such a combination of

---

[6] Some people think sentences have their meanings essentially (see e.g. Simchen, 2012). I find such views difficult to comprehend and incompatible with any plausible theory of human language comprehension. But the central issue here could be discussed entirely in terms of demonstratives—see section 3.2—and it does not seem at all plausible that an uttered demonstrative has its reference essentially.

[7] That said, Field (1994, pp. 275–8) also explores the possibility of using a 'extended' disquotational truth-predicate to avoid this consequence. We'll discuss the extended notion a great deal below.

views. But others, such as Scott Soames (1999), clearly do not, rejecting deflationism about expression and seeking some substantial account of the relation between a sentence and the proposition it expresses. Sentential truth then becomes an equally substantial notion.

Assuming propositional deflationism, then, disquotationalism is all but equivalent to deflationism about the relation of expression, that is, to a view on which the notion of expression can be completely characterized in terms of such apparent trivialities as:

(11) "Snow is white" expresses the proposition that snow is white.

If (11) is indeed a triviality, neither requiring nor admitting of substantial explanation, then so is the notion of sentential truth; if not, then not. But the underlying point is independent of any commitment to propositional deflationism. In effect, what the disquotationalist's acceptance of (8) is telling us is that they regard

(12) The condition that must be met for "Snow is white" to be true is that snow should be white.

as itself a sort of analytic truth, and so one that is true in all possible worlds, no matter what "Snow is white" itself means in those worlds. Ultimately, then, disquotationalism is not just a view about truth but a view about content or, more precisely, about truth-*conditions*. Indeed, Field (1994, p. 253) describes the central question in which he is interested as "whether truth conditions. . . play a central role in meaning and content".

For Field, then, disquotationalism is not motivated by skepticism about propositions, as it is, at least to some extent, for Quine. Field's skepticism extends much further, to encompass *any* notion of representational content robust enough to capture truth-conditions. Quine (1960, pp. 23–4) sometimes seems to doubt that it is even possible to make proper sense of any notion of equivalence between sentences stronger than material equivalence. But that worry emerges from Quine's concerns about the limits of 'radical translation' which, infamously, seem mostly to be imposed by his behaviorism (Chomsky, 1969). Field, on the other hand, is willing to countenance 'theories of content' that Quine would not, namely, any theory that reduces content to physicalistically acceptable notions. In fact, Field (1972; 1978) was for many years a champion of causal theories of reference. It seems to have been his eventual disillusionment with the project of 'reducing' truth-conditional

content to physicalistically acceptable notions that led Field to embrace disquotationalism.

What is at issue in the debate over disquotationalism, then, is no less than this:

> If the notions of meaning and content are to do the work we need them to do, must they be characterized in terms of some representational notion sufficient to determine truth-conditions, in some substantial sense?

I take Field to have made a significant contribution just by realizing that, since disquotationalism precludes any substantial notion of truth-conditions, we can address this seemingly intractable question by asking instead whether we can get by, in the theory of meaning and content, with a disquotational truth-predicate. That seems much more manageable.

What is meant here by a 'disquotational truth-predicate' is one that is stipulatively introduced so as to behave the way disquotationalists think our actual truth-predicate does behave.[8] A disquotational truth-predicate is thus supposed to be one that is wholly characterized in terms of the cognitive equivalence of a given sentence with one attributing truth to that sentence. So the first thesis of disquotationalism is:

DT1 An intelligible notion of truth can be adequately explained in terms of the Disquotation Principle: $\ulcorner$"S" is true$\urcorner$ is fully cognitively equivalent to S (modulo the existence of S itself).

The second thesis concerns the role that such a notion of truth is, and is not, suited to play:

DT2 A disquotational truth-predicate can play only an 'expressive' role and can never play a semantic, 'word–world relating' role.

The standard example of such an 'expressive role' is the use of truth in generalizations, such as:

(13) Everything the Pope says is true.

Disquotationalists want to insist that (13) should be read as the infinite conjunction of such sentences as:

> If the Pope says "Dogs bark", then dogs bark.
> If the Pope says "Pigs fly", then pigs fly.

---

[8] One might reasonably compare the strategy I'm about to describe with a famous 'test' that Saul Kripke (1977, p. 265) uses in his discussion of referential uses of descriptions.

The difficulty is that it is not clear how to express this generalization finitely. The obvious thing to try is something like:

(14) For all S, if the Pope says S, then S.

But that is ill-formed, since S here seems to be unable to make up its mind whether it goes proxy for a sentence (as in the third occurrence) or a name of one (as in the second). The solution is supposed to be provided by the truth-predicate, which allows us to reconstrue (13) as:

(15) For all S, if the Pope says S, then S is true.

This, then, is the 'generalizing' role that disquotationalists take the truth-predicate most fundamentally to play (cf. Quine, 1970b, pp. 11–3).[9]

  An opponent of disquotationalism can happily accept DT1 and DT2, however. Some even do, on the ground that we need a disquotational truth-predicate to express generalizations like (13) and, more importantly, such modal generalizations as "All of the axioms of Euclidean geometry might have been true" (see e.g. McGee, 2005, p. 147). My own view is that this is a mistake, one that results from a failure to distinguish between sentential and propositional notions of truth (see Heck, 2004, §2) . We'll discuss this matter further below.

  At the moment, we need only the weaker point that accepting the legitimacy and utility of a disquotational truth-predicate does not suffice to make one a disquotationalist. The distinctive disquotationalist thesis is rather:

DT3 There are no *legitimate* uses of the notion of truth that are not 'expressive', i.e., that cannot be understood as uses of a disquotational truth-predicate.

What makes one a disquotationalist, then, is the view that the *only* legitimate notion of truth is the disquotational one.

  The dialectic then tends to play out as follows. Opponents of disquotationalism note that there are philosophical and scientific theories and explanations in which the notion of truth seems to play an important role. For example, both in logic and in semantics, we find such claims as

(16) A conjunction is true iff both its conjuncts are true.

_____

[9] It is, in fact, much less clear than usually seems to be supposed what this 'generalizing' is supposed to involve. See Heck (2019, esp. §1).

which is supposed to express, among other things, the truth-functionality of conjunction. What I shall call the *Classical Disquotational Strategy* attempts to unmask this application of the notion of truth as really being of the 'expressive' sort for which the disquotational truth-predicate was designed. If so, then the disquotational notion of truth is adequate for the theoretical or explanatory purpose at issue and no 'substantial' notion of truth is required.[10]

Of course, there are other sorts of strategies open to a disquotationalist. They might argue that the truth-involving theory or explanation is, for one reason or another, independently objectionable. Or they might offer an alternative. In the case mentioned, for example, they might argue that logic and semantics should proceed in terms of proof- or verification-conditions, in which case (16) can be rejected in favor of something like:

(17) A conjunction has been verified iff both its conjuncts have been verified.

Neither of these strategies, however, is proprietary to disquotationalism. An opponent of disquotationalism might have similar complaints.

Indeed, as Field (1994, p. 250) emphasizes, what is distinctive about contemporary disquotationalism is precisely the thought that one needn't follow verificationists in rejecting the notion of truth, nor follow anti-realists like Sir Michael Dummett (1991) in attempting to reduce truth to something else, such as justification. Rather, disquotationalism is a sort of quietism: It allows one to accept the notion of truth *and the uses typically made of it* so long as one can somehow construe those uses as really being just 'expressive'. But that is to say that it is precisely its embrace of the Classical Disquotational Strategy that distinguishes contemporary disquotationalism from earlier views, such as verificationism and anti-realism, that strive either to do without the notion of truth or to reconstruct it in other terms.

To put it differently, the promise of disquotationalism is that, instead of having to articulate an alternative to the truth-conditional conception of content, we can simply make do with the 'naïve' conception of content that is implicit in and wholly explained by such trivialities as

(11) "Snow is white" expresses the proposition that snow is white.

---

[10] We'll not actually consider disquotational accounts of such 'compositional principles' here. I've discussed them elsewhere, however (Heck, 2019).

As Stephen Leeds (1995, p. 4) puts it, disquotationalists believe that (11), and schematic generalizations thereof, tell us "everything there is of interest to know about how our language, and other languages too, connect with the world...". There's simply no need for a theory of content in the sense in which the causal theory was supposed to be one.

My primary goal in this paper is to demonstrate that the Classical Disquotational Strategy fails. I will begin, in section 1, by arguing that translation must play a much greater role for the disquotationalist than is usually acknowledged. This observation is really just a manifestation of familiar points about what the bearers of truth might plausibly be, but I'll argue in section 2 that disquotationalists have paid far too little attention to that issue, and that it has a good deal of significance for their view. The rest of the paper investigates the consequences of incorporating an appeal to translation into the disquotationalist's substitutes for putatively truth-involving explanations. We'll discuss two sorts of cases: The interpretation of context-dependent expressions, which is the focus of section 3, and what might be the most familiar argument against disquotationalism, the so-called 'success argument', which will be the focus of section 4. Ironically, it will turn out that failure matters much more than success.

1   Disquotation and Translation

As Field (1994, p. 260) emphasizes, the thesis that $\ulcorner$"S" is true$\urcorner$ is 'fully cognitively equivalent' to S itself entails that truth applies, in the first instance, only to sentences one understands—and so not, e.g., to sentences of a foreign language one does not speak. That might seem surprising. Even if one does not oneself know what "La nieve es blanca" means, for example, one might have thought that one could nonetheless understand what it means to say that the sentence "La nieve es blanca" is true. But, according to disquotationalism, that is an illusion. The reason is straightforward: If an attribution of truth to a sentence is 'fully cognitively equivalent' to an utterance of that very sentence, then "'La nieve es blanca' is true" is fully cognitively equivalent to—that is, roughly speaking, synonymous with—"La nieve es blanca" itself. Since, by hypothesis, that is not a sentence one understands, one does not understand "'La nieve es blanca' is true" either.

In fact, however, what this argument shows is only that a disquotationalist cannot make do just with what Field calls a 'pure' disquotational

notion of truth: one explained entirely in terms of the Disquotation Principle. Rather, there is a need also for what Field calls an 'extended' disquotational truth-predicate, which is explained in terms of translation. A sentence I do not understand is true in the 'extended' sense if it can be translated by a sentence I do understand that is true in the 'pure' sense.[11] This still has the consequence that I do not myself understand such sentences as " 'If there are measurable cardinals, then there are sets that are not constructible' is true", due to my lack of knowledge of set-theory, which one might find implausible. But I shall set this concern aside here.[12]

One might worry that appealing to translation will bring in content through the back door. But the disquotationalist can insist, with Quine, that the standards of correct translation do not have to be explained in terms of identity (or even similarity) of content, but can instead be explained in broadly pragmatic terms. There might then be no single 'correct' translation, but only a range of equally acceptable ones. That is not obviously a problem, however (Field, 1994, p. 273).

It can seem natural and even obvious that "true" applies, in the first instance, only to sentences one understands. I have often had people say to me: *Surely* we first learn what "true" means by learning to apply it to sentences we understand![13] This is supposed to be easy for us to do because the rule for applying "true" is so easy to learn: One should be prepared to assert ⌜"S" is true⌝ just in case one is prepared to assert S itself. If so, then the truth-predicate we initially acquire seems to be a disquotationalist one, characterized in terms of something like the disquotation principle.

We'll return to the question whether any of that is right. What is quite certain, however, is that the thesis that truth can, in the first instance, only be applied to *sentences* that one understands is, like disquotationalism itself, a *much* more radical claim than is usually acknowledged. The problem is that there are very few *sentences* that are ripe for disquota-

---

[11] That is:

$$\mathsf{T}(\Sigma) \equiv \exists S[S \sim \Sigma \wedge \mathsf{T}(S)]$$

where Σ is a 'foreign' sentence and S a 'native' one. Note the similarity to how propositional deflationists explain sentential truth.

[12] Stewart Shapiro (1998, pp. 55ff; 2003; 2005) has developed an argument that purports to show that this restriction causes trouble for disquotationalism. Field (2001b, pp. 147–8) discusses the argument briefly.

[13] And if I tell them not to call me "Shirley", then they often add: What, you think we start by applying "true" to sentences we *don't* understand?

tion: sentences S that are, at all plausibly, 'fully cognitively equivalent' to the corresponding sentence ⌜"S" is true⌝. The most obvious problem is posed by sentences that involve context-dependence, such as "It is cold here" or "You are a philosopher" or "That is a banana". Familiarly, it makes no sense to ask whether such sentences are true or false.[14] But it has become increasingly clear over the last couple decades that it is not just the usual suspects—"I", "here", "you", "this", and the like—that cause problems, but that context-sensitivity is nearly ubiquitous, affecting even logical terms such as quantifiers. (Consider the many different things an utterance of "Everyone is on the bus" might mean.) In fact, it is difficult to think of any sentence outside mathematics and the 'official' pronouncements of the sciences that does not exhibit some degree of context-dependence, if only for tense.

If I want to say, then, that an utterance U made by someone else, or by myself at some other time, is true, then I need to invoke translation: U is true iff there is some sentence that I understand which, if uttered by me now, would both correctly translate U and be true. The 'pure' disquotational notion of truth thus applies, in the first instance, only to utterances made by me, at the present moment, of a sentence I understand (cf. Heck, 2004, §4; David, 2005, p. 389). Such a notion has limited application, to put it mildly.

A similar problem arises with utterances made by other people, even when these do not involve context-dependence. I cannot simply assume that anyone else means by their words what I mean by mine.[15] I cannot, that is to say, simply assume that any other speaker's utterance of a given sentence will be true just in case an utterance by me of the same sentence would also be true (let alone fully cognitively equivalent).[16] Similar remarks apply to utterances of my own made at other times. If

---

[14] Quine (1970b, pp. 13–4) notes both that disquotation only really applies to 'eternal' sentences and that truth must (largely for that reason) be applied to utterances (or token sentences), but he never seems to consider how those two facts can be reconciled.

[15] It is sometimes suggested that we can just relativize attributions of truth to a scheme of translation and ignore what the speaker's words actually mean. But even Field (1994, p. 274) recognizes that there will be contexts in which it matters very much that the translation be a (if not the) correct one.

[16] Field (2001a) has suggested that, for a disquotationalist, sentences should be typed in terms of their 'computational' properties. That is a theoretical move that can be made in response to the sorts of worries I'm expressing, and we'll discuss a version of it below. But my goal at the moment is just to undermine any sense one might have had that a 'pure' disquotational notion of truth can do any real work by itself.

so, however, then the thesis of 'full cognitive equivalence' between S and $\ulcorner$"S" is true$\urcorner$ almost never holds.[17]

None of that would be news to Field. It is for precisely such reasons that he insists that "true" applies, in the first instance, to sentences of one's own idiolect—or, perhaps, to sentences of one's own language of thought. But what I am arguing is that such a restriction comes at a significant cost, namely, that the 'pure' disquotational notion of truth is all but useless by itself. If a disquotational notion of truth is to do any work, we need to use the 'extended' notion, which embeds an appeal to translation.

2   The Generalizing Role of Truth

There is nothing new in the observations made in the last section. What I was doing was just rehearsing the usual reasons to think that sentences are not plausibly the fundamental bearers of truth. But it seems to me that disquotationalists have failed to appreciate the importance of this familiar point to their views.[18] For one thing, it implies that it is utterly implausible that we 'learn the truth-predicate' by applying it to *sentences* we understand. There are, as we have seen, almost no 'sentences we understand' in the required sense: sentences that are 'fully cognitively equivalent' to attributions of truth to those very sentences. Indeed, as Sir Peter Strawson (1950, p. 130) famously pointed out, attributions of truth to sentences (and even to utterances) are quite rare in natural language. The truth-predicate in natural language is usually applied to propositions.

It is worth belaboring this point, both because it is easy to misunderstand it and because it is relevant to the sorts of claims that disquotation-

---

[17] One might think that context-dependence is simply beside the point, because disquotationalists are only really interested in 'scientific' language. What's worrying Field, on this reading, is whether the notion of truth needs to play a role in serious science, e.g., in physics, in which case context-dependence is not an issue. But notions of truth and content are arguably deployed *within* science, in particular, in linguistics and in cognitive psychology. It has long been an important question in natural language semantics how truth-conditions for (actual and possible) utterances of context-dependent sentences can be compositionally generated. And many cognitive psychologists are interested in questions about the contents of perceptual representations that are naturally understood as questions about their truth-conditions (see e.g. Burge, 1979). Field clearly takes questions of the latter sort very seriously.

[18] In part, I suspect, because of the focus, in much disquotationalist writing, on formal languages.

alists like to make about the 'role' the truth-predicate plays in natural language. Suppose Toni says:

(18)   Not all of the things Obama said were true.

The standard disquotationalist claim is that, in making this utterance, Toni says the same thing she would have said had she instead uttered the negation of the conjunction of the various sentences that Obama uttered, and in some strong sense of "the same" (cf. Gupta, 1993). But, since the sentences Obama uttered were almost certainly context-dependent, that is false. Minimally, then, what's needed here is an appeal to translation. But there is a more serious problem with this sort of suggestion: It assumes that (18) quantifies over sentences, and it pretty clearly does not.

What I mean by this is simply that, if Toni utters (18), then while I could perfectly well ask:

(19)   Were any of them true?

it would be at best bizarre for me to ask:[19]

(20)   Did any of them have plural subjects?

To put it differently, if you were to ask Toni, "And what things did Obama say that were not true?" you would not be likely to get an answer like:[20]

(21)   Obama said "There are pigs in the Rose Garden".

but instead one like:

(22)   Obama said that there were pigs in the Rose Garden.

When I say that (18) does not "quantify over sentences", then, what I mean is that its 'instances' are (related to) things like (22), not to things like (21). This is also what Strawson meant when he said that attributions of truth to sentences are rare in natural language. Ordinary speakers just do not often say things like "The sentence 'John broke the window' is true". They do say things like, "It's true that John broke the window, but it was an accident".[21] This is at least partly because

---

[19] Contrast: Did he use sentences with plural subjects to say any of them? Note also that (20) is a perfectly reasonable question to ask in response to: All of the sentences Obama uttered were really long.

[20] To take a different case, Toni might say, "Obama said that he is the president". That is very different from her saying, "Obama said, 'He is the president'", even if Obama was pointing at himself when he said it.

[21] Compare: The sentence "John broke the window" is true, but it was (is?) an accident. Points in this same vicinity pose problems for Davidson's paratatic analysis of that-clauses (Higginbotham, 1986, p. 39).

of the ubiquity of context-dependence, which implies that attributions of truth to sentences are usually non-sensical. Even "Snow is white" is context-sensitive, at least with respect to "is" and "white", and probably also with respect to "snow".

The issue here is thus not how that-clauses and other devices of 'propositional reference' are to be analyzed. Perhaps that-clauses refer not to propositions but to representations of some sort (though presumably not to sentence-types if only, again, because of context-dependence). Even if so, the point I am making concerns what ordinary speakers need to know to evaluate such utterances as (18), namely, things like:

(23) It is true that there were pigs in the Rose Garden iff there were pigs in the Rose Garden.

They do *not* need to know things like:

(24) "There are pigs in the Rose Garden" is true iff there are pigs in the Rose Garden.

Indeed, such instances of the sentential T-scheme are not only not needed to evaluate (18), but they are worse than useless. Due to the context-dependence of the sentence that occurs on its right-hand side, the *sentence* (24) cannot even be evaluated for truth. Moreover, if it is Obama's utterance of "There are pigs in the Rose Garden" that is at issue, then the best disquotation can give us is:

(25) Obama's utterance of "There are pigs in the Rose Garden" is true iff there are pigs in the Rose Garden.

But no utterance of (25) is guaranteed to be true unless it is made at the same time as Obama's, and even then there are other aspects of context-dependence to consider.

The claim that we need a disquotational notion of (sentential) truth in order to make sense of such statements as (18) thus looks to be false. Plenty of people manage that trick, and there is no evidence that any of them use a disquotational notion of truth to do so. It is a different question, of course, whether one *could* use a disquotational notion of truth to make sense of (18). I doubt it.[22] But that is not what

---

[22] The reason we are supposed to need a disquotational truth-predicate is because we need $\ulcorner$ "S" is true $\urcorner$ to be equivalent to S not just in some weak sense but in the strongest possible sense. Once we introduce translation, though, any such strong equivalence will lapse.

disquotationalists have typically claimed. Rather, they have wanted to argue that one *must* have a disquotational notion of truth in order to be able to make sense of such generalizations, which fact is supposed to force their opponents to accept at least the legitimacy and utility of the disquotational notion (see Heck, 2004, §1) . But there is no 'must' about it. A propositional notion of truth will do just as well—and, contrary to what Field (1994, pp. 266–7) claims, such a notion need not necessarily commit us to "strange entities". Whether it does depends upon how complement clauses are to be analyzed. That is an important issue, to be sure, but it is simply orthogonal to the one presently under discussion.

3   Translation and Linguistic Comprehension

As noted in section 1, sentences uttered by other speakers always might mean something different for them from what they mean for me.[23] As a result, attributions of truth to utterances made by other speakers must use the 'extended' disquotational truth-predicate and so will involve translation. The crucial case, which we will discuss in section 3.2, is that of context-dependent utterances. Before we turn to it, however, it will be worth discussing the more general view that the role typically played in semantic theories by truth-conditions can instead be played by translation—a view with which Field (1994, pp. 278–9) expresses some sympathy. The lessons that emerge will be important in the discussion that follows.

3.1   Translation vs Truth-Conditions

To be more precise, the view I want to consider is that one understands someone else's speech when one knows how to translate it into one's own language. Something like this view is obviously present in Quine (1960, Ch. 2), and it famously surfaces as well in the work of Jerrold Katz and Paul Postal (1964). I do not take this view to need refuting. Ernie Lepore and Barry Loewer (1981) have, to my mind, already completed that task. What Lepore and Loewer do not do, however, is diagnose why such a view should ever have seemed attractive—and why, apparently, it continues to be attractive. That is what I mean to do. I will be arguing that this 'translational conception of interpretation' rests upon a confusion of use with mention. Such confusions are themselves perenially tempting.

---

[23] As Quine (1968, p. 199) famously put it, "...radical translation begins at home".

Donald Davidson (1967) famously suggested that successfully interpreting the utterances of other speakers involves coming to know such things as:[24]

(26) Maria's utterance of "That woman is famous" is true if, and only if, Eva Longoria is famous.

By contrast, Jerry Fodor (1975, pp. 119–121), among others, has argued that nothing so complicated as (26) is needed. Instead, we should think of 'semantic competence' as consisting in an ability to translate natural language into the language of thought.

Here is a very rough model, but one that will do for our purposes. When someone utters a sentence, what I first do is determine how the sentence is composed of its component words. Perhaps this is quite complex; perhaps the 'words' are not at all what we'd ordinarily call 'words'. But once the syntactic analysis is complete, all that is left for me to do is to translate each 'word' that occurs in the original sentence by a corresponding expression of Mentalese, and then to put the Mentalese words together in a way that is determined by the syntactic analysis. When I'm done, I'll have a Mentalese sentence that translates the original sentence of natural language, and I can use it to interpret the present utterance of that sentence.[25]

On the translational conception, then, a central role is played in interpretation by competent speakers' possession of a sort of table that pairs 'words' of natural language with corresponding 'words' of Mentalese.[26] And, in many ways, that seems like a very natural idea. But we need to

---

[24] I'll assume, for present purposes, that all sides are agreed that understanding is constituted by some sort of propositional knowledge, though that knowledge may only be tacit. (My own view is that understanding is constituted by conscious knowledge of such facts as that expressed by (26) (Heck, 2006).) The issue here is what form that knowledge takes, in particular, whether it involves knowing how to translate. Someone who thought that understanding involves knowing that Maria's utterance *meant* that Eva Longoria is famous would thus be on Davidson's side, as far as the present issue is concerned.

[25] This is not unlike the view, which Davidson (1967, pp. 307–8) discusses in "Truth and Meaning", that a theory of meaning can consist just of syntax plus a dictionary. His objection is that such a theory simply fails to address the problems that exercise semanticists. He did not, I think, anticipate that someone might simply deny that those problems are actually of much interest, which seems to be how Field (1994, p. 269) would respond.

[26] If it were not for the creativity and productivity of linguistic competence, we could make do with a table pairing sentences of English with sentences of Mentalese. The issues would be no different.

Big Papi	David Ortiz
The Truth	Paul Pierce
Baby Horse	Alex Morgan

Table 1: Nicknames of Athletes

be careful. There are several different ways one might think of such a table and the information it contains.

Consider, for example, table 1.[27] One way to read the table is as pairing nicknames of athletes with their given names, in which case the first line might be written more explicitly as:

(27) "Big Papi" is a nickname for the person whose given name is "David Ortiz".

So, on this interpretation, the table purports to describe a relation between names, that is, between linguistic expressions. Fully to appreciate the content of the table, so understood, one does not have to understand any of the names contained in it. Even if you have never heard any of these names before, that will not affect your understanding of the table.

A second construal would regard the table as pairing nicknames of athletes with those very athletes, that is, with the people who have those nicknames. In that case, the second line might be written more explicitly as:

(28) "The Truth" is a nickname for Paul Pierce.

In this case, the table describes a relation between between words and people. Fully to appreciate the content of the table understood this way, one *does* have to understand the names that occur in the right column, though not the names that occur in the left column: If one has never heard of Paul Pierce, then one cannot understand (28).

Yet a third construal (we'll stop here) would regard the table as purporting to state a number of true identities, in which case the third line might be written more explicitly as:

(29) Baby Horse is Alex Morgan.

---

[27] Think of the nicknames, on the left, as corresponding to names from natural language and the given names, on the right, to names in Mentalese. So the table is supposed to be telling one something about the nicknames.

In this case, the relation the table describes is wholly worldly. Fully to appreciate the content of the table understood this way, one must understand the names that occur in both columns.[28]

I hope this all seems somewhat confusing. It's not that it's unclear what the three construals are. But how did you understand the table when first you encountered it? In practice, the three construals have an annoying tendency to bleed into one another. Even if the table had been intended the second way, as pairing nicknames with people, someone who was completely unfamiliar with the names on the right could nonetheless read the table in accord with the first interpretation and thereby acquire information about co-reference. And even if the table had been intended the first way, someone who was familiar with the given names could nonetheless read it in accord with the second way and so extract information about the reference of the nicknames. It just is very easy to slide back and forth between use and mention this way, especially where written language is concerned: We simultaneously read the names with understanding and see them as objects in their own right.

I take it, however, that what is at issue between Davidson and Fodor is *what information* competent speakers have that allows them to interpret speech (cf. Peacocke, 1986, 1989). And what information table 1 contains varies enormously depending upon how we interpret it. So we must be very careful how we interpret the 'table' that, according to the translational conception, partially underlies semantic competence. The question we need to ask is: On which construal of such tables do they encode information possession of which would allow someone successfully to interpret speech, i.e., in this example, to understand utterances of the nicknames?

The third construal is clearly no help. On this construal, the information contained in the table is simply a bunch of identities. It is irrelevant that these truths are *recorded in* a mix of, say, English and Mentalese. Worse, because the names in the left column are, on this construal, used, one must already understand them in order to be able to understand the table. The table, so construed, thus cannot play any role in enabling

---

[28] I take it that this case is not what philosophers of language usually have in mind when they speak of 'translation'. That term tends to be reserved for a relation between linguistic expressions, as on the first construal. Colloquial language seems to disagree, however. The third construal is reminiscent of what 'translators' at the United Nations do: They listen to speech in one language and then repeat what they have heard in another language. We might call this enterprise *restatement*.

one to interpret utterances of the nicknames. Only a table in which the names in the left column are mentioned can play that sort of role.[29]

The first construal at least satsifies that condition, but it is no help, either, for the reason given long ago by David Lewis (1970, p. 18): One can know how to translate sentences from one language to another without understanding either of those languages.[30] For our purposes, the right way to formulate this point is: The *information* that a translation manual contains is insufficient to allow one to understand the language being translated. It's true that, if one does understand one of the two languages, then one can parlay one's knowledge of how to translate into knowledge sufficient to allow interpretation.[31] But that, once again, is irrelevant: The question was *what information enables interpretation*, and what Lewis is claiming is that information about how to translate one language into another is insufficient by itself. He is just right about that.

It is when the table is interpreted the second way that it encodes information sufficient to allow interpretation: What you need to know is, for each nickname, who bears that name; that is what the table tells you. To be sure, the table conveys this information by *using* the given names of the people in question. But that does not imply that the table contains (let alone, only contains) *information about* those people's given names. On the contrary, what the table (so understood) does is map nicknames to *people*, the people who have those nicknames. So if, in some sense, comprehension of natural langauge uses such lookup tables, it will have turned out that spanning the great chasm between language and the

---

[29] A different route to this point would begin with the observation that, when we are attempting to understand what someone else has said, we start by identifying the *linguistic expression* they have produced. Certainly, this seems to be the model with which linguists operate.

[30] Davidson (1973, pp. 316–8) makes the same point a few years later—independently, as far as I can tell.

[31] It is far less clear than most people seem to think exactly how this is supposed to happen. How is the transition from the mentioned sentence to the used one made? In the ordinary case, one just 'reads' the quoted sentence. But what if one can only recognize the quoted sentence as a name and not read the embedded sentence with understanding? And what anyway is the analogue of 'reading' in the case of Mentalese? One might think of disquotationalism as offering, among other things, an answer to that question. But it is surprising that we are now supposing that language comprehension involves one's forming *names* of sentences of Mentalese.

world requires no more than a data structure that maps names to their bearers.[32]

If one is tempted to respond that people cannot literally be contained in such a data structure, then one is succumbing to what appears to be an almost irresistable temptation to confuse *what information is contained* in a data structure with *how that information is encoded*. As we are imagining it, the map from names to people is encoded in a table-like structure that competent speakers have in their heads, one that is *written* in Mentalese. So, of course, people are not contained in the table. They are *mentioned* in the table through the *use* of their Mentalese names.[33] To be sure, making use of the information that the table contains involves computing in Mentalese. But, if there is a language of thought, as we are for the moment assuming, then all thought, whatever it is about, involves computing in Mentalese.[34]

Surely it has to be uncontroversial—surely no one wishes to deny—that we are, in some sense, able to think about the external world, e.g., about other people.[35] Nor is a proponent of the translational conception in a position to deny that competent speakers are able to think about linguistic expressions, since such a capacity is presupposed by the ability to translate.[36] But then there cannot possibly be any obstacle in

---

[32] I find it helpful to think, in this connection, about strongly typed programming languages. If you want a program that will 'interpret decimal numerals', for example, then what you will need (at the level of digits), in C++, is a `map<char, int>`, which is very different from a `map<char, char>` (the first construal) or a `map<int, int>` (the third one). Similarly, converting decimal numerals to integers is a very different programming problem than converting them to binary numerals (a much easier one, actually), even if integers are represented, in the machine, in binary.

[33] Indeed, the same is true of the natural language names: They too are only mentioned in the table, through the use of *their* Mentalese names.

[34] This seems to be what Fodor overlooks. He writes:

> Pretty obviously, there are computational procedures which map a representation of the acoustic properties of a speech event onto a representation of the message it encodes. (Fodor, 1975, p. 119)

But *the function being computed* is one from acoustic properties to 'messages', i.e., propositions or truth-conditions or what have you.

[35] The real issue here, of course, is in what sense precisely we are able to think about other people. We'll turn to that issue in section 4. My point here is just that the question whether linguistic comprehension involves translation into Mentalese is orthogonal to that issue.

[36] Independently of that, I take it to be less controversial than it might seem that competent language-users are able to think about expressions. Linguistic theory gives us good reason to suppose that this capacity is innate (see e.g. Chomsky, 1986), though, at the early stages, such 'thought' might occur only sub-personally.

principle to our using whatever mechanism allows us to do those two things—think about people and think about their names—to construct a data structure that pairs expressions of natural language with people. Nor is there any obstacle to our using that information, and related information about other types of linguistic expressions (e.g., predicates), to 'compose' interpretations of complete sentences according to their syntactic structures.

It is still open to Fodor, of course, to suggest that the *mechanism* by means of which knowledge of meaning is generated somehow involves translation into Mentalese. But I know of no very good reason to believe that, and it is far from clear what advantage the translational conception is now supposed to have.[37] What motivated it seems to have been some vague sense that attributing to ordinary speakers an ability to think about word–world relations is too demanding (if not spooky). But if we can think about words and we can think about the world, then we can think about relations between words and the world. Words, after all, are just some of the things we find in the world.

3.2   Translation and Context-Dependence

Since almost all sentences are context-dependent, there are hardly any sentence types to which it makes sense to apply the notion of truth. We can apply the truth-predicate to utterances instead,[38] but in that case no sort of Disquotation Principle is plausible: If U is an utterance of a context-dependent sentence Σ, then an attribution of truth to U typically will not even have the same truth-value as an arbitrary other utterance of Σ, let alone be 'fully cognitively equivalent' to such an utterance. What's needed, therefore, is an 'extended' disquotational notion of truth: An utterance of Σ will be true iff there is some sentence of my language an utterance of which would both correctly translate it and be true—whatever 'correct translation' might be.

Consider a simple example. Suppose Maria utters:

---

[37] It is true is that the compositional problem appears to be far more difficult in the semantic case than it is in the translational case. The latter involves purely syntactic operations on Mentalese symbols, whereas the former involves possibly quite complex compositional axioms. On the other hand, as mentioned in note 31, the translational approach requires a meta-language for Mentalese and then some sort of 'semantic descent'. So it is not obvious that either view has an advantage on grounds of simplicity.

[38] Or to sentence tokens, or something of the sort. Such niceties will not matter in what follows.

(30) That woman is famous.

In order to focus attention squarely on the relevant issue, let us assume that Maria is speaking the same language I am, so that there is no question what sentence I need to utter in order to translate Maria's utterance, namely, (30) itself.[39] But, of course, any utterance of that sentence by me will also be context-dependent, so I need to accompany the utterance by a 'demonstration' of some object.[40] Indeed, given our simplifying assumptions, the *only* thing I need to decide, to figure out how to translate (30), is which person to demonstrate. Suppose the right person is Eva Longoria. Then a correct translation of (30) would be:[41]

(31) That woman [said while demonstrating Eva Longoria] is famous.

Now, the discussion in the previous section shows that the information that Maria's utterance of (30) could be correctly translated by an utterance by me of (31) is of limited utility. That information will, by itself, allow me neither to interpret Maria's utterance nor to understand attributions of truth to it. What I ultimately need to know, for either purpose, is something like:

(32) Maria's utterance of (30) is true iff that woman [said while demonstrating Eva Longoria] is famous.

What the disquotationalist seems to be proposing is thus, once again, that the *mechanism* by which (32) comes to be known involves two steps: translation of (30) by (31) and an application of the Disquotation Principle to a sentence attributing truth to a contemporaneous utterance, by me, of (31).[42]

To focus just on the demonstrative, what I ultimately need to know about it is:

(33) When Maria uttered "that woman", she was talking about that woman [said while demonstrating Eva Longoria].

---

[39] We will also ignore the other sources of context-dependence, besides the demonstrative, that are present here.

[40] In fact, as I have emphasized elsewhere (Heck, 2014), demonstrations are optional, and context-dependence is typically resolved in other ways. But we can again ignore this issue.

[41] A different proposal is that (30) should just be translated as "Eva Longoria is famous". The ensuing discussion would be different in detail but similar in spirit. See note 44.

[42] I shall talk here in terms of what is involved in understanding an utterance of (30). The discussion could also be framed more metaphysically, in terms of the facts about what such utterances mean.

And the proposal is that I come to know (33) by deriving it from:

(34) Maria's utterance of "that woman" can be correctly translated by an utterance made by me, while I demonstrate Eva Longoria, of "that woman".

But that just seems backwards. To know which person I should demonstrate in uttering (34), I need to know which woman Maria was talking about when she uttered the words "That woman"; that is, it seems as if knowledge of (34) rests upon knowledge of (33). Such facts are semantic if any facts are. Translation of context-dependent utterances therefore cannot proceed independently of semantic facts.

A disquotationalist might reasonably object that this argument tendentiously assumes the legitimacy of such semantic notions as *talking about*. But the argument really needs just the following two claims:

(i)   How I should translate an utterance of a sentence like (30)—in particular, whom I should demonstrate—depends, in part, upon *some* relation between the speaker and an object in the world.

(ii)   The mentioned relation must be explicable independently of translation.

To see why, note that there is a 'pure' disquotational notion of reference that is supposed to be characterized by trivialities like:

(35) "Eva Longoria" refers to Eva Longoria.

But, for reasons parallel to those already rehearsed, the 'pure' notion is inadequate in the case of demonstratives and other context-dependent expressions.[43] The disquotationalist thus requires an 'extended' notion of reference, which we may call *E-reference*. This notion, like the extended notion of truth, is presumably to be explained in terms of translation, thus:[44]

---

[43] It is a fairly common view nowadays that even proper names exhibit some sort of context-dependence (see Bach, 2015, for an overview). If that is right, then disquotational principles like (35) are inadequate to account for the reference of utterances of proper names, and all the issues we are discussing here will arise even in that case.

[44] One might suggest that we should instead try:

(*)   Maria's utterance of "That woman" E-refers to Eva Longoria iff her utterance can be correctly translated by an utterance made by me of "Eva Longoria".

But, first, the left-to-right direction arguably fails, for sense–reference-type reasons, though that does depend upon how demanding the standards of 'correct translation'

(36) Maria's utterance of "that woman" E-refers to Eva Longoria iff
      Maria's utterance can be correctly translated by an utterance made
      by me, while I demonstrate Eva Longoria, of "that woman".

Now, the question at issue here, to re-iterate, is what facts determine
which translation of Maria's utterance is correct, which is to say: what
facts determine whom I should demonstrate when translating her. We
obviously cannot just assume that those facts are 'semantic'. But—and
this was the first claim mentioned above—I submit that those facts must
involve *some* relation between Maria and Eva Longoria. That claim does
not beg the question against the disquotationalist. There are plenty of
non-semantic relations between speakers, utterances, and the rest of the
world. Perhaps one of them will do the work the disquotationalist needs
doing.

 If someone wanted to insist that the burden of proof here is on the
disquotationalist—that they need to tell us which relation that is—then
I wouldn't disagree (though I am allergic to burden-of-proof arguments).
It really is very unclear what non-semantic relation might determine
what the correct translation is. There was a time when one might
have been forgiven for thinking it was *pointing at*, but it has long been
appreciated that pointing is not necessary for demonstrative reference.[45]
Nonetheless, *pointing at* is a useful example, because it nicely illustrates
the sort of relation the disquotationalist needs. But here's the thing: If
pointing at were the right notion, then E-reference would be co-extensive
(and not just materially) with *pointing at*. And the same goes for any
substitute that might be offered. It follows—here we apply the second
premise mentioned above—that, by the disquotationalist's own lights, E-
reference must be co-extensive with some relation that is characterizable
independently of translation. But that makes E-reference, whatever the
right relation is, a non-disquotational notion of reference.

 There are several replies a disquotationalist might make. They might
deny the first premise. Indeed, Field (1994, pp. 279–81) did once argue
that in explaining how to translate Maria's utterance of "that woman"
we can make do just with facts about her "internal processing" and so
need not invoke any relation between Maria and the outside world. The

---

are. Second, if names are context-dependent (see note 43), then (*) does not avoid
the problems to be discussed next. Finally, it still seems to be in order to insist that
the correctness of the translation mentioned depends upon there being some relation
between Maria and Eva Longoria.

 [45] Indeed, *pointing at*, in the relevant sense, probably isn't a purely physical relation
but in part a psychological one (cf. Reimer, 1991).

idea was that which woman I should demonstrate will be determined by what other 'mental files' Maria associates with this particular utterance of "that woman". There are several difficulties with this view, but the simplest is that there need not be any other such mental files: Maria may never before have heard of or encountered Eva Longoria (Heck, 2004, pp. 339–41).

A second reply denies that there is any *single* relation between the speaker and an object in the world that grounds E-reference: In each case, there will be some such relation, but there is no single relation that does the job in every case.[46] By itself, however, such an observation would pose no threat to the preceding argument. It is one of the lessons of recent work on demonstratives that many different relations between a speaker and an object can support demonstrative reference (King, 2014; Heck, 2014). So the view would have to be that there is so much diversity that there is really no unified notion there at all, independent of the characterization given in terms of translation. All we can say is that the correct translation is the one that allows us to make the best sense of Maria. We'll consider this sort of reply, which in effect denies the second premise, in section 4.2.

Yet a third reply is to concede both that there is such a relation and that it can be explained independently of translation, but to deny that E-reference has very much to do with *reference* as friends of semantics would understand it. As we will see, this is probably the reply Field would prefer: He holds a similar view about truth-conditions, in part because of problems that arise when one actually tries to explain reference in terms of translation. We'll consider this sort of reply in section 4.3.

---

[46] It is unclear to me whether anyone has ever held such a view. Both Quine and Davidson regard translation (or interpretation) as part of what determines content. But neither rests with vague gestures in the direction of 'making sense' of speakers. Both of them articulate conceptions of the underlying facts about speakers to which a theory of translation (or interpretation) must be sensitive, and radical translation (interpretation) is simply a dramatization of the question how facts about correct translation (interpretation) supervene upon the underlying basis. For Quine, those facts are limited to ones about how Maria's sense organs have been affected by her environment, e.g., how light has stimulated her retina (Quine, 1960, §8). For Davidson, they include facts about what sentences Maria has 'held true' under what external circumstances (Davidson, 1973, pp. 322ff). What grounds the correct attribution of truth-conditions to Maria's utterance of (30) for Davidson, then, does include relations between Maria and objects in the world. If it doesn't for Quine, that is because Quine is a behaviorist, not because he thinks translation plays a role in fixing content.

Before we continue this discussion, however, it is worth making explicit a limitation of the argument we have been considering in this section.

3.3   Intentional Realism

As we have just recalled, for a disquotationalist, the correctness of a translation is to be decided on broadly pragmatic grounds. This feature of the view marks an important difference between disquotationalism and what is sometimes known as *intentional realism*: the view that semantic properties are robust enough to do some kind of causal or explanatory work (cf. Fodor, 1987). It's important to appreciate that the arguments given in this section are *not* intended to provide any support for intensional realism and, indeed, cannot do so. Let me explain why.

Quine (1960, Ch. 2) famously proposed that the way to investigate meaning is to investigate translation: If you want to know what Maria's words mean, and what it is for them to mean that, then the right questions to ask are: (i) how Maria's utterances should be translated into your language and (ii) what the appropriate standards of correct translation are. Quine articulates an answer to (ii) using the notion of radical translation. Davidson (1967), by contrast, thinks we should ask (i′) what the correct theory of truth is for Maria's language and (ii′) what the appropriate standards are by which to adjudicate correctness. Davidson (1973) articulates an answer to (ii′) in terms of radical interpretation.

As concerns the question what it is for Maria's utterances to mean what they do, then, Davidson and Quine are much closer to one another than either is to Fodor or to early Field. There are, of course, plenty of differences between Davidson and Quine here—most importantly, Davidson is no behaviorist—but Davidson's notion of radical interpretation is not just named after but explicitly modeled on Quine's notion of radical translation. In particular, both have a significant pragmatic component. The ultimate test both of a translation manual, for Quine (1970a, pp. 388–9), and of a theory of truth, for Davidson (1973, pp. 320–5), is whether it allows one to make sense of the behavior, linguistic and otherwise, of its subjects, e.g., to communicate smoothly with them.

The real difference between Quine and Davidson concerns how they answer the first question: what form a 'theory of meaning' for Maria's language should take. Quine thinks we can get by with translation, whereas Davidson (1973, pp. 316–8) thinks we must provide a theory of truth for Maria's language, one that will make heavy use of (non-disquotational)

semantic notions. That is why Davidson rejects deflationism (see e.g. Davidson, 1990, 1996), and the argument we have been discussing in this section is meant to be compatible with Davidson's overall outlook. But Davidson is not an intentional realist: Despite his occasional protests to the contrary, his insistence that facts about content essentially depend upon pragmatic factors precludes such a view (Rescorla, 2013, p. 480). So, unless Davidson's views are inconsistent at this point, which I do not think they are, it cannot suffice to establish intentional realism to identify a role for semantic notions in interpretation.

The arguments to be presented in the next section are different: They do favor intentional realism and so are not compatible with Davidson's views.

4   The Role of Falsity in Psychological Explanation

Perhaps the best known argument against disquotationalism is the so-called 'success argument'. The rough idea is that, if we're going to take our existence as rational agents seriously—if we're going to reject behaviorism and the like—then we need to think of our search for food, say, as guided by our beliefs about where food is to be found. Hence, our *success* in finding food will typically depend upon whether those beliefs are true. So if you want to explain how people manage not to starve, and you want to explain it (in part) in terms of people's cognitive capacities, then it looks as if the truth of our beliefs is implicated in that explanation. And that means, or so it would certainly seem, that truth is doing serious explanatory work, which is precisely what disquotationalism cannot abide.

This sort of argument has it origins in the work of Hilary Putnam (1975; 1978, pp. 17–33), but it is perhaps best known from the elaboration, analysis, and defense of it due to Field (1986, §V) just prior to his conversion to disquotationalism. The argument has since received a great deal more attention. It seems to me, however, that most of that discussion has done more to confuse the issue than to clarify it.[47] Indeed, Field describes his own earlier discussion of the success argument as "so abstract and convoluted it couldn't have convinced anyone..." (Field, 2001b, p. 153), and that is when he is being charitable. My goal here,

---

[47] A noteworthy exception is a recent paper by Will Gamester (2018), which I encountered long after this paper was drafted. I'll make some remarks about it in the footnotes.

then, is to offer a version of the success argument that, or so I will be claiming, gives us good reason to accept intentional realism and so to reject disquotationalism.

4.1 Navigation and Cognitive Maps

It will help if we step back a bit and consider the simple case of navigational behavior: our ability to find our way around the world. This is an ability that we share with many other creatures—among them rats, who, for all their other cognitive limitations, have a strikingly good ability to find their way around mazes. How do they do it? Part of my reason for considering this question is that serious scientific work has been done on it, and there are now some developed proposals that have significant empirical support. I'll focus on one of these here, not so much because I'm partial to it (though I am) but because the same sorts of issues would arise with respect to the other accounts on offer, and the one I'll discuss is not just relatively easy to understand but vividly illustrates what is at stake in the present debate.

Here, then, is one popular and reasonably well-supported explanation of how rats navigate mazes: Very roughly, rats find their way around by making use of little maps that they carry around in their heads.[48] These 'cognitive maps' are so-called because, or so some believe, they actually do have a map-like structure: Spatial relations between objects in the environment are represented, on the map, by means of geometrical relations among 'markers' that represent those objects on the map.[49] And rats are very good map-makers. As they move around, they continually update their cognitive maps, and they do so in such a way that, by and large, their maps end up accurately representing the topography of their local environment. So the reason rats are so good at finding their way around mazes is that, once they've had enough time to explore a particular maze, they have a map of it: They know how it is laid out.

There are two features of this account that will be important for what follows. First, the accuracy of the rats' maps is essential to the explanation. That rats have maps in their heads does nothing to explain

---

[48] The evidence for this view is nicely summarized by Michael Rescorla (2018). The classic discussion is in *The Organization of Learning*, by Charles P. Gallistel (1990).

[49] These maps are what Rescorla (2018, p. 381) calls 'cogntive maps in the strict sense'. Note that the geometrical relations need not be spatial but may be more abstract. In fact, though, it would not matter for our purposes if rats only had cognitive maps in the 'loose' sense: other mental representations of topographic aspects of their environments.

their navigational proficiency if the maps are no good. Second, the maps are essential, too. Simply to posit that rats have lots of information about the topography of their environment would simply beg the question how that is possible. The answer is meant to be that this information is encoded, as Edward C. Tolman (1948, p. 192) put it when he introduced the idea, in "a cognitive-like map of the environment. . . indicating routes and paths and environmental relationships. . . ".

What ties these two aspects of the explanation together is the *representational content* of the maps: We can speak of the maps as being correct or incorrect only because they have such content; the maps encode information only because it makes sense to ask whether they are correct. So Tolman's proposed explanation is committed to the claim that cognitive maps have something like truth-conditions. The crucial question, to which we shall turn shortly, is how the disquotationalist proposes to recast this sort of explanation.[50]

Before we address that question, however, let me emphasize that what Tolman offers us is not just an explanation of navigational *success*. Suppose Whiskers has been through the maze several times and can run it quite quickly. But now Peter cruelly decides to change the maze slightly. Off goes Whiskers, and the poor guy ends up crashing into a wall. Why? Because Whiskers thought there was a door there (i.e., his cognitive map represented there as being a door there). What the representational content of Whiskers's map contributes to, most immediately, is thus the explanation of his behavior, successful or otherwise. That is to say: What explains Whiskers's success, in the cases in which he is successful, is the combination of two other facts: (i) that his behavior is guided by his cognitive map of his environment and (ii) that his map is accurate. So, again: Representational content contributes most fundamentally to the explanation of behavior and only derivatively to the explanation of successful behavior. This point will be crucial later.

Let me also note that it does not matter whether one wants to *call* maps 'true' or 'false' or whether one wants to speak of them as having

---

[50] Just to emphasize: There is nothing special about this particular explanation. Similar points could be made about a wide range of psychological explanations that involve 'information processing'. Tyler Burge (1986a) discusses a wide range of such examples, and many of them could be used to make the same sorts of points I will be making here. The central question is how we should understand the notion of *information* that appears in such explanations. (The point I make next is also at least implicit in Burge.)

'accuracy-conditions' instead of truth-conditions.[51] The niceties of English usage are irrelevant. For one thing, this entire discussion could be reformulated, without loss, in terms of an alternative proposal according to which cognitive 'maps' have logical rather than geometrical structure, i.e., are language-like (cf. Rescorla, 2018, §6). But there is a good reason to discuss maps here, namely, because they make it clear that the *pure* disquotational notion of truth really is useless by itself, since what's fundamentally at issue *has nothing to do with language*.[52] The issue is what role, if any, the notion of representational content must play in our theory of the mind, e.g., whether, in explaining behavior, we need to invoke some notion of the representational content of mental states. It does not matter even a little bit whether that content can be linguistically expressed, if that's what it takes for it felicitously to be called 'true'.[53]

4.2   Translation and the Classical Disquotational Strategy

Disquotationalists are of course free to reject Tolman's explanation of rats' navigational proficiency and, indeed, to reject all explanations of mental processes in terms of computations over structured mental representations. But the question whether to accept the so-called 'representational theory of mind' has nothing obvious to do with truth: There have, in fact, been adherents of RTM who thought that we could do without the idea that mental 'representations' have content (Stich, 1983).[54] But the promise of disquotationalism, or so I argued above, is supposed to be that it does not require us to reject the sorts of uses to which truth is put in (cognitive) science, but only to recognize them as 'expressive'. So the question in which I'm interested here takes the form: Is disquota-

---

[51] Several people who attended presentations of this material become quite indignant about this matter, for reasons I have never quite understood.

[52] I.e., the real issue here is utterly independent of linguistic deflationism, which is a thesis about the meaning of the word "true", as it applies to linguistic objects (i.e., sentences).

[53] It is an empirical issue whether all mental representations have syntactic structure or whether some of them have other sorts of structure, such as the geometrical structure that maps have (see e.g. Fodor, 2007; Heck, 2007; Rescorla, 2009). It would be very unfortunate for disquotationalists if their views had implications for this debate.

[54] Although it is many years too late, I'd like to take a moment to thank Fred Dretske for introducing me to the issues we are discussing here. He taught a graduate seminar at Duke somewhere around 1985, when I was an undergraduate, and one of the core texts was Stich's then recent book. Both Stich's book and Dretske's reaction to it have stayed with me ever since.

tionalism compatible with Tolman's account? If it turns out that it is not, then I will be happy to rest my case. Disquotationalists should not be in the business of telling (cognitive) scientists how to do their jobs.[55]

How, then, might a disquotationalist understand the role apparently played by the notion of truth in Tolman's explanation of rats' navigational abilities?

The Classical Disquotational Strategy would be to show that the use made of the notion of truth here is merely 'expressive', i.e., that a disquotational notion of truth is all we need. Now, it should be clear that the 'pure' disquotational notion of truth will do us no good at all.[56] The maps that rats have in their heads are not 'sentences I understand', so I will need to use the 'extended' disquotational notion of truth if I'm to make sense of attributions of correctness to those maps. And that means, of course, that I need to consider how to translate the rats' maps into sentences of my language. A disquotational explanation of the rats' success would thus have the following form:

> Rats are good at running mazes because they navigate by little maps in their heads, and these tend, by and large, to be constructed in such a way that they are correctly translatable by true sentences of my language.

Or, to consider just the explanation of behavior:

> Whiskers attempted to run through a wall at location L because he was navigating by a cognitive map that is correctly translatable, in part, by the sentence of my language "There is a door at location L".

That is: If we pursue the Classical Disquotational Strategy, then the disquotational substitute for an explanation of a creature's behavior in terms of the information encoded in its mental representations will have

---

[55] There is a hilarious parable, due to Lewis (1991, p. 59), about a philosopher who seeks to convince mathematicians to change their ways by confronting them with philosophy's litany of successes. A similar warning applies here. A sharper one would emphasize the fate of Quine's *a priori* arguments for behaviorism.

[56] Gamester (2018, §§3–4) discusses the form the disquotationalists' substitute must take if this worry is waived—the basic idea goes back to Horwich (1990, pp. 22–3)—and argues that it fails anyway. But his argument is focused much more on truth than on truth-conditions, which are what I'm discussing here. (To put it differently, Gamester in effect grants disquotationalists a claim about what the contents of the mental states in question are. What I am questioning is their right to any such claim.)

to invoke translation. That should be no surprise. Translation is what disquotationalism substitutes for content.

I submit that the substitute explanations offered above on behalf of the disquotationalist clearly fail. It is plainly false that Tolman's rats were good at running mazes because the maps they constructed are translatable by true sentences of English.[57] That there are such things as people and natural languages has nothing to do with rats' navigational abilities: Rats would have been just as good at running mazes if there had never been any people, which is to say that the substitutes get the counterfactuals wrong.

The obvious reply is that what explains a given rat's behavior is not the existence of a certain translation but the underlying facts, whatever they may be, that make the translation correct, appropriate, or adequate. Talking about translation here is just a way of gesturing in the direction of those facts. Leeds (1995, pp. 28–9) compares this case to that of the meter stick: We can fix the length of a meter in terms of the length of a rod in Paris, but that doesn't imply that explanations that speak of meters somehow implicate that rod. Fair enough. But, as Leeds (1995, pp. 29–30) himself seems to recognize, that only raises the question which facts about the relations between a rat's map and its environment are really doing the explanatory work.[58] So we find ourselves, much as in section 3.2, returned to the question what makes a given translation of the rat's map correct, and my own suspicion is that it is the representational content of the map. The disquotationalist owes us some other

---

[57] Field (1986, p. 79) offers a version of this objection in his pre-conversion discussion of the success argument. I stumbled upon it independently around 2005 when thinking specifically about cognitive maps and only later became aware of the back and forth between Field and Leeds, which I'll discuss shortly. This argument is not discussed in "Deflationist Views of Meaning and Content" (Field, 1994), and, while Field mentions Leeds as an influence, Leeds's paper on these topics was only published the following year. (Hence, there is no reference to Leeds's paper in Field's.)

[58] I confess that I find Leeds's discussion of this issue extremely confusing. He ends up arguing that the crucial question is whether we can make sense of objective reasons:

> If we could find a way to make sense of th[e] statement [that A is a good reason for B] without mention of ourselves and our conceptual scheme...then we would have found our correspondence theory. The idea of looking for an 'objective' notion of reasons remains...the correspondence theorist's best hope.... (Leeds, 1995, p. 31)

But I cannot see why that should be. What do objective reasons have to do with questions about the representational content of cognitive maps in rats? Not that Leeds discusses this sort of case. But that is itself a symptom of the fundamental problem with Leeds's discussion: that he over-intellectualizes the issue.

answer. Otherwise, they haven't actually offered an alternative to the content-involving explanation. They have just expressed the hope that there might be one.[59]

As we'll see in the next section, Field does sketch such an explanation (suggesting that he takes the need for one seriously). Before we discuss his proposal, however, I want simply to point out that, the moment the need for such an explanation has been conceded, the Classical Disquotational Strategy has failed. For what has been conceded is precisely that the appeal to truth- or correctness-conditions in the sort of explanation we have been considering cannot be replaced by (or understood in terms of) 'disquotational' truth-conditions, whether these are pure or extended. I take this to show that we do at least *understand* a non-disquotational notion of truth and that there are possible (even actual) explanatory projects for which it is *prima facie* necessary.

The reason this point has been missed is that, as we saw earlier, disquotationalists tend to think of the truth-predicate as fundamentally a device of generalization. So their response to the success argument has tended to have two parts: First, *particular* explanations of success need not use the notion of truth at all, because applications of the truth-predicate to *particular* sentences can be eliminated via the Disquotation Principle; Second, *general* reliability can be explained by generalizing over particular explanations of particular successes. Gupta (1993, p. 67) has raised serious questions about the second part. But the first part has not received the same scrutiny,[60] and that has made it seem as if the only role plausibly played by the truth-predicate here is the very

---

[59] Field (2001b, p. 154, fn. 13) raises the question, crediting Leeds, why we "should think that there must be some means of specifying the relevant correlation [with external conditions] that doesn't go via translation...". But the disquotationalist has allowed that some 'correlation with external conditions' is what is doing the explanatory work, so they owe us an account of what kind of correlation that might be if it is not a representational relation. To put it differently: Even if we cannot *specify* what a meter is except in terms of some rod in Paris, *what it is* to be a meter long had better not essentially involve relations to that rod, since otherwise explanations that speak of meters *would* implicate such relations (and so, derivatively, that rod). So disquotationalists need there to be some account of what such a 'correlation with external conditions' might be that (i) can play the role the rest of us think truth-conditions play, (ii) does not involve semantic notions such as truth and reference, and yet (iii) is not explained in terms of translation. Disquotationalists can hardly expect the rest of us to take their word that there is some such notion available, so they need to say *something* substantial about what the 'correlation with external conditions' actually might be.

[60] Gamester (2018) is again an exception. A very similar sort of explanation is sometimes offered of the reliability of our mathematical beliefs. That makes me suspect that

one that disquotationalists like to emphasize. In fact, however, the first claim is also mistaken: The Disquotation Principle, by itself, cannot help us to explain behavioral success, even in particular cases, because serious explanations of behavioral success invoke mental representations, and mental representations are not 'sentences we understand'. So the explanation has to invoke translation, but then the problems discussed in this section doom the Classical Disquotational Strategy.

4.3   Truth and Indication

The question we have been discussing is how a disquotationalist might understand the role apparently played by representational content in map-based explanations of navigation. Appealing to translation does not help. It just begs the question what makes a given translation correct, and, whatever that is, that is what does the real explanatory work. So what exactly is it that does the real explanatory work, according to the disquotationalist?

It would be unfair to complain that Field nowhere gives us a full answer to this question. It is not as if it is clear how representational content enters into psychological explanation. But the request is not for a philosophical account of psychological explanation, but just for some reasonable indication of what the first-order explanation itself might be. We know very well what first-order explanation of rats' navigational proficiency the proponent of content-involving psychological explanation is offering. It was already outlined by Tolman.

Field himself has made important contributions to the study of how mental representations contribute to the explanation of behavior, beginning with his now classic paper "Mental Representation" (Field, 1978). In that paper and elsewhere, Field claims that (assuming something like the language of thought hypothesis) intentional explanations can always be recast as purely computational explanations (Field, 1986, p. 84; 2001b, pp. 155–6; 2001c, pp. 72–6). If we assume, as Field does, that mental processes are implemented computationally, then of course some computational process must always underlie any given intentional explanation. Moreover, as Field (2001c, p. 74) observes, the familiar objection that the computational story by itself cannot explain rats' navigational abilities depends upon our regarding computation as 'narrow'. If we instead take it to be 'wide', in the sense that "features of the external world [can]

---

the argument Øystein Linnebo (2006) gives against that view could also be deployed here. (Indeed, in some ways, Linnebo's remarks anticipate Gamester's.)

appear in the computational story...", then relations in which the rats stand to their environments become part of the computational story, and it is no longer so obvious that we can't explain rats' navigational abilities in computational terms.[61] The question thus becomes, once again: What relations between a given rat's map and the environment in which it lives might take over, in a disquotational explanation, the role that representational content plays in Tolman's account?[62]

Consistently throughout his writings on this topic, what Field has offered as an alternative to representational content are what he calls *indication relations* (e.g. Field, 1994, pp. 254–5). I have been unable to find in Field a detailed explanation of what indication relations are supposed to be. For our purposes here, however, what we will need to know is just this: What a particular belief-state indicates is a matter of how the world tends to be when the subject is in that state, and so is uncontroversially reducible to physicalistically acceptable materials.[63] In many cases, of course, belief-states will be reliable indicators of their truth-conditions. That is just another way of saying that many of our beliefs are true. But, as Field emphasizes, there will also be cases in which what a belief-state indicates has little to do with its truth-condition. Someone's political beliefs might indicate only what has recently been said by their favorite political commentator (Field, 1986, p. 89); my beliefs about what is happening in Bosnia might correlate only with what has recently been written in the *New York Times*. That, says Field (1994, p. 255), makes any proposed reduction of truth-conditions to indication relations "at best a gleam in the eye of some theorists".

Despite this divergence, Field (2001b, p. 154) argues, we can still use indication relations to explain behavioral success. Consider someone who is trying to land a plane. An intentional explanation would posit beliefs about airspeed and about what actions it is appropriate to take under various conditions. But we can instead simply posit that there are mental states C_{low}, C_{good}, and C_{high} with the following properties: When Goldilocks is in C_{low}, she increases speed; when she is in C_{good},

---

[61] That is: Field's view is not vulnerable to the criticisms that Burge (1979; 1986a) makes of methodological solipsism (Fodor, 1980), because Field's view is not individualistic.

[62] The disquotationalist thus now seems to be in the very predicament in which verificationists and semantic anti-realists like Dummett have long been stuck: They need to articulate some notion of content as an alternative to the truth-conditional notion that they reject. That is the cost of the failure of the Classical Disquotational Strategy.

[63] Indication relations thus seem to be of a piece with what other philosophers, such as Dretske (1981), have meant by *information*.

she maintains speed; when she is in C_{high}, she decreases speed. And Goldilocks tends to be in C_{low} when the speed is too low; C_{high}, when it is too high; and C_{good}, when it is *just* right.

Obviously, none of this is to be taken terribly seriously. Still, it is hard not to be struck by its quasi-behavoristic simplicity. The story sketched is one of stimulus and response. To be sure, Field is as aware as anyone that mental states, in general, do not 'correlate' with external conditions in the simple way that this sketch supposes they do, either on the input side or on the output side. His response, I take it, would be to admit the over-simplification and to insist that an appropriately similar story can be told by a functionalist: the computational story mentioned a few paragraphs ago. Still, one might wonder what gleam is in whose eye.[64]

A more serious problem is that there are well-known reasons to believe that any notion of content suitable for the purposes of psychological explanation must be compositional: systematicity, productivity, and all that (see e.g. Fodor and Lepore, 2002). But it is at best unlikely that indication relations are compositional: Only certain of my beliefs about Bosnia reflect what is written in the *Times*. Still, this is a well-known problem for views that characterize content in terms of conceptual role, and Field's view is a kind of wide conceptual role theory. So, however serious a problem one might think compositionality poses for such views, it's one they had anyway.

But it is really quite easy to see that behavior cannot be explained in terms of indication relations. This is an immediate consequence of the fact, mentioned above, that indication relations and truth-conditions can come dramatically apart. To take a real case: In December 2016, Edgar Welch drove from Salisbury, North Carolina, to the Comet Ping Pong pizzeria in Washington D. C.—about 350 miles, or 560 kilometers—where he opened fire with an assault rifle. He did so because he believed that the restaurant was being used by associates of (then recently defeated presidential candidate) Hillary Clinton as a front for a child sex ring. What Welch's beliefs *indicated*, it would seem, had rather more to do with what was being posted on certain conspiracy-obsessed websites than with anything approximating reality.[65] To be sure, a complete account of why Welch did what he did would have to dig deep into 'fake news' and the culture of the alt-right. But that has nothing to do

---

[64] It is, in fact, far from clear that there is a content-free notion of computation to be had. (Rescorla, 2014, 2015, 2017).

[65] See https://tinyurl.com/CometPing for the account of this episode in the *New York Times*.

with why, *psychologically* speaking, Welch did what he did: He drove to Washington D. C., etc., because of *what he believed*, namely, that a child sex ring was operating out of Comet Ping Pong. (That was the truth-condition of Welch's belief.) Similarly, even if what my beliefs about Bosnia 'indicate' is what's been reported in the *Times*, what I *do* as a result of my having those beliefs—travel to Bosnia, avoid Bosnia, send money to Bosnia—will be determined by *what* it is that I've come to believe (by the truth-conditions of my beliefs), not by what those beliefs indicate.

To take one of Field's own examples, imagine that my beliefs about height are

> systematically exaggerate[d], so that my believing a sentence of the form 'It is n feet high' is strongly correlated with the object before me being $f(n)$ feet high, where $f(x)$ starts dropping off rapidly from x after about 6 feet or so. (Field, 1994, p. 255)

To be even more concrete, suppose that, if I see an object that is 6 feet 2 inches tall, I believe that it is 6 feet 4 inches tall. Now suppose that I need to fit that object through a portal that I have just measured at 6 feet 3 inches tall. Will what I do be determined by how tall I believe the object is or by what my belief indicates about how tall it is?

The reason this point has been missed, I suggest, is because of the focus on explanations of behavioral *success*. If a particular piece of behavior is successful because the beliefs that explain it are true, then that more or less implies that indication relations and truth-conditions will coincide in that case, which will make it hard to choose between them. But, as I emphasized earlier (see page 29), what matters here is the role that the truth-conditions of our beliefs play in the explanation of behavior, not just the role that the *truth* of our beliefs plays in the explanation of *successful* behavior. It is not just successful behavior that is explained in terms of representational content but *all* (intentional) behavior. Indication relations cannot substitute for truth-conditions in the general case, for the simple reason that the two diverge. One might well say, then, that it turns out to be the role that *falsity* plays in the explanation of *un*successful behavior that is crucial.[66]

---

[66] Indeed, there seems to be something deeply satisfying about the way in which falsity emerges here as the crux: It is, after all, the possibility of misrepresentation that seems to demand some notion of representational content. Otherwise, we could just make to with correlations (see e.g. Dretske, 1986).

To circle back: All these points apply just as well to rats' cognitive maps. It was a large part of Tolman's point that the role cognitive maps play in the explanation of navigation cannot be reduced to stimulus and response:[67] There is no simple story to be told of the form "If the map has this feature, then the rat will turn right".[68] And even if we were somehow to arrange, in the case of some particular rat, for indication relations to diverge from truth-conditions, then it would still be the content of the map that explained the rat's behavior, not what the map indicated.

I conclude, then, that indication relations cannot play the role that truth-conditions are supposed to play in intentional explanation. The argument just given seems, in fact, to show that *no* relation between mental representations and the external world that diverges from truth-conditions can play that role.[69] That brings us very close to showing that nothing but truth-conditions can play that role, but it will be enough for now to observe that nothing else seems to be on offer.

5   Disquotation and Reduction

Field (1994, pp. 249–51) tells us that attempts to articulate an alternative to truth-conditional approaches to content have tended to flounder because they have accepted that very burden: to articulate a notion of content other than the truth-conditional one and, in some cases, even to re-characterize the notion of truth itself. Disquotationalism is meant to be an heir to such views that allows us to reject the demand for such alternatives. It insists that there are utterly unproblematic notions of truth and of truth-conditions, characterized in terms of disquotation, that everyone must accept. And it conjectures that, on examination, it will turn out that the only theoretical role the notions of truth and truth-conditions are ever really needed to play—whether in philosophy, linguistics, logic, or psychology—is the 'expressive' role for which the disquotational notions are custom-built. What I have been calling the Classical Disquotational Strategy amounts to an attempt to prove that conjecture, one case at a time.

---

[67] Tolman made the proposal we have been discussing at a time when behaviorism dominated American psychology, and his proposal is explicitly presented as an alternative to a behaviorist account. Indeed, it is one of the earliest examples of what would later come to be called 'cognitive psychology'.

[68] Special thanks to Dilip Ninan and Marcus Gianquinto here.

[69] To make fully explicit the connection to the end of section 3.2: E-reference cannot play the role reference plays in intentional explanation if the two come apart.

I have argued here that the Classical Disquotational Strategy fails in at least two cases: that of context-dependent utterances and that of psychological explanation. It is, of course, open to disquotationalists to offer alternatives to truth-involving explanations rather than to attempt to reconstrue them disquotationally. But, if so, then it is hard to see what disquotationalism is contributing to the effort. To give what is perhaps an extreme example: Quine wouldn't have been troubled in the least by the question what role truth-conditions play in psychological explanation, but that's not because he was a disquotationalist; it's because he was a behaviorist.

All of that is purely negative, of course, and I admit that I have done nothing here to address the question how representational content can play a significant role in psychological explanation. But I've started to wonder whether part of what underlies the disagreement between disquotationalists and intentional realists is a different disagreement about what work actually needs to be done here.

In early discussions of deflationary theories of truth (e.g. Horwich, 1990), they were often contrasted with correspondence theories, coherence theories, and the like. But it was quickly pointed out that opponents of deflationism need have no interest in any of those alternatives (see e.g. Davidson, 1990, 1996). There is a dangerous ambiguity in the phrase "theory of truth". Such a theory can be one about the nature of truth, or it can simply be one that is 'about truth' in much the same way that Peano arithmetic is a theory about the natural numbers. Maybe no theory of truth's 'nature' is possible, not because truth is 'insubstantial', but because, as Gottlob Frege (1984, opp. 59–60) thought, truth is too fundamental to be definable in other terms. To defend the semantical viewpoint, then, we do not need a theory of truth's nature. We just need a theory that tells us some true things about truth.

All of that is widely appreciated nowadays. But the corresponding points about content are not. I mentioned earlier that Field's own embrace of disquotationalism seems to have been motivated, at least in part, by the failure of various attempts, throughout the 1980s, to reduce semantical notions to broadly physical ones. Those, of course, were 'theories of content' in the sense in which the correspondence theory is a theory of truth. But one might wonder why the failure of the reductionist project should have surprised anyone. As Fodor (1989, p. 413) famously quips, "...nothing ever seems to reduce to anything...", and yet that does not disqualify unreduced notions from doing serious explanatory work, or so Fodor (1974) famously argues elsewhere. No doubt, there's

a great deal more to be said about this, but that is very much my point. Absent further argument, the irreducibility of semantic notions to physical ones merits no more than a shrug unless you assume some strong form of reductionism (cf. Chomsky, 2000).

I suggest, then, in closing, that we should at least consider the possibility that it is with content as it is with truth (cf. Burge, 1986b, p. 719). To defend the representational (truth-conditional) viewpoint, we do not need a theory of content's nature, one that reduces content to more basic notions. It is enough to have a well-motivated theory in which representational content has a central role to play. Such a theory would tell us about representational content by making some true claims about it.[70] Cognitive science, or so I have argued, contains at least one example of such a theory, and there are many, many more (see Burge, 1986a), including in the branch of cognitive science known as linguistics. Of course, any such theory might be false; maybe all of the extant ones are. But the sorts of arguments presented here plausibly apply to any view that makes serious use of a notion of representational content. To be sure, the idea that the mind is an information-processing device may eventually prove ill-founded. But that sort of question will not be settled by conceptual analysis of the word "true".[71]

References

AmourGarb, B. and Beall, J., eds. (2005). *Deflationary Truth*. Chicago, Open Court.

Bach, K. (2015). 'The predicate view of proper names', *Philosophy Compass* 10: 772–784.

Burge, T. (1979). 'Individualism and the mental', *Midwest Studies in Philosophy* 4: 73–122. Reprinted in Burge, 2007, Ch. 5.

---

[70] Nor, I would argue, is reduction required if we're to understand 'the place of content in a natural world'. It might be good enough to have a theory of mind and language in which notions like truth and content play an important role.

[71] Talks based upon this material were presented at the University of St Andrews, McGill University, the Institute of Philosophy at the University of London, the University of Toronto, and Washington University in St Louis. Thanks to everyone who attended for their questions and comments, which did much to improve the paper. Special thanks are due to Imogen Dickie, Marcus Gianquinto, Dilip Ninan, Crispin Wright, and Byeong-uk Yi.

—— (1986a). 'Individualism and psychology', *Philosophical Review* 95: 3–45. Reprinted in Burge, 2007, Ch. 9.

—— (1986b). 'Intellectual norms and foundations of mind', *Journal of Philosophy* 83: 697–720. Reprinted in (Burge, 2007, Ch. 10).

—— (2007). *Foundations of Mind: Philosophical Essays, Volume 2*. Oxford, Oxford University Press.

Carnap, R. (1947). *Meaning and Necessity: A Study in Semantics and Modal Logic*. Chicago IL, Univeristy of Chicago Press.

Chomsky, N. (1969). 'Quine's empirical assumptions', in D. Davidson and J. Hintikka (eds.), *Words and Objections: Essays on the Philosophy of W. V. Quine*. Dordrecht, Reidel, 53–68.

—— (1986). *Knowledge of Language: Its Origin, Nature, and Use*. Westport CT, Praeger.

—— (2000). 'Naturalism and dualism in the study of language and mind', in *New Horizons in the Study of Language and Mind*. New York, Cambridge University Press, 75–105.

Church, A. (1950). 'On Carnap's analysis of statements of assertion and belief', *Analysis* 10: 97–99.

Cohen, J. and McLaughlin, B., eds. (2007). *Contemporary Debates in Philosophy of Mind*. Oxford, Blackwells.

David, M. (2005). 'Some T-biconditionals', in AmourGarb and Beall 2005, 382–419.

Davidson, D. (1967). 'Truth and meaning', *Synthese* 17: 304–23. Reprinted in Davidson, 1984, Ch. 2.

—— (1968). 'On saying that', *Synthese* 19: 130–46. Reprinted in Davidson, 1984, Ch. 7.

—— (1973). 'Radical interpretation', *Dialectica* 27: 313–28. Reprinted in Davidson, 1984, Ch. 9.

—— (1984). *Inquiries Into Truth and Interpretation*. Oxford, Clarendon Press.

—— (1990). 'The structure and content of truth', *Journal of Philosophy* 87: 279–328.

—— (1996). 'The folly of trying to define truth', *Journal of Philosophy* 93: 263–278.

Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge MA, MIT Press.

—— (1986). 'Misrepresentation', in R. Bogdan (ed.), *Belief: Form, Content, and Function*. Oxford University Press, 17–36.

Dummett, M. (1991). *The Logical Basis of Metaphysics*. Cambridge MA, Harvard University Press.

Field, H. (1972). 'Tarski's theory of truth', *Journal of Philosophy* 69: 347–375. Reprinted in Field, 2001d, Ch. 1.

—— (1978). 'Mental representation', *Erkenntnis* 13: 9–61. Reprinted in Field, 2001d, Ch. 2.

—— (1986). 'The deflationary conception of truth', in G. MacDonald and C. Wright (eds.), *Fact, Science, and Morality*. Oxford, Blackwell, 55–117.

—— (1994). 'Deflationist views of meaning and content', *Mind* 103: 249–285. Reprinted in Field, 2001d, Ch. 4.

—— (2001a). 'Attributions of meaning and content', in Field 2001d, 157–174.

—— (2001b). 'Postcript to "Deflationist views of meaning and content"', in Field 2001d, 141–56.

—— (2001c). 'Postscript to "Mental representation"', in Field 2001d, 83–103.

—— (2001d). *Truth and the Absence of Fact*. Oxford, Clarendon Press.

Fodor, J. (1974). 'Special sciences (or: The disunity of science as a working hypothesis)', *Synthese* 28: 97–115. Reprinted in Fodor, 1981, Ch. 5.

—— (1975). *The Language of Thought*. Cambridge MA, Harvard University Press.

—— (1980). 'Methodological solipsism considered as a research strategy in cognitive psychology', *Behavioral and Brain Sciences* 3: 63–73. Reprinted in Fodor, 1981, Ch. 9.

—— (1981). *Representations*. Cambridge MA, MIT Press.

—— (1987). *Psychosemantics*. Cambridge MA, MIT Press.

—— (1989). 'Review essay: *Remnants of Meaning* by Stephen Schiffer', *Philosophy and Phenomenological Research* 50: 409–23. Reprinted in Fodor, 1990, Ch. 7.

—— (1990). *A Theory of Content, and Other Essays*. Cambridge MA, MIT Press.

—— (2007). 'The revenge of the given', in Cohen and McLaughlin 2007, 105–16.

Fodor, J. and Lepore, E. (2002). *The Compositionality Papers*. Oxford, Clarendon Press.

Frege, G. (1984). 'Thoughts', tr. by P. Geach and R. H. Stoothoff, in B. McGuiness (ed.), *Collected Papers on Mathematics, Logic, and Philosophy*. Oxford, Basil Blackwell, 351–72. Also in Frege 1997, 325–45.

—— (1997). *The Frege Reader*, Beaney, M., ed. Oxford, Blackwell.

Gallistel, C. R. (1990). *The Organization of Learning*. Cambridge MA, MIT Press.

Gamester, W. (2018). 'Truth: Explanation, success, and coincidence', *Philosophical Studies* 175: 1243–65.

Grover, D., Camp, J., and Belnap, N. (1975). 'The prosentential theory of truth', *Philosophical Studies* 27: 73–125.

Gupta, A. (1993). 'A critique of deflationism', *Philosophical Topics* 21: 57–81.

Heck, R. K. (2004). 'Truth and disquotation', *Synthese* 142: 317–52. Originally published under the name "Richard G. Heck, Jr".

—— (2006). 'Reason and language', in C. MacDonald and G. MacDonald (eds.), *McDowell and His Critics*. Oxford, Blackwells, 22–45. Originally published under the name "Richard G. Heck, Jr".

—— (2007). 'Are there different kinds of content?', in Cohen and McLaughlin 2007, 117–38. Originally published under the name "Richard G. Heck, Jr".

—— (2014). 'Semantics and context-dependence: Towards a Strawsonian account', in A. Burgess and B. Sherman (eds.), *Metasemantics: New Essays on the Foundations of Meaning*. Oxford, Oxford University Press, 327–64. Originally published under the name "Richard G. Heck, Jr".

—— (2019). Disquotationalism and the compositional principles.

Higginbotham, J. (1986). 'Linguistic theory and Davidson's program in semantics', in E. Lepore (ed.), *Truth and Interpretation: Perspetives on the Philosophy of Donald Davidson*. Oxford, Basil Blackwell, 29–48.

Horwich, P. (1990). *Truth*. Oxford, Blackwell.

—— (1998). *Meaning*. Oxford, Clarendon Press.

Katz, J. J. and Postal, P. M. (1964). *An Integrated Theory of Linguistic Description*. Cambridge MA, MIT Press.

King, J. C. (2014). 'Speaker intentions in context', *Noûs* 48: 219–37.

Kripke, S. (1977). 'Speaker's reference and semantic reference', *Midwest Studies in Philosophy* 2: 255–76.

Langford, C. H. (1937). 'Review of E.W. Beth, "The significs of pasigraphic systems. A contribution to the psychology of mathematical thought processes"', *Journal of Symbolic Logic* 2: 53–4.

Leeds, S. (1995). 'Truth, correspondence, and success', *Philosophical Studies* 79: 1–36.

Lepore, E. and Loewer, B. (1981). 'Translational semantics', *Synthese* 48: 121–133.

Lewis, D. (1970). 'General semantics', *Synthese* 22: 18–67. Reprinted in Lewis, 1985, Ch. 12.

—— (1985). *Philosophical Papers Volume 1*. Oxford, Oxford University Press.

—— (1991). *Parts of Classes*. Oxford, Basil Blackwell.

Linnebo, Ø. (2006). 'Epistemological challenges to mathematical platonism', *Philosophical Studies* 129: 575–4.

McGee, V. (2005). 'Afterword: Trying (with limited success) to demarcate the disquotational–correspondence distinction', in AmourGarb and Beall 2005, 143–52.

Moore, G. E. (1953). *Some Main Problems of Philosophy*. London, Allen and Unwin.

Peacocke, C. (1986). 'Explanation in computational psychology: Language, perception, and level 1.5', *Mind and Language* 1: 101–123.

—— (1989). 'When is a grammar psychologically real?', in A. George (ed.), *Reflections on Chomsky*. Oxford, Basil Blackwell, 111–131.

Putnam, H. (1975). 'What is "realism"?', *Proceedings of the Aristotelian Society* 76: 177–94.

—— (1978). *Meaning and the Moral Sciences*. London, Routledge and Kegan Paul.

Quine, W. V. O. (1956). 'Quantifiers and propositional attitudes', *The Journal of Philosophy* 53: 177–87. Reprinted in Quine, 1976, Ch. 17.

—— (1960). *Word and Object*. Cambridge MA, MIT Press.

—— (1968). 'Ontological relativity', *Journal of Philosophy* 65: 185–212. Reprinted in Quine, 1969, Ch. 2.

—— (1969). *Ontological Relativity and Other Essays*. New York, Columbia Unversity Press.

—— (1970a). 'Methodological reflections on current linguistic theory', *Synthese* 21: 386–98.

—— (1970b). *Philosophy of Logic*. Englewood Cliffs NJ, Prentice Hall.

—— (1976). *The Ways of Paradox and Other Essays*. Harvard University Press.

Reimer, M. (1991). 'Demonstratives, demonstrations, and demonstrata', *Philosophical Studies* 63: 187–202.

Rescorla, M. (2009). 'Predication and cartographic representation', *Synthese* 169: 175–200.

—— (2013). 'Rationality as a constitutive ideal', in E. Lepore and K. Ludwig (eds.), *A Companion to Donald Davidson*. Oxford, Wiley-Blackwell, 472–88.

—— (2014). 'The causal relevance of content to computation', *Philosophy and Phenomenological Research* 88: 173–208.

—— (2015). 'The representational foundations of computation', *Philosophia Mathematica* 23(3): 338–366.

—— (2017). 'Levels of computational explanation', in T. M. Powers (ed.), *Philosophy and Computing: Essays in Epistemology, Philosophy of Mind, Logic, and Ethics*. Cham, Switzerland, Springer, 5–29.

—— (2018). 'Maps in the head?', in K. Andrews and J. Beck (eds.), *The Routledge Handbook of Philosophy of Animal Minds*. New York, Routledge, 34–45.

Shapiro, S. (1998). 'Proof and truth: Through thick and thin', *Journal of Philosophy* 95: 493–521.

—— (2003). 'The guru, the logician, and the deflationist: Truth and logical consequence', *Noûs* 37: 113–32.

—— (2005). 'Gurus, logical consequence, and truth-bearers: What is it that is true?', in AmourGarb and Beall 2005, 153–70.

Simchen, O. (2012). *Necessary Intentionality: A Study in the Metaphysics of Aboutness*. Oxford University Press.

Soames, S. (1999). *Understanding Truth*. Oxford, Clarendon Press.

Stich, S. (1983). *From Folk Psychology to Cognitive Science: The Case Against Belief*. Cambridge MA, MIT Press.

Strawson, P. F. (1950). 'Truth', *Proceedings of the Aristotelian Society* supl vol 24: 129–56. Reprinted in Strawson, 1971, Ch. 10.

—— (1971). *Logico-Linguistic Papers*. London, Methuen.

Tolman, E. C. (1948). 'Cognitive maps in rats and men', *Psychological Review* 55: 189–208.