What is Truth?

Edited by
Richard Schantz
The Metaphysics of Deflationary Truth

MICHAEL DEVITT

1. Introduction

What exactly is deflationary truth and how does it differ from correspondence truth? This question has been much discussed and yet, I shall argue, the answer remains unclear. That unclarity arises from insufficient attention to the distinction between the metaphysics of truth and the linguistics of the truth term; hence from insufficient attention to what deflationary theories say, or should say, about the metaphysical issue. In arguing this, I shall emphasize that deflationism is similar to a sort of "nonfactualism".

2. Four Problems

There are four related problems in locating the difference between the deflationary theory and the correspondence theory. The first is that the two theories have opposite foci. Whereas the focus of the correspondence theory is on the nature and role of truth, the focus of the deflationary theory is on the nature and role of the truth term; for example, of "true". The former focus is metaphysical, the latter, linguistic. So, an awful lot of what deflationists say does not bear directly on what the correspondence theorists say, and vice versa.

A simple explanation of this difference in focus – too simple as we shall soon see – is as follows. Deflationism is really a sort of eliminativism, or anti-realism, about truth: it deflates truth itself. We might say, very roughly, that according to deflationism, there is no reality to truth. Since there is no reality to truth there is nothing positive to be said about the nature of truth. However, unlike some early eliminativists, deflationists have no objection to the use of the truth term. Indeed, they are enthusiastic about the term and have a great deal to say about its linguistic role and semantics.¹ In contrast, correspondence theorists are realists about truth and therefore struggle to explain its nature. But, for them, the truth term is just another one-place relational predicate – like, say, 'warranted' or 'factual' – with the standard sort of semantics of such predicates. This semantics is likely to start from the assumption that the term denotes the property truth or applies to all true things. This is so unexciting as to be hardly worth saying and the theorists are not usually inclined to say anything more exciting.

In sum, the deflationist has little to say about the metaphysics of truth but much to say about the linguistic role of 'true', whereas the correspondence theorist has a lot to say about the metaphysics of truth but little to say about the linguistics of 'true'.

I am here describing a real difference in focus between the two theories. The second problem in distinguishing the theories is that this difference is often not apparent. Discussions of deflationism tend to blur the distinction between the linguistic and the metaphysical. In particular, remarks that should be about the truth term are often presented as being about truth: there is use/mention slippage, even confusion. So it can often seem that discussions are talking about truth when they are not really.

The third problem is that when discussions of deflationism do address the metaphysical issue, rather than merely appealing to when addressing the linguistic issue, what is said is often unsatisfactory. And this is not surprising because it turns out to be rather hard to capture the deflationary metaphysics of truth. That is the fourth problem. A sign of this problem is that my characterization of the metaphysics of deflationism in describing the first problem really is very rough.

To appreciate the third and fourth problems it helps to notice that deflationism is similar to "nonfactualism" in ways to be explained (section 4).

Despite these four problems, the difference between deflationary and correspondence theories on the linguistic issue of the truth term is relatively clear. Not so, the difference on the metaphysical issue of truth. As a result of the four problems there is a good deal of uncertainty, if not confusion, over the difference between deflationary and correspondence views of the nature of truth. This is serious because this metaphysical difference is deeper than the linguistic one; it is explanatorily prior.

Section 3 will be concerned with the linguistics of the truth term. Sections 4 to 6 will be concerned with the deflationary metaphysics of truth and the third and fourth problems. A standard characterization of the deflationary metaphysics will be criticized and a better one attempted. These accounts of the linguistic and metaphysical issues are necessary background for appreciating evidence in section 7 of the very tricky second problem: a use/mention slippage that obscures the real metaphysics of deflationism. The uncertainties and confusions arising from these four problems will emerge as we go along but will be particularly prominent in section 7.

In talking of uncertainty and confusion I speak from bitter experience, for I am only too well aware that my own writings on truth have provided some examples.²

3. The Truth Term

The deflationists have some very interesting things to say about the truth term. They have persuasively demonstrated that the term has an extremely useful "logical" or "ex-

¹ The "semantics" of a term concerns its meaning. So also does its "linguistics" but the latter may also concern other aspects of the term's nature and role.

² For example, 1991; 1997: 34.
pressive" role. Thus, suppose that Jack says, "We all lie about our sex lives," and Jill replies, "That is true." Intuitively, the role of the truth term here is to enable Jill to "say the same thing" as Jack without repeating his words (and whilst admitting his priority). Attention to such examples encourages the simplest deflationary theory, the "redundancy" theory, for they make it seem as if we could dispense with the truth term altogether. However, other examples show that the term is very useful. It enables us to assert briefly something that may otherwise be tedious, if not impossible, to assert.

Suppose that Imogen wishes to express general but qualified agreement with a certain article. She can say simply, "Most of what that article says is true." Consider what would be required to say this without using "true." Her claim entails that at least half the claims in the article are true, but is not specific about which half. So her claim is equivalent to a long disjunction of conjuncts, each conjunct consisting of a different set of more than half the claims in the article. If she could remember all the claims, she could, in time, manage to express this disjunction. If not, she needs the truth term. So does a person who has forgotten Goldbach’s Conjecture but nevertheless wants to express agreement with it. Can one say, "Goldbach’s Conjecture is true." A person who has lost track of all the utterances of the Great Helmsman can nevertheless express her commitment, ‘Every time Chairman Mao said was true.’ Without the truth term, she faces the impossible task of asserting an infinite conjunction. So also does a logician in asserting each instance of a schema that has an infinite number of instances.3

The truth term can play its logical role because it yields equivalences like the classic one between "Snow is white" is true and "Snow is white." When the term is attached to the quotation name of a statement it yields a statement that is equivalent to that statement: it undoes the effect of quotation marks. (Attention to this led to the name, ‘the disquotation theory of truth’.) Indeed, when the truth term is attached to any device for referring to a statement it yields a similar equivalence; it is a "denominalizing" device. Thus Jill’s remark, "That is true," is equivalent to Jack’s, "We all lie about our sex lives." If I were to say, "Jack’s remark is true," my remark would be equivalent to Jack’s. The person who said, "Goldbach’s Conjecture is true," said something equivalent to the Conjecture. In general, the deflationary view supports "the equivalence thesis": all appropriate instances of the “equivalence schema”

\[ s \text{ is true iff } p \]

hold, where an appropriate instance substitutes for ‘p’ s translation of the statement referred to by what is substituted for ‘s’.4

What is the “meaning” of the truth term? The deflationists have offered a variety of answers. Thus, Paul Horwich (1990) proposes a "minimalist" theory according to which "true" is an unusual "logical" predicate implicitly defined by its use in the appropriate instances of the equivalence schema. Dorothy Grover (1992) urges a "prosentential" theory according to which "true" is not a predicate at all. Rather, it is a syntagmatic part of an anaphoric "prosentence," where pronouns are to sentences as pronouns are to nouns. This ingenious theory has the unhappy consequence that ‘that’ in ‘that is true’ does not refer to some statement, as one would naturally suppose (and as I supposed in introducing the equivalence thesis). This led Robert Brandom to propose a variation on the prosentential theory that avoided this consequence: the truth term should be treated as a prosentence forming operator (1988: 88-90).

So much for deflationary views of the meaning and role of the truth term. What are deflationary theorists to make of this? The important thing to notice is that they can, and should, go along with most of it. Certainly, they cannot go along with a deflationary theory of meaning of the truth term, whether of the Horwich, Grover, Brandom, or any other variety. They think that the term has the standard semantics of a one-place relational predicate, very likely explained in terms of reference to the truth property or to true statements, as I noted. Still they can and should accommodate the rest of the deflationary story. In particular, they should accept the equivalence thesis: that is a constraint on any theory of the truth term. And if the correspondence theory meets that constraint it can account for the logical role of ‘true’ that the deflationists have so persuasively demonstrated.5

So although the correspondence theorist disagrees with the deflationist over the meaning of the truth term, he should agree that the term has the logical role explained by the deflationist. There will probably be one other important disagreement. The deflationist will insist that the truth term does not have any role other than the logical one; in particular it does not have the "descriptive" role of a normal predicate. The correspondence theorist is likely to think that the term has a substantial descriptive role in some theory of the world.

These linguistic differences between a deflationary and a correspondence theory over the truth term are striking and obvious. The metaphysical differences between the two theories over the nature of truth are much less so. Yet the metaphysical differences are explanatorily prior because they largely motivate the linguistic ones.

4. The Deflationary Metaphysics of Truth: the Problem

We have seen what the deflationists say about the truth term, but what is their view of truth? Where do they stand on the metaphysical issue? I have said that deflationism is a sort of eliminativism or antirealism, and roughly characterized it as denying that there is any reality to truth (section 2). But the inadequacy of this is apparent when we note that deflationists are as ready to talk about statements being true as correspondence theorists; thus many deflationists will say that Jack’s remark is true, because to say this is just to express the common belief that we all lie about our sex lives, and they will all

3 The expressive power that we get from the truth term could also be obtained by introducing sentence variables into our language, see Horwich 1990, pp. 4-5, for an influential argument.

4 This is rough; for example, we need to guard against the semantic paradoxes and allow for individuals.

The formulation talks of translation because an appropriate instance might refer to a statement that is not in the language of the instance; e.g., 'Schnee ist Weiss' is true iff snow is white.

5 As Mark Lance in effect, points out (1997: 184-5).
agree that ‘Snow is white’ is true, because to say this is just to say that snow is white. So what does their antirealism consist in? The focus of deflationist literature is not on answering this question and the little literature says is often unsatisfactory; that was our third problem in section 2. The question turns out to be rather hard to answer; that was our fourth problem.

To appreciate these two problems it helps to realize that analogous problems arise elsewhere. For, deflationism about truth is similar to the ‘nonfactualism’ exemplified by ‘noncognitivism’ about morals, ‘projectivism’ about causality, positivistic instrumentalism about science, and Simon Blackburn’s "quasi-realism" (1984, 1993a, 1993b). Characterizations of the metaphysics of nonfactualism tend to be unsatisfactory and it is difficult to give a satisfactory one. I have discussed these problems for nonfactualism elsewhere and will draw on those discussions in what follows.

Deflationism has two defining features of this kind of nonfactualism. The first is that the linguistic level is very explicit in the literature. Nonfactualism in an area has a revisionist view of the language in that area: the language is not "descriptive" as we would naturally take it to be. This view is expressed in a variety of ways, some rather unsatisfactory, but the key idea is clear: terms that appear to be predicates in the area do not have the standard semantics of a normal predicate; perhaps they are not predicates at all. Because these terms are in this way "non-descriptive" they are not like a normal predicate in pursuing to "describe reality"; they have some other role. Thus, the most famous nonfactualism, noncognitivism about morals, has a revisionist view of the semantics of ‘good’ as a result of which indicative sentences containing it are not assertions or statements. Rather, those sentences express attitudes or emotions, or prescribe norms or rules. The deflationist view of the truth term, discussed in the last section, is a similar sort of revisionism. The truth term does not have the standard semantics of a normal predicate. And its role is not to describe sentences; its only role is logical or expressive.

What is meant by "the standard semantics"? Typically a philosopher's standard semantics will be truth-referential but it need not be: it might be verificationist, for example. And the standard semantics of a deflationist cannot be truth-referential because, for her, truth is not explanatory (section 5). So the non-truth-referential meaning that the attributer to a normal predicate – for example, a certain sort of use condition – she does not attribute to the truth term.

Despite the linguistic similarities between deflationism and nonfactualism, there are important differences which should make one reluctant to treat deflationists as a specie of nonfactualism. First, the deflationist's "expressive" role for ‘true’ is nothing like the noncognitivist's "expressive" role for ‘good’; the former is logical, the latter emotive. Second, the noncognitivist holds that because ‘good’ is nondescriptive, sentences of the form ‘x is good’ are not factual. In contrast, the deflationist does not hold that because ‘true’ is nondescriptive, sentences of the form ‘S is true’ are not factual. For her, whether these sentences are factual depends on whether S is factual. So, if S is ‘x is good’, it is factual for a deflationist who is a cognitivist but not factual for one who is a noncognitivist.

The second defining feature of nonfactualism is at the metaphysical level and is often more implicit than explicit in the literature. Nonfactualism in an area is antirealist about that area. Thus noncognitivists are antirealist about goodness. Deflationists are similarly antirealist about truth.

Consider the problem of characterizing the antirealism of nonfactualism. The most straightforward way of characterizing antirealism in general, using the ordinary language for denying ontic commitment, obviously does not capture the metaphysics of nonfactualism. Thus the noncognitivist does not claim that there are no good people, right actions, and so on. She is as ready as the realist to say, "This person is good" and "That action is right," for she thinks that these utterances express appropriate emotions or prescriptions. We have already made the analogous point about deflationism: the deflationist does not claim that there are no true statements. The nonfactualist and the deflationist talk like a realist but give that talk a revisionist interpretation. This is what poses the problem of distinguishing these doctrines from realism at the metaphysical level; cf. our fourth problem.

But perhaps the problem is illusory. Maybe my confident claim that nonfactualism and deflationism are antirealist is mistaken. Perhaps the focus of these doctrines is so linguistic because they really have no commitment one way or the other on the metaphysical issue. If this were so our enterprise of attempting to characterize their antirealism would be misguided. There are two reasons why this dismantling of the problem must be rejected. The first is that the doctrines are presented in opposition to realist views; thus, deflationists oppose correspondence truth. And, despite the linguistic focus, the doctrines are accompanied by claims that are clearly intended to be antirealist even if, as we shall see, the claims are often not adequate to the intention. The second reason is that an antirealist metaphysics is needed to motivate the revisionist view of language urged by these doctrines. If there were not something problematic or defective about the area of reality that ‘true’ or ‘good’ appear to concern why suppose that they do not have the standard semantics of a descriptive predicate? Of course, the semantic revisionism is typically supported by some purely linguistic considerations: evidence of a non-descriptive role for the language in question. Thus, deflationists are motivated by the logical role of the truth term and noncognitivists by the action-guiding role of moral language. But what is to stop language covered by the standard se-
antirealist one that we are attempting to characterize. Finally, consider the unselective
realist who thinks that there is a property for each predicate. A nonfactualist might
accept that ‘good’ is a predicate, as indeed Blackburn does (1993a: 206), and a deflationist
might accept that ‘true’ is, as indeed Harwich does. If such a person is an unselective
realist she will think that there is a property of goodness or truth, thus disagreeing
with the popular characterization. And even if the nonfactualist denies that ‘good’
is a predicate, and the deflationist that ‘true’ is, hence that there are properties of
goodness and truth, the popular characterization of their antirealism is dubious: it ‘runs
the wrong way’. It finds a defect in reality because of something special about language
where we need to find a defect in reality to motivate the view that the language is spe-
cial.

The general issue of realism about properties is independent of the issues of non-
factualism and deflationism. It should be possible for someone to embrace or reject
the metaphysics of these doctrines whatever her position on this general issue. There
should be a way of stating that metaphysics that is appropriate whatever the truth of
the matter about the reality of properties.

So far, then, we have made no progress characterizing the antirealism of nonfactual-
ism and deflationism. The most straightforward statements of realism, using the
ordinary language of ontic commitment, are not denied by these doctrines because they
are reinterpreted so that they have no such commitment. We have just seen the failure
of a characterization using more “philosophical” talk of properties. In general, the
nonfactualist/deflationist practice of talking like a realist while giving that talk a revisionist
interpretation makes progress hard. We are attempting a characterization of the meta-
physics that must motivate the special semantic treatment that the doctrines give to a
certain area of language. Yet our attempts seem doomed to vitiation by that very
semantic treatment. Nonfactualism and deflationism are supposed to be a sort of antire-
alism and yet it seems impossible to give a metaphysical statement of their antirealism.
Realism issues begin to evaporate. Indeed, Blackburn sometimes comes very close to
claiming that they have evaporated (1993a: 4, 15-34, 55-9; 1993b: 368).

5. The Deflationary Metaphysics of Truth: the Solution

To avoid the evaporation of realism as a metaphysical issue and to characterize
the metaphysics of nonfactualism in an area, we must first find some language in that area
that is not just apparently descriptive but is treated by the nonfactualist as really
descriptive. We must then examine her statements using that language to find ones that
disagree with realist statements about the area.

10 Concerning realism issues in general, I have argued that, from a naturalistic perspective, we should al-
ways “put metaphysics first” by establishing a metaphysical base with near enough no appeal to
semantics and by arguing from that base for a semantics. For we know far more about the world than we do
about meanings (1997, Devitt and Sterelny: 11.4, 12.4).

11 Or that there are any facts in the area. That characterization has similar problems to the one about prop-
erties. For examples of these characterizations, see Ayers 1952: 89; Wright 1988: 29-30; Sayre-McCord
1988: ix-x; Brandon 1988: 90-1; Boghossian 1990a: 157-9, 161-2; Gevers 1992: 14; Blackburn

12 Richard Kirkham notes the problem that the realism issue about properties poses for the popular char-
acterization. His solution is to characterize deflationism at the thesis “that ‘true’ is not a genuine predi-
cate” (1993: 311). One objection to this is that some deflationists—e.g., Harwich—think that
‘true’ is a genuine predicate. A more serious objection is that it is a linguistic characterization and we
need a metaphysical one.
I have argued elsewhere (1996: 165-70; 1997: 313-18) that two sorts of realist claim are the most promising candidates for denial by the nonfactualist. First, the typical realist offers explanations of the nature of the problematic reality in language that the nonfactualist should accept as factual. For, the realist thinks that the problematic reality is constituted by, or supervenes on, a reality that should be unproblematic for the nonfactualist. Even though the nonfactualist claims to be able to accept many sentences that seem to describe the problematic reality, taking them as expressive, prescriptive, or whatever, she does not accept the need for, or possibility of, these substantial "broadly reductive" explanations. Thus, moral realists claim that there are things about a person in virtue of which she is good, that make her good; for example, being kind, considerate, generous, honest, etc. And there are things about an action in virtue of which it is wrong, that make it wrong; for example, leading to unhappiness, being contrary to socially accepted rules, and so on. The noncognitivist must reject all such "in virtue of" claims as totally misconceived.

The deflationist has a similar disagreement with the typical realist about truth. The realist will claim that there is something common and peculiar to true statements: a statement is true in virtue of some sort of correspondence relation to the world; this relation makes it true. A substantial theory is then needed to describe and explain this correspondence, a theory that may include, for example, causal theories of reference. Deflationists should reject any such reductive explanation of truth. Horwich does so in denying that truth has an "underlying nature" or some "hidden structure waiting our discovery" (1990: 2): "being true is insusceptible to...scientific analysis" (p. 6).

Grover claims that "truth talk...can be explained without appeal to any kind of analysis of the nature of truth" (1992: 3). This is not to say that the deflationist rejects all statements of the form 'p explains that S is true'. The deflationist, like everyone else, accepts the need for, and possibility of, explanations of "worldly facts" such as that snow is white. Explanations that appeal to laws of nature. Suppose that E explains that snow is white. So, given the deflationary theory, E explains that 'snow is white'. But this sort of explanation, varying from truth to truth, is not what the correspondence theorist seeks. He seeks an explanation of what all true statements have in common, an account of "correspondence to the world." That is the sort of explanation that the deflationist must reject.

The second sort of realist claim that the nonfactualist should deny concerns causal role. The typical realist thinks that the problematic reality is the cause or effect of some unproblematic reality. The nonfactualist should not accept these claims about the role of the problematic reality because on her view there is no reality that could play such a role. Thus, the typical moral realist thinks that there are causes and effects of a person being good or bad. He thinks that it is because Hitler and his associates were depraved that we believe that they were depraved. And it is because they were depraved that they behaved as they did and that millions of people died in concentration camps. The noncognitivist must reject all such explanations.

Once again, the deflationist has a similar disagreement with the realist about truth. The typical realist will give truth important explanatory roles; for example, to explain the success of science or the success of people in meeting their goals; or to explain meaning, where meaning itself plays a role in the explanation of behavior. A deflationist must reject all such explanations and Branden clearly does reject them all (1988: 91-2). 13

This is not to say that the deflationist cannot use the truth term in explanatory statements. The logical role of the truth term makes an explanation of the form 'p because it is true that q equivalent to one of the form 'p because q'. But the appearance of 'true' in the former sort of explanation does not make truth explanatory of p. Consider an example: 'Clinton was impeached because he is hatred by the religious right' can be rewritten as 'Clinton was impeached because it is true that he is hated by the religious right'. Manifestly, what is explanatory here is hatred not truth. Even where the expressibility provided by the truth predicate is essential to an explanation - because without it the explanation would be infinite - it is not truth that is explanatory.

In sum, the typical realist thinks that there is a reality to truth which, like any other reality, has a nature and causal role; and that this nature and role need explanations. The deflationist reveals her antirealism by rejecting the need for and possibility of such explanations. Although she can join with the realist in accepting ordinary truth claims, she cannot join with him in his explanation of the reality which he takes those claims to describe. The deflationist should have nothing that is positive and substantial to say about truth.

Sadly, this account of the distinction between realism and nonfactualism/deflationism has a flaw, reflected in the frequent use of 'typical'. There are doubtless some philosophers who claim to be moral realists and yet join the noncognitivists in denying the need for an explanation of moral reality and in denying that this reality has any causal role; it is inexplicable and epiphenomenal. One can imagine an analogous claim from someone who sees himself as a realist about truth. Such positions are deeply antirealistic, of course. They are also hard to motivate: Why believe in a truth or goodness that can do nothing and cannot be explained? Still, the positions are possible. And if they have a standard semantics for 'true' and 'good' they surely are realists, for they accept the straightforward statements of realism without interpreting away the ontic commitment of the statements. So, the flow in my account is that it does distinguish this atypical realism from nonfactualism and deflationism at the strictly metaphysical level. I suspect that this realism cannot be so distinguished. If not, we must conclude, disappointingly, that to fully capture the antirealism of nonfactualism/deflationism we have to add a little semantics: what makes these doctrines antirealism is not only their denial of explainable nature and causal role but also their adoption of a nonstandard semantics that removes the commitment from apparently straightforward statements of realism.

The nonfactualist/deflationist and the atypical realist agree that in a certain area there is no reality with an explicable nature and a causal role. Despite this failure, the atypical realist holds that there is a reality in that area: the reality is simply inexplicable and epiphenomenal. In contrast, the failure motivates the nonfactualist/deflationist to

---

revert the reality altogether by revising the semantics for what would otherwise be straightforward statements of realism.

6. The Equivalence Thesis

The difference between deflationism and the correspondence theory should emerge in their responses to a demand for an explanation of the equivalence thesis. Let us take the most famous instance of the equivalence schema as our example: 'Snow is white' is true if snow is white. In virtue of what is this so? The core of the correspondence theorists' answer is a reductive theory of the nature of true statements - that is, common and peculiar to these statements. This will be an account of the relation that true statements stand in to the world. When this theory is applied to 'Snow is white' it shows that this statement is related to the world in such a way that the statement is true if snow is white. So the theory of truth, together with facts about the statement 'Snow is white', explain why 'Snow is white' is true if snow is white.

The deflationist, in contrast, cannot accept any appeal to a theory of the nature of truth in her explanation because she dismisses the possibility of saying anything substantial about that nature. So, what explanation does she offer? Basically, none. She thinks the demand for an explanation here is misguided: that 'Snow is white' is true if snow is white is a "brute fact" needing no explanation. However, she has something further to say to make this provocative claim palatable: a diagnosis of the error of thinking that we need an explanation here. The diagnosis moves up to "the semantic level," considering the way the brute fact is expressed. Although the deflationist denies the need to explain why 'snow is white' is true if snow is white, she accepts the need to explain why people wrongly think that the statement 'snow is white' is true if snow is white expresses something that needs explaining. The error arises from treating 'true' as if it were a normal descriptive relational predicate, thus taking the truth of 'Snow is white' to depend on some relation that statement has to snow being white. Once the non-descriptive meaning of 'true' is appreciated, we see that to say that 'Snow is white' is true is not to relate the statement in some way to the world but simply to say that snow is white. So, of course 'Snow is white' is true if snow is white, just as snow is white if snow is white.15 No more needs to be said (unless "the logical structure of the world" is to be explained). Similarly, it might be claimed that the following are brute facts needing no explanation: that Schnee is Weiss if snow is white; that all bachelors are unmarried; and that Hesperus is Phosphorus. However, someone might think otherwise because she failed to appreciate the relevant semantic facts: that 'Schnee is Weiss' is synonymous with 'snow is white'; that the meaning of 'bachelor' includes the meaning of 'unmarried'; and that 'Hesperus' and 'Phosphorus' rigidly designate the same object.

The contrast between the two theories should not be that the correspondence theory must offer a substantive explanation of why 'Snow is white' is true if snow is white where the deflationary theory offers a trivial one appealing to the meaning of 'true'. The contrast should be that the correspondence theory must offer an explanation where the deflationary theory appeals to the meaning of 'true' to explain why no explanation is necessary.

The position I am attributing to the deflationist on this matter is undoubtedly hard to grasp. The position is developed and modified a little in the next section.

I started this paper by mentioning four related problems in distinguishing deflationism from the correspondence theory. The first problem was a difference in focus: the focus of deflationism is on the linguistics of the truth term, the focus of the correspondence theory on the metaphysics of truth. The second problem was the unsatisfactory nature of attempts to characterize the metaphysics of deflationism and the fourth was the difficulty of such a characterization. I have said a lot about these three but nothing yet about the second problem. We now have the background to discuss it.

7. Use/Mention Sloppiness

The second problem was use/mention sloppiness, even confusion, in the literature: deflationist remarks that should concern the linguistics of the truth term are often misrepresented as being about the metaphysics of truth, thus obscuring the real metaphysics of deflationism. In giving examples of this sloppiness I do not mean to suggest that all of them amount to real confusions in thinking. Some surely are just insignificant careless or convenient rhetoric. Still I want to show, first, how pervasive this sloppiness is. Second, I want to make it plausible that there are some cases of real confusion: that there should be a theory of the truth term is really being taken as a theory of truth, not simply carelessly expressed. This shows, it seems to me, how very difficult it is to handle the use/mention distinction in discussing truth.

In the light of our discussion so far, it is easy to spot the sloppiness. The deflationist is talking about truth itself, and saying something inappropriate, when she denies that truth has a nature or causal role that needs or can have an explanation (section 5). And she is talking about truth itself, but saying something inappropriate, when she denies that truth is a property (section 4). Anything else she says, particularly anything positive, that is represented as being about truth should very likely be about the truth term.

1. The problem starts with the very names of some deflationary theories. The name 'the redundancy theory of truth' implies that truth is redundant but what is really redundant according to the theory is the truth term. Similarly, what is really disqua-

---

14 Note that we are asking in virtue of what 'Snow is white' is true if snow is white, not in virtue of what 'Snow is white' is true if snow is white is true. Deflationism and the correspondence theory would give similarly different responses to the latter question, but the responses would be more complicated.

15 Compare: "All the anaphoric (prosaistential) theory of truth tells us about what it is for 'Snow is white' to be true, is that it is for snow to be white" (Lance 1997: 188).
tional according to 'the disquotational theory of truth' is the truth term not true. What is presentential according to 'the presentential theory of truth' is not true but a linguistic expression including the truth term. The generic name, 'the deflationary theory of truth', does refer to theories that deflate truth not the truth term, and so the name does not confuse use and mention. Still, the name is a bit misleading because only a small part of what deflationary theories actually say concerns truth. What they say mostly concerns the truth term. They deny a descriptive role for the term but emphasize other roles that were largely unnoticed or ignored by correspondence theories.

On balance, deflationary theories inflate the truth term. Consider this historically important but notoriously difficult case of deflationism: Alfred Tarski. A special difficulty is that Tarski does not see himself as a deflationist but rather, it seems, as a correspondence theorist (1956: 153, 404).

On the opening page of "The Concept of Truth in Formalized Languages," Tarski variously describes his enterprise as the definition of truth, of the term 'true sentence,' and of the concept of truth (1956: 152). The last two can be taken to be the same but, prima facie, they are different from the first. Defining truth is a matter of explaining its nature, a metaphysical matter, whereas defining the term and the concept are linguistic matters. We have use/mention sloppiness.

What does Tarski actually do? He defines the meaning of 'true-in-L', where L is any of a certain range of formal languages. Does this have anything to do with explaining the nature of truth? See aside worries arising from the fact that he has defined 'true-in-L' not 'true' and suppose that he had defined 'true'. Could that have shown anything about truth? It depends on the definition. In certain cases we can move straight from a definition of a word's meaning to an explanation of the nature of the reality that the word concerns. For example, we can move straight from defining 'bachelor' as 'adult unmarried male' to the explanation: to be a bachelor is to be an adult unmarried male. So moving back and forth between talking of defining 'bachelor' and defining bachelordom would be an insignificant use/mention sloppiness, of interest only to pedants. But a linguistic definition licenses this move to metaphysics only if it treats the word in question as a descriptive predicate. Where the definition amounts to a revisionist view of the word's meaning, the definition cannot yield a substantial explanation of the nature of the reality that the word appears to concern. Indeed, an antirealist view of that reality is necessary to motivate the revisionist semantics (section 4). Consider an example: we could not move from a noncognitivist definition of 'x is good as compared to y' to an explanation of goodness as 'usefulness', and noncognitivism is partly motivated by an antirealist view of goodness. One lesson I think that we should draw from Harry Field's classic article, "Tarski's Theory of Truth" (1972), is that Tarski's definition of 'true' is of the revisionist sort and so, as it stands, does not show us anything substantial about truth. Tarski's use/mention sloppiness is of more than pedantic interest.

Tarski's definition of 'true-in-L' rests on list-like definitions of various referential words along the lines of the following definition of 'designate':

\[ N \text{ designates } x \iff 'N' \text{ is 'France'} \text{ and } x \text{ is France or...or } N \text{ is 'Germany'} \text{ and } x \text{ is Germany} \]

By comparing such definitions with a similar one for 'valence', Field brings out dramatically that the definitions do not yield satisfactory reductive explanations of the nature of reference. So, the definition of 'true' in terms of the referential words does not yield a satisfactory reductive explanation of the nature of truth. In the light of subsequent discussions, we can see why: the list-like definitions are essentially deflationary and so could not yield anything substantial about reference. Indeed, in offering these definitions Tarski is implicitly committed to antirealism about reference: only if there were something problematic about reference would there be adequate justification for not treating the referential terms as ordinary two-place relational predicates; for not saying, for example, that 'designate' designates the relation designation, or applies pain-value to all ordered pairs where the first member designates the second. Tarski shared the physicalism of the positivists and clearly did think that there was something problematic about both reference and truth. And that was the thought that drove his enterprise. Although Tarski seemed to view himself as a correspondence theorist about truth, the theory he actually presented is deflationary, as I think is now generally agreed. So there is a fair amount of use/mention confusion in representing Tarski's definition as a theory of truth, as Tarski and others do. Tarski's definition tells us a lot about 'true-in-L'. It tells us nothing about truth-in-L, because it is implicitly committed to the view that there is nothing to tell.
3. Scott Soames begins "What is a Theory of Truth?" (1984), an important defense of Tarski from Field's criticisms, with the report: "Alfred Tarski's theory of truth and its successors...are commonly believed by philosophers to provide analyses of the nature of truth" (p. 411). If Soames is right that this belief is common—and I think he is—the use-mention confusion I have just noted is widespread. Soames does not share my view that this belief misrepresents Tarski's achievement, but he notes that "there is considerable doubt about whether, or in what sense, 'Tarski's theory is a theory of truth.' He goes on:

One main reason for this uncertainty is the difficulty of determining what a theory of truth ought to be. Generally, theories of truth have tried to do one or the other of the three main things:

(i) to give the meaning of natural-language truth predicates
(ii) to replace such predicates with substituents, often formally defined, designed to further some reductionist program; or
(iii) to use some antecedently understood notion of truth for broader philosophical purposes... (p. 411)

This is striking. Suppose that we wondered what a theory of, say, genes tries to do. Two things occur: (a) it tries to describe the role of genes—state the laws about genes—which is what Mendelian genetics does; (b) it tries to say what genes are—explain their nature—which is what molecular genetics does. Now explaining the role of the genes is, near enough, analogous to Soames' (iii). But explaining what genes are has no analogue on Soames' list! The metaphysical task of explaining what truth is, which is surely what correspondence theorists and many others were trying to do, has become one or other of the two linguistic tasks, (i) and (ii). Use has become mention.

4. The theory Horwich proposes in his influential book, Truth (1990), is explicitly deflationist. Yet he talks positively of "the minimalist function of truth" (p. xii), of the "entire conceptual and theoretical role of truth" (p. 6), of "the properties of truth" (p. 26), and of "all the facts involving truth" (p. 7). Strictly speaking, on his antirealist theory, truth can have no function, role, or (nontrivial) property, and cannot be involved in any facts. The truth term is what has the function, role, properties and involvement. Horwich claims that his minimal "theory of truth...involves nothing more than the equivalence schema" (p. 12); it is "what is expressed by [the schema]'s uncontroversial instances" (p. 7).24 But this is misleading at best for his theory of truth is really to be found in various negative remarks about the nature of truth, some of which I have quoted (section 5). Of course, the word 'true' is used not mentioned in each instance of the equivalence schema and this might suggest that these instances explain the nature of truth. But that suggestion treats 'true' as a normal descriptive predicate which is precisely what Horwich, like all deflationists, denies: he has a revisionist theory of the truth term.25 This theory of the truth term is what really "involves nothing more than the equivalence schema"; it holds that every fact about the role of the term can be explained simply by taking the meaning of the term to be implicitly defined by its use in the appropriate instances of the equivalence schema.26 The equivalence schema has nothing to do with his theory of truth, everything to do with his theory of the truth term.

An analogy with goodness may help. Suppose that a noncognitivist were to talk positively of the function, role and properties of goodness, and of the facts involving goodness. Suppose that she claimed to be giving a theory of goodness that was quite clearly based on her views about the nondenominative meaning and expressive role of 'good'; for example, she claimed that her theory of goodness involves nothing more than the view that to say that x is good is just to express a pro-attitude toward x. It would be obvious that she was misdescribing her position: on her view, such remarks should really apply only to 'good' not goodness. For her theory of goodness is, roughly, that there isn't any; more precisely, it is the view suggested in section 5.

5. Stephen Leeds' "Theories of Reference and Truth" (1978), and Robert Brandt's "Fragmatism, Phenomenalism, and Truth Talk" (1988)28 are two of the best brief presentations of deflationary truth. Leeds sketches a disjunctive theory of the sort famously suggested by Quine (1970). He comments: "What we have sketched is not a theory of truth...but a theory of the concept of truth" (1978: 222). But then he spoils this assessment by claiming that his account explains "facts about truth-in-English" and "what we ordinarily say about truth" (p. 123). It doesn't. But it might explain facts about 'true' in English and ordinary uses of 'true'. Brandt takes the 'central theoretical focus' of deflationism to be "on what one is doing when one takes something to be true, that is, our use of 'true'." He goes on: "It is then denied that there is more to the phenomenon of truth than the properties of such takings" (1988: 77). But, stric-

24 Consider also: "The basic idea for deflationary theories of truth is roughly that there is no more to truth than the equivalence there..." (Devitt 1997: 50). "The deflationist tells us... Truth's 'nature', such as it is, is (pretty much) exhausted by the equivalence of a claim p with the claim p is true" (Richard 1997: 57).

25 Horwich is happy to go along with the unselective realist about properties, holding that 'true' is a predicate referring to a 'logical' property (p. 38). So, in that respect, instances of the schema are 'about truth'. But it is still a mistake to think that the instances say anything substantial about the nature of truth. Truth as a logical property has no nature open to reductive explanation. Indeed, it has no properties except trivial ones like being logical and being a property. And although we might perhaps take the equivalence schema to yield an explanation of truth, it does not yield a substantial reductive one; cf note 20.

26 On this see pp. 54-7 (abstracting from the confusion of meaning with a speaker's knowledge of meaning, cf. Devitt and Sterfley 1999, eh. 8).

27 Kirkham takes Horwich at his word and so sees his remarks about the equivalence schema as an answer to the metaphysical question about the nature of truth (1992: 329). As a result of this, and Horwich's acceptance that truth is a property, Kirkham does not classify Horwich as a deflationist. Soames takes the 'leading idea' of deflationism to be that the equivalence schema "is in some sense definitional of the notion of truth" (1997: 4). The talk of 'notions' makes this appropriately linguistic. But the talk immediately follows the inappropriately metaphysical "the equivalence schema... use crucial in explaining what truth contains in" (p. 5). And it is immediately followed by the claim that the statement, 'there is no such property as truth', which is straightforwardly metaphysical, is a variation of it (p. 4).

28 Brandt's excellent paper is sadly neglected; it gets no mention, for example, in Kirkham's encyclopedic discussion (1995).
ly speaking, on the deflationary view the properties are not any part of the phenomenon of truth because, roughly, there is no such phenomenon. The only phenomena are truth takings.

6. Finally, consider Marian David's *Correspondence and Disquotation* (1994), the most detailed and informed critique of the disquotational theory of truth available. David starts his description of the disquotational theory by claiming that it is inconsistent, that the correspondence theory, is "an antitheory of truth": its view is, "Truth has no nature." So far, so good. But then he continues: "The correct explanation of truth is that it requires less extravagant resources." The correct explanation is that truth is *disquotation* (pp. 3–4). But the disquotational view does not require less extravagant resources to explain truth, it does not require any because, properly understood, it is the view that truth does not need and cannot have an explanation. That is the respect in which it "has no nature." And disquotation does not explain truth, it explains the truth term.

Not surprisingly, when David sets out to find the unextravagant disquotational theory of what it is for a sentence to be true, he finds the theory "a bit elusive" (p. 62). The core of the disquotational theory is, of course, the equivalence schema. David worries away at the schema trying unsuccessfully to find in it a theory of truth other than a correspondence theory. Sometimes he comes close to realizing that he is seeking something that the disquotationalists think is not there to be found: they think that "sentence-truth is in a sense 'nothing'" (p. 65):

Strictly speaking, the question [about truth] will not even receive a response with the right logical form to count as an answer to this question, for the grammatical truth predicate does not function like an ordinary predicate... Given that the standard way of answering "What is F?" questions does not work when it comes to truth all one can do is describe the linguistic role that the term 'true' plays in our language. (pp. 68–69)

Just so. Still he remains puzzled: "where does the deflationary idea that truth is *nothing but* disquotation come from?" (p. 69). Deflationists have given him reason to be puzzled, as we have seen. Despite what is often suggested, the disquotational view should not be that truth is nothing but disquotation. The view should be that truth is

In this section, I have indicated how pervasive use/mention sloppiness is in the discussion of deflationary truth. Some of this sloppiness is surely insignificant. Yet I hope to have shown that some of it is not: a theory of the truth term is really being taken as a theory of truth. This helps to obscure the metaphysics of deflationary truth and hence the difference between deflationism and the correspondence theory described in sections 5 and 6.

8. Summary

I have attempted to bring out the real difference between deflationism and the correspondence theory by emphasizing the similarity between deflationism and nonfactualism. At the linguistic level, the real difference is fairly apparent. The correspondence theorist can, and should, grant that the truth term has the logical role emphasized by

the deflationist. But the correspondence theorist does not accept the deflationary view that the term has no other role: he holds that it has a descriptive role. Furthermore, he thinks that the term has the standard semantics of a one-place descriptive predicate, a view that the deflationist rejects. At the metaphysical level, the real difference between the two theories is much harder to discern. The typical correspondence theorist thinks that truth has a nature and causal role that need explaining. The deflationist should reveal her antirealism in the characteristic nonfactualist way by rejecting the need for and possibility of any such explanation. Finally, the metaphysical difference motivates the linguistic one, implicitly if not explicitly: it is largely because of her antirealism that the deflationist rejects a standard semantics and a descriptive role for the truth term.

I have located the difficulty in discerning the metaphysical difference in four problems. The first problem is a difference in focus: the focus of deflationism is on the linguistics of the truth term, the focus of the correspondence theory on the metaphysics of truth. The second problem, just illustrated in some detail, is that use/mention sloppiness in discussions of deflationism tends to obscure the real metaphysics of deflationism. The third problem is that when discussions do address the metaphysical issue, rather than merely appearing to when addressing the linguistic issue, what is said is often unsatisfactory. And this is not surprising because it turns out to be rather hard to capture the nonfactual metaphysics of truth, as it is to capture the metaphysics of nonfactualism. That is the fourth problem.

References


---

29 Special thanks to Harry Field whose doubts about the main thesis of this paper have led to many changes. I am grateful for conversations at the University of Sydney and the Graduate Center of City University of New York when versions of the paper were presented. My thanks also to Marian David, Paul Horwich, Mark Lance, Bill Lycan, Paul Pietroski and George Rey for helpful comments. Devitt 2001 is an expanded version of this paper.