Field on the Concept of Truth

Anil Gupta
José Martínez-Fernández
University of Pittsburgh

Hartry Field's views on the concept of truth have evolved substantially over the years. Two major phases may be distinguished in this evolution. First, in the 1970s, Field articulated *inflationist* theories of truth; more particularly, he proposed a distinctive correspondence account of truth. Second, in the 1990s, Field became an advocate of *deflationist* theories. We shall explain and critically examine these two phases below: the inflationist phase in section I, and the deflationist in section II. Through all the changes, however, some constant themes are visible: a firm commitment to physicalism, a clear interest in the bearing of views about truth on accounts of the mind, and a concern with Quinean skepticism about meaning. Also visible is a constant philosophical attitude: an admirable openness to competing ideas.

I. The Inflationist Phase

In his very first published paper, "Tarski's Theory of Truth" (1972), Field sketched a correspondence account of truth, one that was shaped by Alfred Tarski's definition of truth.[1] Tarski had constructed, for certain formal languages, a definition that is *provably* materially adequate: the concept defined by Tarski is bound to be extensionally equivalent to truth. A remarkable feature of Tarski's definition is that it uses no unclear or problematic notions – at least none beyond those that occur in the language for which truth is defined. Thus, a physicalist, confronted with Tarski's definition for her austere language, is bound to acknowledge that the definition uses only legitimate resources and yet defines a notion extensionally equivalent to truth: Tarski's definition is physicalistically unimpeachable. Nonetheless, Field argued that the physicalist should not accept the definition as providing a satisfactory *reduction* of the notion of truth. Field observed that in the core of Tarski's definition two parts are distinguishable. First,

---

[1] Unless indicated otherwise, Field's papers that we cite are collected in his book *Truth and the Absence of Fact* (Oxford, 2001). Parenthetical references are to the pages of this book.

there are the recursive clauses, which explain the truth and falsity of sentences in terms of the denotations of primitive constituents, that is, in terms of such things as what names refer to and what predicates are true of. Second, there are the base clauses, which enable Tarski to eliminate the notion of denotation from his definition. Field pointed out that the base clauses are essentially lists: amalgamations of such claims as " 'Socrates' denotes Socrates," " 'Plato' denotes Plato," and " 'Man' denotes men." The base clauses, Field argued, do not provide a genuine reduction of the notion of denotation. According to Field, Tarski had only reduced truth to denotation. To obtain a full reduction, a genuine physicalistic reduction will have to be given of denotation. Field maintained that such a reduction can indeed be given. He argued that just as in chemistry the usefulness of the notion of valence provides a substantial reason to expect that it has a physicalistic reduction, similarly the usefulness of the notion of truth (in, e.g., learning about the world from the utterances of others) provides a good reason to expect that truth can be physicalistically reduced (25). Field pointed to Saul Kripke's recent "causal" theory of names as providing some hope for a physicalistic reduction of denotation.

The theory of truth just sketched aims to "explains truth in terms of . . . correspondence relations between words and the extralinguistic world," which Field takes to be the "central" feature of correspondence theories of truth (199). Field's theory is a significant improvement over earlier correspondence theories. The earlier theories typically offered only vacuous accounts of the correspondence relation (e.g., 'Caesar' corresponds to Caesar and 'Caesar crossed the Rubicon' corresponds to the fact that Caesar crossed the Rubicon). But the theory envisioned by Field contains substantial empirical and semantical components: it contains a substantial theory of how primitive terms acquire their denotations and a substantial semantic theory that explains how truth of complex sentences derives from the denotations of their primitive constituents. Field went on to argue (in "Mental Representation," 1978) that this sort of correspondence theory provides the best hope of solving the problem of intentionality, that is, the problem of explaining, materialistically, intentional relations such as belief. Field's proposal was straightforward: The relation "X believes that P" should be viewed as a product of two relations, one a relation between X and a mental representation, say S, and the other a relation of meaning between S and the proposition that P. The first relation presents no special problems for a materialist. Perhaps it can be understood in terms of dispositions to affirm or accept certain sentences, or something

similar. The second relation – "*S* means that *P*" – could be explained, Field argued, in a materialistically adequate way through the correspondence theory sketched above. This theory will yield not only the denotations of names, but also the *properties* represented by the primitive predicates. And truth-conditions can be recovered from this information. The early Field brought a breathtaking simplicity and optimism to a complex and discouraging subject.

Two features are the defining characteristics of *inflationism* about truth. First, an inflationist holds that a substantive account of truth and reference can be given. In particular, questions such as "In virtue of what does 'Aristotle' refer to Aristotle?" can have, according to the inflationist, informative and useful answers. Second, the inflationist holds that truth and truth-conditions "have a central role in a theory of mind (111)." In particular, they have a central role in an account of meaning and content. Field's early work on truth is thoroughly inflationist.

Even in the early work, Field recognized – indeed, stressed – that the inflationist picture needs to be a little more complicated. The concept of denotation cannot be the primary one for an inflationist, for terms do not invariably have a unique denotation. For example, Newton's word 'mass', Field argued, is *referentially indeterminate*. It does not determinately refer to anything in this relativistic world. Still, the term is not completely denotationless, in the way that 'phlogiston' and 'Zeus' are. Field proposed that we should view Newton's word 'mass' as having two *partial denotations*: relativistic mass and proper mass. Referential indeterminacy of one sort or another is ubiquitous in language. A vague term, for instance, 'heap', does not have a determinate extension. We can, by linguistic legislation, remove the vagueness in 'heap' and thereby obtain a definite extension for the term. But different legitimate ways of sharpening 'heap' will yield different extensions. These extensions are the partial denotations of 'heap'. Another complication must be recognized: the partial denotations of different terms may be interlinked. Two extensions U and V may each result from legitimate sharpenings of respectively 'red' and 'orange', but there may be no legitimate joint sharpening of 'red' and 'orange' that yields these extensions (for instance, if U and V overlap). An inflationist must recognize what Field called *correlative indeterminacy*: the indeterminacy in one term may be constrained by the indeterminacy in other terms. This complicates the theory of truth a little but, fortunately, we can keep (a form of) the Tarskian recursive clauses. Let an *admissible interpretation* of a language be an assignment I of entities to terms such that each term is assigned a partial denotation in a manner that respects the constraints

of correlative indeterminacy. Crudely, there is one global sharpening of all terms that yields the assignments made by I.[2] Now truth can be defined via Bas van Fraassen's supervaluation technique: a sentence is true (false) if and only if it is true (false) under all admissible interpretations; where truth (falsity) relative to an admissible interpretation receives the usual Tarskian treatment. In summary, the fundamental notion for an inflationist is not "denotation" but "partial denotation" – or, better, "admissible interpretation."[3] To gain a physicalistic reduction of truth, one needs to work at a physicalistic reduction of "partial denotation."

Field used the idea of partial denotation to make a striking argument against W. V. Quine: even if Quine is right about the existence of radical indeterminacy in meaning and reference, Field argued, we should still accept a correspondence theory of truth (see "Quine and the Correspondence Theory," 1974). All that the existence of radical indeterminacy would show is that terms had a wider range of partial denotations than we had expected. For example, our term 'rabbit' might have the set of undetached rabbit parts as one of its partial denotations. But this does not require any change in the account of truth given in the previous paragraph.

Still, this way of coping with Quinean indeterminacy hardly deserves the label of "correspondence" theory. If the Quinean point is accepted, we have no real explanation of truth in terms of partial denotation. If anything, the explanation goes the other way around: we explain the partial denotations of terms via the sentences that are true (or are held true). Doubt about the correspondence character of Field's theory remains even if we set Quinean indeterminacy aside. For it is doubtful that the notion of partial denotation can be explained independently of the notion of truth. We can grant that the denotation of some terms is prior to the sentences we take to be true. But plainly this is not always so. The partial denotations of *some* terms (e.g., those introduced in certain theories) is settled in part by the (basic) sentences that we accept about them. So, truth does not appear to be reducible to denotation. A related point is this. Even if we accept physicalism, there is little reason to expect a *uniform* reduction of "partial denotation." Our words hook up to things in myriad ways. We have names for all sorts of things, near and distant, small and large, concrete and abstract. And we acquire these names in all sorts of ways.

---

[2]We are deviating here from Field's exposition of the idea, for which see p. 211.

[3]Henceforth we drop the mention of "admissible interpretation."

Sometimes by naming things ourselves, sometimes in formal introductions, sometimes from a passing remark, and so on. The physicalistic reduction of "denotation," even if it is forthcoming, is not likely to be uniform.[4]

The analogy between valence and truth, then, though striking, can be highly misleading. In the case of valence, the answer to the question "In virtue of what does sodium have the valence of +1?" is precisely what explains why sodium exhibits important chemical characteristics. In the case of truth, however, the causal-historical answer to the question "In virtue of what does 'Aristotle' denote Aristotle?" may explain little other than how and why our term 'Aristotle' got linked to a particular ancient philosopher. The answer may contain all sorts of accidental curiosities that are of interest only to a lexicologist. (Perhaps Aristotle's parents chose his name by throwing dice.)[5] The usefulness of the concept of truth, to which Field draws attention, does not provide a good reason to think that truth and denotation have a uniform or explanatory reduction.[6]

Field now repudiates the argument for physicalistic reduction, but for very different reasons based on the deflationary view of truth (29). We shall explain this view below. Its basic idea is that truth is a *logical* notion, not a causal-explanatory one. The function of truth, the deflationists say, is to enable us to express certain things that would remain inexpressible without it: a language without 'true' is comparable to a language without 'not'. Hence, though truth is a useful notion, the deflationists maintain, it is misguided to look for its physicalistic reduction. Deflationists make parallel claims about other semantic concepts, such as reference.

Even in his early writings, Field considered, albeit briefly, the deflationary viewpoint, but he is dismissive. In "Mental Representation" (1978), he says of a central thesis of deflationism, "it

---

[4]In "Mental Representation" Field basically recognizes both the points made here. He writes favorably about a *functional* account of truth and reference, and says that such an account is desirable for "it would allow for the possibility that the reference relation is realized by different physical relations in different organisms (62)." We wish he had added that the realization might vary from term to term within one organism.

[5]The move to inner representation (or "language of thought") does not affect the point. The inner representation that corresponds to "Aristotle" may be an accidental product of the goings on in the brain.

[6]Note that the above considerations do not cast any doubt on inflationism in general.

seems to me rather weak, and I would like to do better (61)." In "The Deflationary Conception of Truth" (1986),[7] Field "flirted" with deflationism – this is Field's self-description (153) – but ends up rejecting it. The flirtation is evident in "Stalnaker on Intentionality" (1986), where he writes, "I am tempted to view such notions as truth and reference in their correspondence senses . . . as having no important value (95, fn. 14)." Field's move to deflationism is complete by the time he publishes "Deflationist Views of Meaning and Content" (1994). Now he not only rejects his earlier argument for a physicalistic reduction, he challenges all the distinctive claims of inflationism.[8] Henceforth, Field's writings on truth fall squarely in the deflationist camp.[9]

II. The Deflationist Phase

Field accepts the basic tenets of the brand of deflationism that finds its inspiration in Quine.[10] Quine held that the principal function of 'true' is to enable us to express certain generalizations. Consider how we might generalize

(1)   If Socrates is a man then Socrates is a man

in first-order logic, the only logic that Quine's philosophy recognizes as legitimate. If we wish to generalize on 'Socrates', we can easily obtain the first-order analog of 'Every man is a man':

---

[7]In G. McDonald and C. Wright (eds.), *Fact, Science, and Morality* (Oxford: Blackwell, 1986), pp. 55-117. This paper is not reprinted in Field's book.

[8]The philosophers whose work moved Field to embrace deflationism are Quine, Dorothy Grover, Joseph Camp, Nuel Belnap, Paul Horwich, and, especially, Stephen Leeds. See Field's book for references.

[9]Truth and other semantic concepts do appear to play a substantive role in Field's later writings. For instance, his "anti-objectivism" about set theory rests on the idea that the set-theoretic discourse does not have a determinate enough interpretation to fix the truth-values of undecidable statements. If taken at face value, the notion of interpretation is playing a substantive role here. Field is committed to reconstructing this and all other uses of semantic notions within the confines of deflationism (see Chapters 8-10).

[10]Horwich's deflationism contains many of these tenets also. The deflationism of Grover, Camp and Belnap, and of Christopher Hill, is significantly different.

For all x, if x is a man then x is a man.

But what if we wish to generalize on the sentence 'Socrates is a man'? In first-order logic, there is no way of doing so directly, for all variables here are nominal and, thus, there is no analog of

For all z, if z then z.

We can express the generalization, however, if 'true' is available. For (1) is equivalent to

If 'Socrates is a man' is true then 'Socrates is a man' is true,

which in effect converts the sentence position into one that first-order variables can occupy. And we can express the desired generalization thus:

For all sentences x, if x is true then x is true.

More interesting examples of this maneuver can be given, but the essential point is this: 'true', according to Quine, enables us to generalize on *sentence* positions while using *nominal* variables. Crucial to this function is the equivalence,

(T)   'p' is true iff p,

which deflationists take to be the defining principle of 'true'. It is this that leads Quine to call truth a device of *disquotation*. A curiosity is worth noting here. The idea that (T) is central to the function of 'true' enables deflationists to answer Field's seemingly unsurmountable objection to Tarski, the objection that the base clauses in Tarski's definition are unacceptable because of their list-like character. The deflationists point out that without these clauses Tarski's definition does not yield equivalences of the form (T). The base clauses need to be just as they are in Tarski, if 'true' is to play its special role.

     Field's deflationism has some distinctive features. In the first place, it has a strong first-

person orientation: the fundamental truth predicate for a person *S* – what Field calls "pure disquotational truth" – is "true-as-*S*-understands-it." Field explains this notion as follows.

> [A] person can meaningfully apply 'true' in the pure disquotational sense only to utterances that he has some understanding of; and for such an utterance u, the claim that u is true (true-as-he-understands-it) is cognitively equivalent (for the person) to u itself (as he understands it).[11]

Field recognizes other notions of truth but holds that they must all be explained in terms of pure

---

[11]Page 105. This is not the only formulation of disquotational truth to be found in Field's book. Indeed, within the space of one printed page (114-5), Field suggests no fewer than three others. One version takes the instances of (T) to hold by conceptual necessity. Another permits the use of schematic variables in sentence positions, both in and out of quotes, and takes (T) itself to be a principle governing truth. The third allows these variables to be bound by quantifiers and works with the universal generalization of (T). This last version hardly counts as a version of deflationism, for it assigns to truth no essential logical role; it makes truth redundant. Anyhow, the three versions are all different from the canonical version, introduced above, that is the focus of Field's essay. The canonical version takes *utterances* to be the objects of truth, but the other three take these to be *sentences*. (In Chapter 5, Field proposes that quotational names be read as referring to *computational types*, but he provides no precise account of these entities.)

Now each of these versions overcomes some of the problems facing the others. Still none is free from obvious but serious problems. For instance, if utterances are the objects of truth, then the problems of ambiguity and indexicals can perhaps be overcome. But these problems are acute if sentences are taken to be the objects of truth. On the other hand, with utterances, the Quinean point about the generalization function of 'true' loses whatever plausibility it had with sentential truth. And, of course, the burden of reconstructing other notions of truth in terms of utterance truth is well-nigh unbearable.

In the Postscript to Chapter 4, Field expresses a preference for the schematic-variable version. And he allows the truth schema to apply to sentences one does not understand. This is a considerable move away from the original pure disquotational truth, which gives primacy to the notion "true-as-I-understand-it." And it raises urgent questions: What precisely is the truth schema? It cannot be (T) for the trivial reasons of ambiguity and indexicals. So what is it? The answer is not obvious because our understanding of sentences can change over time. Further, what is the motivation for the first-person orientation if schemata are allowed to define predicates? And how can an open-ended schema, some of whose instances are meaningless (for me), define a predicate (for me) in the first place? It is interesting that if the truth-concept is available, we can make some sense of the idea of schemata defining predicates. But how to make sense of the idea *without* the truth-concept? See section V of Gupta's "A Critique of Deflationism" (*Philosophical Topics* 21, 1993, 57-81).

A technical point about the schematic-variable view: It seems to us that rule (ii), stated on p. 115, needs to be modified. It yields an undesirable result when applied to a formula such as

' '*p*' is true' is true iff '*p*' is true.

disquotational truth (123). Second, Field places severe restrictions on the resources available to the deflationist. In particular, he disallows the use of "synonymy." The motivation for this comes in part from Quinean worries about translation.[12] Third, Field links his deflationism about truth to a deflationism about content.[13] The principal thesis of the latter is that truth-conditions do not play a central role in an account of content. Content, it maintains, can be explained in terms of such things as computational role and indication relations (108-112). The deflationist challenge to the early Field to show that truth is a causal-explanatory notion now takes this form: Why cannot an account of content make do with computational role and indication relations? Why are truth-conditions needed in a computational theory of mind (72-82)? Field's deflationist has the notion of truth-conditions available – via the schema (T) or something similar – but assigns no important role to it in his theory of mind.

Let us note a curiosity. The early Field downplayed the significance of the base clauses in Tarski's definition of truth. The later Field, with his emphasis on the schema (T) and its analogs, downplays the significance of the recursion clauses in Tarski's definition (123-6, 141-3).

It seems to us that deflationism about truth is quite independent of deflationism about content. Deflationists about content need not accept the characteristic claims of the deflationists about truth: claims about the centrality of (T) to the content of 'true' or about the priority of pure disquotational truth. Deflationists about content can – and should – hold the entire range of computational roles and indication relations of sentences containing 'true' to be relevant to the content of 'true'. They should treat 'true' as they would any other word. They should refuse to see certain uses of it (e.g., in truth-attributions to utterances one understands) as particularly central to its content. To be sure, deflationism about truth promises them an easy argument

---

[12]Even with the notion of synonymy at hand, it is difficult to recover a full notion of truth. Field's way of dealing with untranslatable utterances (148-51) rests on an erroneous model of what it is to pick up a name (e.g., 'George') – or a sentence – from another (e.g., Mary). Field says that in picking up the name we defer to Mary in that we base our beliefs on her beliefs, or what we take to be her beliefs. But this is not right. We defer only on the reference of 'George', not on beliefs. We may have a low opinion of Mary's cognitive abilities and can still pick up a new name from her. The requisite idea of deference cannot be explained, it seems to us, without appealing to reference and truth.

[13]Strictly, we should speak of content and *meaning*. We shall for the sake of brevity equate the two, for nothing in our argument hinges on the distinction.

against their opponents. But the promise is illusory. There is nothing easy about sustaining deflationism about truth.

Equally, deflationists about truth, even those who are physicalistically inclined, need not accept the characteristic claims of the deflationists about content. They need not accept the idea that content is to be explained in terms of computational role and indication relations. Instead, they can, for example, follow Quine and reject altogether the notion of content. More interestingly, the deflationists about truth can give a truth-conditional account of content. These deflationists have available the notion of truth-conditions, as Field himself notes (111, fn. 9). And nothing precludes them from using it to construct an account of content. The point is perhaps easiest to see with names. Nothing prevents a deflationist about truth (and reference) from holding that the content of 'Socrates' is the man Socrates, as opposed to, say, an individual-concept or a conceptual role. Generalizing this idea, the deflationist can hold that the content of a name is its referent. This does not conflict in any way with the disquotational account of reference.[14] Similarly, there is no conflict between a truth-conditional account of meaning and what Field's deflationist says about pure disquotational truth.[15] The point holds also for such other deflationary ideas as that the schema (T) is central to the meaning of 'true'.[16] Let us note that

---

[14] It might even be said that the disquotational account is needed to express the desired generalization. But we think this is going too far – see Gupta, *op. cit.*

[15] Field's deflationist holds that the claim that an utterance u is true is cognitively equivalent, for a subject that understands u, to u itself. Two sentences are cognitively equivalent for a subject X, according to Field, when X's inferential procedures permit a "fairly direct" intersubstitution of the occurrences of one sentence for those of the other (with the usual restrictions about quotational contexts, etc.). Field requires that these inferences should be "more or less" indefeasible on empirical and conceptual grounds (106, fn. 2). Now, if "cognitive equivalence" is understood in this way, there is nothing in the claim of cognitive equivalence that conflicts with a truth-conditional account of content. A believer in truth-conditional content can happily grant – even insist – that *P* and ' '*P*' is true' are cognitively equivalent for a subject who understands *P*.

[16] There is an argument due to Michael Dummett that attempts to establish a tension between the deflationism about truth and the truth-conditional accounts of meaning. This arguments goes roughly as follows. Suppose one can explain the meaning of a sentence *P* to someone X who does not understand it by giving the truth-conditions of *P*. Then, X must have a prior understanding of truth. But, according to deflationism, this is not possible: one's understanding of truth requires a prior understanding of *P*. This argument works with an

truth-conditions offer some advantages to a theorist of content. They are free from irrelevant factors that threaten to contaminate the rival deflationist account. For example, indication relations are a product of haphazard circumstances and can vary radically with collateral belief. And computational role makes reference to specific signs, which are plainly irrelevant to content.

Accounts of content are quite independent, then, of deflationism about 'true'. They should be evaluated entirely on the basis of the theory that invokes them. Field struggles, early and late, with the deflationists' challenge to show that truth-conditions are *needed* in a computational theory of the mind. But this challenge seems to us misconceived. It imposes on the believer in truth-conditions a particular theory that he has reasons to set aside or even reject. The structure of the truth-conditional theory is quite different from that of a computational theory. In the latter theory, the notion of content seems to play no essential role; whereas in the former, it is crucial. As things stand, it is an open question whether a computational theory is even feasible. Of course, it is also an open question whether a good theory of mind will invoke truth-conditions. The point we wish to insist on is that this question is independent of issues such as whether 'true' is a logical term. Even if 'true' is a logical term, nothing follows about the causal-explanatory status of *statements* such as 'True beliefs promote success' or 'None of Fred's beliefs about beer is true'. Nothing follows about the nature of the objects of propositional attitudes, nor about the shape of an adequate psychological or semantic theory.[17] If we let our commonsense ways of

---

implausibly strong deflationism. And, even here, it establishes a tension only with a rather crude account of *understanding* meaning, not of meaning itself. See Gupta, *op. cit.*, for further discussion and for references.

[17]Reminder: causal-explanatory statements can contain logical terms! We acknowledge this much: if 'true' is a logical term, then its physicalistic acceptability does not rest on the success of the reductionist program Field outlined in his early papers. On the other hand, the idea that 'true' is a logical term does not by itself undermine that program either. For the crucial notions in that program are partial denotation and supervaluational truth, not denotation and truth *simpliciter*.

The early Field linked the fate of truth-conditional accounts of content to the viability of his reductionist program. The later Field has not explicitly repudiated the link, but it should be repudiated. Psychology and semantics may find a truth-conditional account of content useful for stating laws, even though the notion of content is not physicalistically reducible in the manner of "valence" – that is, even though there is no uniform physicalistic reduction of content, nor an explanatory one. Psychology may have a more complex conceptual relationship to physics than does chemistry.

understanding each other be our guide, then it is natural to suppose that truth-conditions will play at least *some* role in content. Our commonsense ways have, of course, their limits. But if we venture beyond them, then we are on a sea of speculation. We can affirm this much, however. Even if we are led, as Quine was, to reject altogether meaning and reference – even if we are led to view ourselves as mouthing sounds without content or any "aboutness" – we can regain a truth-like concept in the Tarskian way. Truth has a certain peculiar robustness. Even if the conceptual system that gives it life is destroyed, its ghost lingers on.[18]