Keith Simmons

Deflationism and the autonomy of truth

In *Thought and World,* Chris Hill offers us the most sophisticated version of minimalism to date. Hill's starting point is Horwich's minimal theory of truth (MT), according to which propositions – or *thoughts* in Hill's preferred terminology – are the truth-bearers, and the axioms of the theory are the thoughts expressed by sentences like "The thought that aardvarks amble is true if and only if aardvarks amble". But Hill's theory is superior to Horwich's in a number of crucial respects. For example, while Horwich gives up on a finite formulation of MT, taking it that an infinity of axioms is needed, Hill's basic theory (*simple substitutionalism* or SS) is composed of just one axiom. And unlike MT, SS can accommodate certain generalizations about truth (for example, the generalization that only thoughts are true, and general propositions about the truth-conditional properties of logical connectives.) Moreover, Hill provides a detailed deflationary theory not only of truth, but also of related semantic notions such as reference and concept-denotation, notions that have received relatively little attention in the literature on deflationism; he takes correspondence intuitions much more seriously than most deflationists, and is at pains to show how those intuitions may be accommodated within his framework; he provides the most developed deflationary account available of indexical concepts and indexical thoughts; and unlike many deflationists, he is not content to merely set aside problems concerning bivalence and the Liar paradox. Those of us interested in the proper treatment of our semantic concepts, whether we are deflationists or not, are in Hill's debt.

It is a distinctive feature of deflationism that truth is explained in terms that do not invoke other philosophically weighty notions. According to Horwich, for example, the virtue of minimalism is that "*it provides a theory of truth that is a theory of nothing else*"[1]. As Hill puts it:

> If minimalism is correct, then there is no particular set of concepts that one must acquire prior to acquiring the concept of truth… (p.4)
>
> … minimalism represents the concept of truth as autonomous and presuppositionless. (*ibid*.)
>
> Although there are important differences of detail, all versions of deflationism share with minimalism the optimistic message that the concept of truth is philosophically and empirically innocuous… . (*ibid*.)

And according to SS, truth and other semantic concepts are explained in terms of the logical device of substitutional quantification, and "[a] theory that explains truth and other semantic concepts in terms of a logical device is paradigmatically deflationary" (p.23). On this deflationary view, then, we should not expect to find substantive explanatory relations between truth and other notions to which it is traditionally tied – meaning, assertion, belief, successful action, and the rest.

Deflationism is sometimes characterized in metaphysical terms (that there is no substantive *property* of truth) and sometimes in *linguistic* terms (for example, that the word 'true' is a device for disquotation or a prosentence-forming operator). In my view, however, it is the *concept* of truth with which we should primarily be concerned, and so I applaud Hill for his focus on our semantic concepts. But in what follows I'd like to suggest that the concept of truth is not as autonomous or explanatorily inert as Hill's substitutionalism might suggest.

SS consists of just one axiom:

(S) For any object x, x is true if and only if (Σp)((x=the thought that p) and p).

Here Σ is the existential substitutional quantifier. On pain of circularity, the substitutional quantifiers cannot be characterized in the usual way in terms of truth. Instead, Hill's characterization proceeds in terms of rules of inference, modeled on the elimination and introduction rules for the standard objectual quantifiers. For example, one form of the Universal Elimination rule is this:

(UE)
$$\frac{(\Pi p)(\ldots p\ldots)}{(\ldots T\ldots)}$$

where T is a particular, determinate thought, and (…T…) is the particular, determinate thought that comes from replacing all free occurrences of the propositional variable p in the open thought (…p…) with T. Hill gives an example of an inference authorized by this rule (p.21):

$$\frac{((\Pi p)(\text{if Terry believes that p, then p})}{\text{If Terry believes that the universe is expanding, then the universe is expanding.}}$$

As Hill puts it, the variable p occupies positions that can be occupied by full thoughts, just as variables in rules for objectual quantification occupy positions that can be occupied by nominal concepts (see p.20). In Hill's terms here, the thought *the universe is expanding* replaces all free occurrences of p in the open thought *if Terry believes that p, then p* (see p.21).

But I think this way of putting things is potentially misleading. Hill's characterization of the rules of inference for the substitutional quantifiers goes forward at the level of thoughts, with no mention of language. But if we are after a fully rigorous characterization of these rules, we cannot, I think, avoid the explicit mention of

sentences. The variable p is a linguistic item, part of our logical vocabulary, and when we instantiate UE we replace p by a *sentence* (in Hill's example, the English sentence "The universe is expanding"). What we should say, then, is that p is replaced by a sentence which expresses the thought that the universe is expanding, just as in the parallel case of objectual quantification we say that the variable is replaced by a singular term which refers to an object (and which may be associated with a nominal concept, depending on one's view of singular terms). Moreover, the two occurrences of p in the open sentence 'if Terry believes that p then p' must be replaced by the *same* sentence. It will not be enough to characterize sameness here in purely syntactic terms, because of the possibility of orthographically identical sentences with different meanings. So sameness must be characterized in part as sameness of meaning - though that too is not enough, since, for example, we cannot replace one occurrence of p by "The universe is expanding" and the other by "L'univers s'etend". This suggests that we need a more precise formulation of UE which requires that all free occurrences of p are to be replaced by sentences that are syntactically identical and share the same meaning or interpretation. Corresponding adjustments will be needed for precise formulations of the other rules.

Consider these inference rules in action. Take, for example, an application of Existential Introduction (EI) in Hill's derivation via (S) of the equivalence:

(*) The thought that snow is white is true if and only if snow is white.

In a step of the derivation of the right-to-left direction (p.34), we use (EI) to move from

(3) (The thought that snow is white = the thought that snow is white) and snow is white

to

(4) (Σp)((the thought that snow is white = the thought that p) and p).

Let (…p…) be the open thought expressed by

(5) (The thought that snow is white = the thought that p) and p.

How should we understand the relation between (…p…) and the thought expressed by (3)? Of the three occurrences of "Snow is white" in the sentence (3), only the third expresses a thought – the first two are part of a singular term referring to a thought. Given this difference, it is hard to see how we can follow Hill and regard the thought expressed by (3) as the result of replacing *both* occurrences of p in (…p…) by the *thought* that snow is white. Rather, it seems that we should regard the *sentence* (3) as coming from the open *sentence* (5) by replacing all free occurrences of p by tokens of the English *sentence* "snow is white". The proper formulation of EI should proceed in terms of sentences, not thoughts. And the talk here of tokens of the English sentence "snow is white" is one way of requiring that all the free occurrences of p are replaced by syntactically identical sentences with the same meaning or interpretation.

Let's take stock. Axiom (S) provides a characterization of truth via substitutional quantification, and substitutional quantification is in turn characterized by way of inference rules. So whatever notions are part of the characterization of the inference rules will also be part of the characterization of truth. Consequently SS ultimately accounts for truth in terms that include *sameness of meaning*, and related terms such as *interpreted sentence*, and *sentence-types and sentence-tokens of English*. This seems to compromise the autonomy that Hill claims for the concept of truth; even if we adopt SS, we still need substantive semantic notions for an adequate characterization of truth. Moreover, these notions must themselves be explained independently of truth, on pain of circularity. So, for example, it would seem that a truth-conditions account of *meaning* is

unavailable – perhaps a use theory will serve instead. But then SS seems not to provide a "philosophically innocuous" characterization of truth, but rather an account which carries with it some philosophical commitments. The point here is *not* the familiar, general one that deflationists must reject a truth-conditions account of meaning. The point is rather that the notion of meaning (and other related notions) are built into SS's very definition of truth, so that the definition is adequate only if meaning can itself be characterized independently of truth - and that is a controversial philosophical thesis.[2]

Hill claims that the concept of truth is "autonomous and presuppositionless"; I have suggested that his definition of truth does not meet the standard set by this deflationary claim. Now the deflationist claim that truth is autonomous runs, I think, in two directions. I have been concerned with the first direction: truth is not to be explained in terms of philosophically substantive concepts and theories. In the other direction, truth is to have no substantive role in the explanation of philosophical concepts like meaning, belief, assertion, and the rest. I turn now to this second direction.

Hill writes that "Quine has identified the main functions of the concept of truth" (p.89). Hill takes it to be a virtue of simple substitutionalism that "it dovetails beautifully" with Quine's "extremely plausible conjecture about the role that the concept of truth plays in our descriptive and explanatory practices" (p.24). According to Quine, truth is a device that enables indefinite and generalized endorsements. Following Quine, Hill (along with many other deflationists) focuses on propositions such as *Fermat's last theorem is true*, *What John said yesterday is true*, and *Everything Gandhi said is true*. These propositions do not directly present the evaluated propositions, unlike *The*

proposition that penguins waddle is true; instead, the evaluated propositions are indirectly referred to, or belong to a domain that is quantified over.

In all these cases, the concept of truth applies to propositions, whether they are directly presented, referred to indirectly, or quantified over. Call such applications of the concept of truth *first-order*. Deflationists tend to be concerned almost exclusively with first-order uses of truth. But there are other uses of the concept of truth that are not first-order – uses that are more reflective or theoretical or *second-order*. We say things like

(1) True sentences correspond to reality.

(2) True beliefs engender successful action.

and

(3) To assert p is to present p as true.

Philosophers sometimes use the substantive 'truth' to make claims like

(4) Truth is one, but beliefs are many.

and

(5) Truth is a property definable in the language of some eventual physics.

Deflationists must have something to say about these second-order uses. Perhaps some of them may be dealt with quickly. For example, Brandom argues that (4) and (5) are the result of a mistake: a false analogy is drawn between "true" and ordinary predicates, with the result that a property of truth is hypostatized.[3] But (1)-(3) appear to articulate highly plausible intuitions that forge connections between truth and other concepts, threatening the autonomy of truth. What should the deflationist say about them?

Turning first to (1), Hill provides in Chapter 3 an instructive discussion of (1) in the form:

(CP) For any thought x, if there exists a state of affairs y such that x semantically corresponds to y, then x is true if and only if there exists a state of affairs y such that x semantically corresponds to y and y is actual.

Hill's strategy is this: to construct a derivation of (CP) from (S) and other premises acceptable by deflationist lights. One of those premises is Hill's definition of the intuitive notion of semantic correspondence:

(SC) For any thought x and any state of affairs y, x bears R to y if and only if
(Σp)(x=the thought that p and y=the state of affairs that p).

The remaining premise introduces the notion of actuality:

(A*) (Πp)(if the state of affairs that p exists, then the state of affairs that p is actual if and only if p).

Hill admits that the addition of (SC) and (A*) to (S) compromises simple substitutionalism to some degree, since deflationists are hostile to the notion of a state of affairs and the notion of actuality. Accordingly, he labels this ecumenical position *extended substitutionalism*, and regards it as a quasi-deflationist compromise between classical deflationism and the traditional correspondence theory (see pp.56-7). And one can indeed appreciate the deflationary flavor of (SC): the correspondence relation is captured merely by observing that the name *the proposition that p* and the name *the state of affairs that p* both have the thought *snow is white* as a constituent. (SC) simply pairs up the thought and the state of affairs via this common thought.

One could take issue with the specifics of Hill's treatment of correspondence and truth. For instance, one could take issue with the choice of explanandum. We might take the correspondence intuition to be expressed not by (CP) but by

(CT) For any thought x, x is true if and only if there exists a state of affairs y such that (a) x semantically corresponds to y and (b) y is actual.

(CT) is a stronger proposition than (CP), and Hill does not attempt to produce a derivation of it from (S). Instead he claims that we do not have intuitions that commit us to (CT). For it is a consequence of (CT) that to any true thought there corresponds a state of affairs – but we intuitively accept that there are true thoughts which do not correspond to states of affairs, or at least where it is far from clear that they correspond to states of affairs. We feel comfortable about applying truth to normative thoughts, for example, but we need not feel committed to there being corresponding states of affairs or robust truth conditions, since we may doubt that there are real properties to which normative concepts refer (see pp.54-5). It does seem to me that Hill's argument here is questionable. (CT) is a very natural articulation of the correspondence intuition, and it might well be supposed that if we became convinced that normative propositions did not have robust truth conditions, we would be reluctant to apply the truth concept to them (as indeed non-cognitivists are). Should the fact, if it is a fact, that we apply truth to propositions that do not have robust truth conditions count against (CT), or should it rather count against the appropriateness of some of our everyday applications of the truth concept? The larger issue here is whether a theory of truth should, like Hill's, be "concerned only with descriptive questions about the structure of our actual conceptual scheme" (p.120). The danger with such a purely descriptive account is that we may be misled by ordinary, unreflective applications of truth – including, perhaps, applications to normative propositions.[4]

But whatever particular concerns one may have with Hill's account of correspondence and truth, the general strategy is impeccable: the construction of a derivation of the explanandum from premises acceptable to the deflationist, with the

deflationary axiom (S) among them. Truth plays no more of an explanatory role than the denominalizing role supplied by (S). We see exactly the same strategy in Horwich's explanation of (2) above, that true beliefs engender successful action.[5] Horwich considers the following instance:

> If all Bill wants is to have a beer, and he thinks that merely by nodding he will get one, then, if his belief is true, he will get what he wants.

At one point in his explanation, Horwich makes "the familiar psychological assumption" that if one has a desire, and believes that a certain action will satisfy that desire, one will perform the action.[6] That is, conceptual connections are assumed between belief, desire and action. But all that is assumed about truth in Horwich's explanation is its denominalizing role, as embodied in the axioms of the minimal theory. In the course of the explanation, we move from "The proposition that if Bill nods then Bill has a beer is true" to "If Bill nods then Bill has a beer"; and a little later we move from "Bill has a beer" to "The proposition that Bill has a beer is true". These are the only steps where truth has a role to play, and it is the denominalizing role given to it by the minimal theory.

But this deflationary strategy doesn't always work, or so it seems to me. Consider (3) – the claim that to assert is to present as true. Since this is a natural application of the concept of truth, Hill's descriptivist account should accommodate it. With SS in mind, we might represent (3) as

(Πp)(to assert that p is to present that p as true),

and claim that this is equivalent to

(Πp)(to assert that p is to present that p).

But this would commit us to the claim that to present as true is just to present; for example, to present as true the thought that aardvarks amble is just to present the thought

that aardvarks amble. But this claim is false, for there are many ways to present a thought. I can present it as worthy of your consideration, or as a working assumption, or as a remote possibility, or as outrageous – and I can also present it as true. Presenting as true is just one way of presenting. We have a second-order use of the concept of truth that cannot be denominalized away. In applying the concept of truth here, we are not even purporting to evaluate some sentence or thought. Rather we are offering a general explanation of what speakers are *doing* when they use language in certain ways. No reference is made to any specific proposition or domain of propositions; instead, we are trying to identify a distinguishing feature of a class of *acts* – namely, assertions. The correspondence theory has no difficulty here: what distinguishes the act of asserting from the act of assuming or conjecturing, say, is that when we assert p, we present p as corresponding to how the world is – when we assert, we tell it like it is (or at least how we think it is). On the correspondence conception, this second-order use of truth is of a piece with first-order uses. In contrast, the denominalizing role that deflationists confer on first-order uses of truth does not carry over to this second-order use. There's more to be said, of course, but it does seem that the connection between truth and assertion cannot be explained along deflationary lines.[7]

In my view, assertion is not a special case – I believe that truth plays an indispensable role in the explanation of other concepts too, for example *meaning* and *truth-aptness*, where that role cannot be captured along deflationary lines.[8] But the case of assertion is enough to put pressure on the claim that truth has no explanatory power beyond its denominalizing role. I am suggesting, then, that Hill's substitutionalism neither escapes nor accommodates truth's involvement with substantive philosophical

concepts. Doesn't escape it, because a fully spelled-out version of the axiom (S) will require the concepts of meaning, interpreted sentence, and the like; and doesn't accommodate it, because truth has an explanatory role beyond the one that substitutionalism permits.

References

Bar-On, D., Horisk, C. and Lycan, W.G. (2000): 'Deflationism, Meaning, and Truth-Conditions', *Philosophical Studies* 101.

Bar-On, D., and Simmons, K. (forthcoming): 'Deflationism, Truth, and Assertion', forthcoming in *The Illocutionary Role of the Concept of Truth*, edited by D. Greimann and G. Siegwart, de Gruyter.

Blackburn, S. and Simmons, K. (1999): *Truth*, in the series Oxford Readings in Philosophy, Oxford University Press, Oxford.

Brandom, R. (1994): *Making It Explicit*, Harvard University Press, Cambridge, Mass.

Hill, C. (2002): *Thought and World*, Cambridge University Press, Cambridge.

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

Porubcansky, D. (2004): 'Deflationism and Truth-aptness', MA thesis, UNC Chapel Hill.

Simmons, K. (1999): 'Deflationary Truth and the Liar', *Journal of Philosophical Logic* 28, pp.455-488.

Endnotes

[1] Horwich 1990, p.26, original emphasis.

[2] Since SS's definition of truth ultimately leads us back to sentences, this raises the question: Why not work with sentences in the first place? After all, propositions are regarded with suspicion in certain quarters. And for the deflationist there is the special worry that the very notion of a proposition may not be as innocent of involvement with truth as the minimalist requires. (For more on this, see Blackburn and Simmons 1999, pp.18-22.) Hill says relatively little about the nature of propositions (see pp.2-3), but what he says about their logical structure seems to tie them closely to sentences anyway: "… it is appropriate to view propositions as having constituent structures that parallel the logical structures of sentences" (p.3). Unnecessary dealings with propositions could be avoided by adopting a disquotational version of the axiom (S), say:

For any object x, x is true if and only if (Σp)(x="p" and p),

where we associate a substitution class of sentences with the variable p.

[3] Brandom 1994, p.323.

[4] As a consequence of the Liar paradox and the descriptivist character of his theory, Hill is led to accept that "substitutionalism is committed to the incoherence thesis – that is, to the claim that our conceptual scheme has basic structural features which permit the derivation of contradictions" (pp.118-9). So more is needed: either we need to show that despite appearances our conceptual scheme needs no revising, or acceptable revisions must be found. In my view, deflationists face even more of an uphill struggle with the Liar than substantivists (see Simmons 1999).

[5] Horwich 1990, pp.23-24.

[6] See (5), Horwich 1990, p.24.

[7] More is said about this in Bar-On and Simmons (forthcoming).

[8] See Bar-On et al. 2000 on truth and meaning, and Porubcansky 2004 on truth and truth-aptness.