Chapter 10

Consequences for Deflationism

10.1 Deflationary truth

The contemporary debate about the nature of truth pits the substantivist against the

deflationist. According to the substantivist, truth is a substantive property shared by all truths.

For example, according to the traditional correspondence theory, all truths share the property of

corresponding to a fact or *corresponding to a state of affairs that obtains*. A correspondence

theorist will say that the sentence 'snow is white' is true in virtue of its corresponding to a fact,

and so is the sentence 'grass is green'. Of course, to make the correspondence theory precise, we

will need to say what the truth-bearers are, what facts or states of affairs are, what the

correspondence relation is, and what it is to obtain. But we are already in a position to appreciate

the deep difference between the correspondence theorist and the deflationist. The deflationist is

opposed to the correspondence conception - and any substantivist conception of truth - in a

special, radical way. The deflationist does not present a competing view of the nature of truth;

rather, the deflationist denies that truth has a substantial nature at all.[1]

To focus our discussion, consider one prominent kind of deflationist - the

disquotationalist. Take the familiar T-sentences:

'snow is white' is true iff snow is white

'grass is green' is true iff grass is green

'whales are mammals' is true iff whales are mammals

.

.

.

1

According to the disquotationalist, there is no more to truth than is revealed by these T-sentences. One may think of each of them as a partial definition of truth. Tarski writes:

> "... every equivalence of the form (T) obtained by replacing 'p' by a particular sentence, and 'X' by a name of this sentence, may be considered a partial definition of truth, which explains wherein the truth of this one individual sentence consists."[2]

The sentence 'snow is white' is true if and only if snow is white - and that's all there is to the truth of 'snow is white'. And the sentence 'grass is green' is true if and only if grass is green - that's all there is to the truth of 'grass is green'. These partial definitions do not ascribe a common feature to the true sentences. For the correspondence theorist, true sentences share a common property: they all correspond to a fact. But for the disquotationalist, there is nothing they share in virtue of being true. Truth is not a genuine property.

On this kind of view - championed by Quine, among others - truth is merely a *device for disquotation*. According to Quine, saying

'snow is white' is true

is just an indirect way of saying something about the world - namely, that snow is white. Truth-talk is *eliminable* - we can semantically descend. Talk about the truth of sentences can be eliminated in favor of direct talk about the world.

So why employ 'true' at all? According to the disquotationalist, the usefulness of the truth predicate emerges when we consider sentences like

(a) What Claire said yesterday is true,

and

(b) Every alternation of a sentence with its negation is true.

In these cases, we refer to sentences not by their quote names, but via a definite description or

quantification. Joe may use the definite description in (a) because he cannot remember what Claire said yesterday, but still wishes to endorse it – a case of what is often called *blind ascription*. But whether we mention sentences directly via quote-names, or indirectly via definite descriptions or quantification, the truth predicate reminds us that "though sentences are mentioned, reality is still the whole point"[3]. The truth predicate serves "to point through the sentences to the reality"[4].

And according to the disquotationalist, truth-talk is eliminable even in cases like (a) and (b). Following Tarksi's lead[5], the disquotationalist will point out that every sentence has a quote name, and go on to offer this eliminative analysis of 'true' in (a):

What Claire said is true iff what Claire said = 's_1' & s_1

or what Claire said = 's_2' & s_2

or ...,

where 's_1', 's_2', ... abbreviate sentences.[6] This preserves the spirit of the disquotational account: the truth predicate in (a) points through Claire's utterance to the reality, as expressed by 's_1' or 's_2' or ... .

The disquotationalist will offer a similar account of generalizations over sentences. An eliminative analysis of (b) is given by:

for any sentence x, if x = 's_1' then (s_1 or not-s_1)

or if x = 's_2' then (s_2 or not-s_2)

or ...,

where 's_1', 's_2', ... abbreviate the sentences over which x ranges. We can see (b) as a way of expressing an infinite conjunction, say

3

> Tom is mortal or Tom is not mortal
>
> and Snow is white or snow is not white
>
> and ...,

where each of these conjuncts is about the world.[7] Quine writes:

> ... if we want to affirm some infinite lot of sentences that we can demarcate only by talking about the sentences, then the truth predicate has its use. We need it to restore the effect of objective reference when for the sake of some generalization we have resorted to semantic ascent.[8]

The disquotational treatment of (a) and (b) suggests that 'true' is not just a device for disquotation; it is also a device for expressing infinite disjunctions and conjunctions. This is often referred to as the *expressive role* of truth, and according to deflationists it is the only role that truth plays. For example, according to Michael Williams, the function of truth talk is "*wholly* expressive, *never explanatory*"; truth talk only serves to allow us "to endorse or reject sentences (or propositions) that we cannot simply assert"; "what makes deflationary views deflationary is their insistence that the importance of truth talk is exhausted by its expressive role".[9]

The expressive role of truth suggests a natural disquotational definition of 'true':

DefT   x is true iff (x='s_1' & s_1) or (x='s_2' & s_2) or ... ,

where 's_1', 's_2', ... abbreviate sentences.[10] And falsity may be defined as follows:

DefF   x is false iff (x='s_1' & ~s_1) or (x='s_2' & ~s_2) or ... .

Notice that the T-sentences are (easy) logical consequences of the truth definition, given a suitable infinitary logic - so Tarski's criterion of adequacy is satisfied.[11] Notice also that these definitions are infinitary. It is not clear that there is a satisfactory finite statement of disquotationalism.[12] Given the infinitary character of *DefT* and *DefF*, we finite beings cannot

hope to fully grasp the definientia, and so the disquotationalist must provide some other account of how we come to understand what 'true' means.[13]

In what follows we shall be working with the disquotational account of truth suggested by *DefT*.   There are two features of this account that are worthy of special attention - they capture the basic disquotational intuitions about truth. First, *there is no more to the truth of a sentence than is given by the disquotation of its quote-name*. Second, *the truth predicate is in principle eliminable*: truth-talk can in principle be eliminated in favor of direct talk about the world.

There is another way in which we might present the disquotational account, in terms of a theory of truth rather than a definition. We specify the disquotational theory by specifying its axioms - and its (infinitely many) axioms are just the T-sentences, the instances of the truth-schema

T 'p' is true iff p.[14]

This is, for example, how Horwich presents his minimal theory of truth for utterances.[15] Clearly, the theory is driven by the same disquotational intuitions: the truth of a sentence is still given by disquotation, and 'true' does not appear on the right hand side of the T-sentences. Though for convenience I shall work with *DefT*, it will make no difference whether we think of disquotational truth via the disquotational definition or via the disquotational theory. Either way, disquotation and eliminability are essential features.

10.2 Deflationism extended

Field writes:

> "If truth conditions play no central role in meaning, and truth is fully explained by the disquotation schema (and of value only as a logical device ...), then the same is true of reference...".[16]

And Horwich extends deflationism (or, in his terminology, minimalism) to 'true of' as well as 'refers':

> "For minimalism is primarily the view that the equivalence schema is conceptually basic vis-a-vis the truth predicate - and, analogously, that parallel schema are conceptually basic vis-à -vis 'is true of' and 'refers'" [17]

Of truth, reference and satisfaction (where satisfaction is the converse of the 'true of' relation), Horwich writes:

> "We should expect no deeper analyses of any of these semantic phenomena than are provided by their minimal theories..."[18]

The idea is that just as the T-schema tells us everything there is to tell about truth, so the schemas for 'refers' and 'true of' tell us everything there is to tell about reference and the 'true-of' relation. In his presentation of a Tarski-style disquotational truth theory, Resnik provides disquotational versions of these schemas, along the following lines:

D   't' denotes x iff t=x

S   'F' is true of x iff Fx,

where each occurrence of 't' ('F') is to be replaced by one and the same denoting expression (predicate).[19] And corresponding to each of these schemas there are infinitary definitions of 'denotes' and 'true of':

DefD   x denotes y iff (x='t_1' & y=t_1) or (x='t_2' & y=t_2) or ...

where 't_1', 't_2', ... abbreviate denoting expressions, and

DefS   x is true of y iff (x='F_1' & F_1y) or (x='F_2' & F_2y) or ...

where 'F_1', 'F_2', ... abbreviate predicates.

As in the case of truth, disquotationalists may present their account in terms of these definitions

or as a theory whose axioms are the instances of *D* and *S*.

Any thorough-going deflationist who deflates 'true', 'denotes' and 'true of' will deflate

'extension' too. The notion of an extension is part of the same semantic family as truth,

denotation and satisfaction. Indeed the connections are very tight: x is in the extension of 'F' iff

'F' is true of x. And the relation between a predicate and its extension is analogous to the relation

between a name and its denotation. Of course, the deflationist will not deny that there are, say,

dogs, and need not deny that there is a set of dogs, any more that she will deny that, say, Obama

exists. What she will deny is that there is a substantive semantic relation between the predicate

'dog' on the one hand and dogs or the set of dogs on the other, just as she will deny that there is a

substantive semantic relation between the name 'Clinton' and Clinton.   A disquotationalist about

extensions will point to the schema

E   x is in the extension of 'F' iff Fx,

and present the disquotational account either in terms of a theory whose axioms are instances of

E, or in terms of the corresponding definition:

DefE   y is in the extension of x iff (x='F_1' & F_1y) or (x='F_2' & F_2y) or ... ,

where 'F_1', 'F_2', ... abbreviate predicates.

Either way, the notion of extension is treated as eliminable via disquotation, along with truth,

denotation and the true-of relation.

Deflationists emphasize the expressive role of truth, and typically say far less about

7

denotation, true-of, and extension. But deflationists about these other semantic notions would

presumably claim the same kind of expressive role for them too. Suppose, for example, that Joe

says 'Everyone the headmaster named is in big trouble'. (This may be a case of blind ascription,

where Joe doesn't know to whom the headmaster referred, but knows that, whoever they are,

they're in big trouble.) Then this is a way of expressing what this infinite conjunction expresses:

> if the headmaster used the name 'Smith', then Smith is in big trouble
> and   if the headmaster used the name 'Jones', then Jones is in big trouble
> and   … .

10.3 Deflationism and semantic paradox

Consider the result of putting a pathological expression for x in one of these

disquotational definitions. For example, suppose we put the denoting expression C for x in the

definition *DefD*. The only true disjunct in the definiens - the only one with a true identity

statement as its first conjunct - is this:

> C = 'the sum of the numbers denoted by expressions on the board in room 213'
> and y = the sum of the numbers denoted by expressions on the board in room 213.

And so a use of 'denotes' appears in the definiens. The disquotational definition is circular. The

same circularity arises for any of the disquotational definitions whenever we put a pathological

expression for x. Or consider the axiomatic presentation of the disquotational theory of, say,

denotation, where each instance of the schema *D* is an axiom. Now the problem is that when we

instantiate to a pathological referring expression, we obtain a biconditional in which 'denotes'

appears on the right hand side.

There is a *prima facie* threat of circularity even from expressions that are not

pathological. If we put for x the innocent denoting phrase 'the successor of 0', the definiens will

still contain uses of 'denotes', since every denoting expression is disquoted in the definiens, including those containing 'denotes'. However, in this case the disquotationalist has a ready response: the circle here is not vicious. Only one disjunct in the definiens has a true first conjunct (the conjunct "'the successor of 0' = 'the successor of 0'") - and so all the other disjuncts are false. And 'denotes' does not appear in this one true disjunct. So the entire infinite disjunction is equivalent to a single disjunct in which 'denotes' does not appear.

The point can be extended to certain denoting expressions that themselves involve 'denotes'. For example, consider the expression:

(e) the number denoted by 'the successor of 0'.

In this case the one true disjunct is this (where we put y=1 in *DefD*):

(i) "the number denoted by 'the successor of 0'" = "the number denoted by 'the successor of 0'" & 1 = the number denoted by 'the successor of 0'.

Now this disjunct itself contains a use of 'denotes'. But by *DefD*, we can eliminate this use of 'denotes' by putting 'the successor of 0' for x. We obtain the one true disjunct (again putting y=1):

(ii)   'the successor of 0' = 'the successor of 0' & 1 = the successor of 0,

and this disjunct is free of 'denotes'. The idea is this: in the case of the expression e, we can "selectively disquote". That is, we can trace a path through the disjunctions (i) and (ii) to the denoting expression 'the successor of 0'. In the case of (e), disquotation ultimately leads to a denoting expressions free of 'denotes'.

The same is true for grounded denoting expressions generally: disquotation leads to a denoting expression, or denoting expressions, in which 'denotes' does not appear. The general story will be more complicated, though. A denoting expression may make reference to more

than one denoting expression, and reference may be made by means other than quote-names, such as definite descriptions or quantification. For example, the denoting phrase 'the sum of the numbers denoted by expressions on the board' makes reference to however many denoting expressions there are on the board, and does so by means other than quote-names. The path we trace by selective disquotation must pass through the disquotation of the quote names of each denoting phrase on the board. When describing this path in general terms, the disquotationalist might find the notions of a *determination set* and a *primary tree* useful.[20] But however the story is told, the leading idea is this: while it remains true that the definiens contains uses of 'denotes', in the case of grounded expressions we can always trace a path throught the true disjuncts to denoting expressions free of 'denotes'. In this way, the disquotationalist may respond to the charge of vicious circularity. Parallel accounts can be given for the cases of truth and extension.

If disquotationalism is presented in the axiomatic way, the disquotationalist can make a similar response by adjusting the idea of selective disquotation. Consider the instance of (D) associated with (e):

(a)   "the number denoted by 'the successor of 0'" denotes 1 iff 1 = the number
        denoted by   'the successor of 0'.

Here 'denotes' appears on the right hand side. But the disquotationalist can point out that the right hand side of (a) is the left hand side of another instance of (R):

(b)   'the successor of 0' denotes 1 iff 1 = the successor of 0.

And the right hand side of (b) contains a denoting phrase free of 'denotes'. So here we can trace a path through instances of (R), arriving at a denoting phrase in which 'denotes' does not appear.

But this disquotationalist response will not work for the pathological expressions. For these expressions, we cannot carry out the procedure of selective disquotation. We never arrive

10

at a denoting expression free of 'denotes'. In the case of C, for example, we are constantly led

back to C, and a use of 'denotes'. The threat of circularity is real.

10.4 Three deflationist responses

How might the disquotationalist deal with semantically pathological expressions? One

way is to take the eliminability of 'denotes', 'extension' and 'true' *very* seriously, as a necessary

condition for meaningfulness. If, for example, 'true' is ineliminable by disquotation from an

expression, then the expression is meaningless. So all pathological expressions are meaningless,

and the threat of paradox is removed.[21] A second way is to take eliminability *less* seriously –

accept the ineliminability of the semantic terms from ungrounded expressions. According to this

second response, no restrictions are placed on the disquotational definitions or on instantiations

of the schemas. Here, the result is an uncompromising disquotationalism which embraces all the

instances of the truth schema -- in the case of *DefT every* T-sentence is a consequence, and in the

axiomatic version, every T-sentence is an axiom.[22] A third way is to place the pathological

expressions beyond the scope of the disquotational definitions -- or, on the axiomatic version,

limit the instances of the schemas that can serve as axioms of the theory.[23]   I'll consider these

responses in turn.

10.4.1   According to the first response, an ungrounded expression is meaningless, *because* its

semantic term is ineliminable. It is true that here the disquotationalist is providing independent

grounds for the claim that pathological expressions, such as the Liar sentence, are meaningless.

But the trouble is that this claim is highly implausible, given empirical versions of the paradoxes.

It is hard to see how expressions such as

(C) the sum of the numbers denoted by expressions on the board in room 213

or

(P) unit extension of a predicate on the board in room 213

or

(L) The sentence written on the board in room 213 is not true,

are *meaningless*. Their defectiveness is the result of the empirical circumstances, and not

anything intrinsic to their syntax or semantics. Other tokens of the same type as C

straightforwardly denote a number; other tokens of the same type as P straightforwardly have an

extension; other tokens of the same type as L are straightforwardly true or false. And we can

reason with C, P and L, as we've seen throughout this book.

So the claim that ungrounded expressions are meaningless is implausible. And the claim

is clearly incompatible with the singularity theory. According to the contextual analysis of the

repetition and *rehabilitation* discourses, pathological expressions like C, P and L are not only

meaningful, but they have a definite semantic value when assessed by an appropriately reflective

schema – C denotes π+6, the extension of P has sole member the Moon, and L is true.

10.4.2 According to the second disquotationalist response, no restrictions are placed on the

truth schema, or the denotation and extension schemas. This will require a departure from

classical logic. As we saw in Chapter 8, Field maintains all instances of the truth schema by

adopting a paracomplete logic, in which <A> and True<A> are fully intersubstitutable in non-

opaque contexts. The disquotationalist will welcome this modeling of naïve truth. However, as

we saw, Field's theory introduces the notion of *strong truth or determinate truth*, articulated by a determinacy hierarchy. This greatly complicates the picture for the disquotationalist. According to Field, the general notion of determinate truth is ultimately unintelligible, and so there can be no account of it -- and in particular no disquotational account. So the disquotationalist who follows Field here will have to deny the intelligibility of the general notion of determinate truth, along with the general notion of semantic defectiveness. The disquotationalist will also have to deny the intelligibility of exclusion negation and truth gappiness. I argued in Chapter 8 that this is to deny the intelligibility of notions that appear quite intelligible to us, where that denial is motivated by the need to protect the theory from the threat of paradox.

Alternatively, one could adopt dialetheism, and accept that an unrestricted truth schema leads to contradictions. This really does seem to yield an uncompromising disquotationalism, free of the complications of further notions of truth beyond disquotational truth. On this approach, for truth, denotation or extension, grounded and ungrounded expressions alike are admitted into the scope of *DefD*, *DefE* and *DefT* (or, on the axiomatic version, instantiations of the schemas to ungrounded as well as grounded expressions serve as axioms). A disquotationalist who responds this way will reject the idea that eliminability is at the heart of disquotationalism. Armour-Garb and Beall have argued that this need be no real compromise.[24] Consider the case of truth. If we take the core of disquotationalism to be the twin theses that truth is a device of disquotation, and that the purpose of truth is to talk about the world through talk about sentences, then disquotationalism need not be committed to the eliminability of 'true'.[25] According to Beall, pathological expressions inevitably arise in natural languages that contain semantic terms, but the occurrence of, for example, 'true' in a liar sentence, does not

serve truth's core purpose. That purpose is served only in connection with grounded sentences, where truth can lead us out to the world. Pathological expressions are unintended by-products – Beall calls them *spandrels* – and the semantic terms within them do not play their standard disquotational role.

The picture here is this: we start with a language free of the truth predicate, and then introduce the truth predicate whose only role is to serve as a device of disquotation that allows us to express generalizations that could not otherwise be expressed. With the introduction of the truth predicate there arises spandrels from which 'true' cannot be eliminated – but these ineliminable occccurrences are special cases where 'true' does not play its intended role. In my view, this deflationary picture is mistaken. Truth is more than a disquotational device introduced merely to play a logical role. Truth is a basic concept that has important conceptual work to do, and there are *explanatory* uses of 'true' that cannot be treated disquotationally.[26] But this raises large issues beyond the scope of the present discussion. Putting these issues aside, the dialetheist disquotationalist has to embrace as true the contradictions generated by L and other pathological expressions.

INSERT along the following lines: According to the singularity theory, L is a truth that contains an occurrence of 'true' that cannot be eliminated. Not even God could dispense with the truth predicate here. The truth of L turns on irreducibly *semantic* features of L – that it's pathological, that it cannot be assessed by its associated schema. This compromises the disquotationalist's claim that truth's core purpose is to lead us out from sentences to the world. The truth of L leads to us to semantic facts about the sentence L, expressed in terms of truth. Truth cannot be disquoted away.

So the dialetheist route holds out the promise of a pure disquotationalism.   But this

comes with obvious costs.   Not least is that, as Beall and Armour-Garb point out, the dialetheist

not only accepts that L is true and false, but must also accept that L is neither true nor false.[27]

There are true contradictions in the theory, not just in the object language.[28] Moreover, as we

saw in Chapter 11, dialetheism is subject to revenge paradoxes. And so the present dialetheist

response embraces true contradictions without escaping paradox.   Better, then, to reject this

disquotational response to paradox.

The singularity theory itself has clear consequences for both Field's response and the

dialetheist response.   As we saw in earlier chapters, the singularity theory preserves classical

logic. There is no need to abandon the law of excluded middle or the law of non-contradiction.,

or the principle of bivalence. More generally, the singularity theory is at odds with any

paracomplete, substructural or paraconsistent theory.

10.4.3   The third disquotational response is this: we should place the pathological expressions

beyond the scope of the disquotational definitions, or, on the axiomatic version, limit the

instances of the schema that can serve as axioms of the theory. Horwich writes:

> "we must conclude that permissible instantiations of the equivalence schema are
> restricted in some way so as to avoid paradoxical results ... certain instances of the
> equivalence schema are not to be included as axioms of the minimal theory... ."[29]

Horwich speaks here of avoiding paradoxical results, but the problem is broader than that. As

we have seen, not all pathological expressions generate paradox and contradiction - consider

cases of chains, or the predicate 'is a self-member' or the Truth-Teller 'This sentence is true'. The problem for the disquotational theory goes beyond the strictly paradoxical expressions and the contradictions they produce.

Suppose then that the disquotationalist restricts the scope of the disquotational definitions. One way to avoid semantic pathology is to define 'denotes' or 'extension' or 'true' or 'true of' for a language in which these terms do not appear. According to Horwich:

> "We know that this restriction need not be severe. It need have no bearing on the propositions of science - the vast majority of which do not themselves involve the concept of truth."[30]

But the restriction to languages without these semantic terms, such as the language of science, is surely too severe. The deflationary theory is supposed to apply to natural languages or speakers' idiolects, and semantic terms are present in these languages. Disquotationalists do disagree over the scope of the disquotational definitions: for some, their scope includes foreign languages as well as the home language;[31] for some, it does not go beyond the home language;[32] and for others, it applies only to the sentences or utterances of a single speaker's idiolect.[33] But it seems that no disquotationalist will exclude an utterance like "the number denoted by 'the successor of 0'" or "'Snow is white' is true" - these are utterances of English and of at least some English speakers' idiolects. Indeed, Horwich himself takes his theory to apply to *all possible extensions* of a natural language like English.[34]

So disquotationalists cannot limit themselves to languages free of semantic terms. And whatever restrictions the disquotationalist suggests, they must not *ad hoc*, but principled. The disquotationalist will be hard-pressed to find any intrinsic semantic or syntactic criterion for excluding pathology. 'Empirical' paradoxes are pathological not in virtue of their intrinsic

16

syntactic or semantic form, but in virtue of the empirical facts -- and other tokens of the same type as C or P are unproblematic, and have definite denotations or extensions.

The larger issue here, in the case of truth, is that of *truth-aptness* – and this notion of aptness can be extended to denoting expression and predicates. A simple account of truth-aptness is *syntacticism*, according to which a sentence is truth-apt if it displays the appropriate syntax. If a sentence is declarative in form – if it can be embedded in conditionals, negation, propositional attitude constructions and so on – then it is truth-apt.[35] Wright and Boghossian have proposed the strengthening of syntacticism to *disciplined syntacticism*. For a sentence to be truth-apt, it must not only be declarative, but it must also be part of a discourse that is disciplined, a discourse where "there are firmly acknowledged standards of proper and improper use of its ingredient sentences".[36] But empirical paradoxical sentences clearly pass the tests of syntacticism and disciplined syntacticism.[37] And similarly, paradoxical denoting phrases like C and paradoxical predicate like P clearly pass the tests of syntacticism and disciplined syntacticism carried over to denoting expressions and predicates. Beall and Armour-Garb have argued that disquotationalists (and deflationists generally) can appeal to no more than disciplined syntacticism – with the result that deflationism leads to dialetheism, since the deflationist has no principled reason to exclude liar sentences (and all other ungrounded sentences).[38] The question is whether the disquotationalist can do better than this.

One standard way of characterizing truth-aptness is to say that a sentence is truth-apt if it is either true or false.[39] Perhaps, then, the disquotationalist can offer a principle of exclusion in terms of gaps, taking pathological sentences to be gappy. Gappy sentences are excluded from the scope of DefT, or, on the axiomatic version, only instantiations of the truth-schema to

17

sentences that are true or false can serve as axioms.   This is a natural principle of exclusion for

the disquotationalist. For consider again the disquotational definitions of truth and falsity, and

suppose that the scope of these definitions is limited to a class S of sentences; that is,

DefT   x is true iff (x='s_1' & s_1) or (x='s_2' & s_2) or ... ,

and

DefF x is false iff (x='s_1' & ~s_1) or (x='s_2' & ~s_2) or ... ,

where 's_1', 's_2', are the members of the given substitution class S. It follows easily from these

definitions that a sentence is neither true nor false only if it is not a member of S.[40]   Gappy

sentences are excluded from the scope of the these definitions.[41]

There are two reasons why this way of excluding pathological expressions does not go far

enough.   The first is that, as we've already observed, pathological expressions, even if gappy,

are as syntactically and semantically well-formed as other unproblematic tokens of the same

type. The second is that, if the disquotationalist says no more about gaps, gappy sentences will

be lumped along with, say, Julius Caesar, who is also outside S. But Julius Caesar does not

suffer a truth value gap (except perhaps in a highly attenuated way). So the disquotationalist

must say more. Gappy sentences must be distinguished from tokens of the same type that are

admitted into S, and they must be distinguished from other entities that also fall outside S, but

are not gappy in any natural sense.

At this point, I think the best line for the disquotationalist is to embrace the notion of

ungroundedness, and make a virtue of the circularity that threatened the disquotational

definitions (or axioms). The disquotationalist should observe that when we instantiate the

disquotational definitions to an ungrounded expression, a circle results. And this observation

provides the basis for a disquotational characterization of ungroundedness. We can present informally a procedure for determining whether an expression σ is ungrounded in this disquotational sense. Suppose σ is an expression in the substitution class of the disquotational definition for 'denotes', 'extension', 'true' or 'true of' - i.e. σ is one of the t_i's, F_i's, or s_i's. Now instantiate the relevant disquotational definition to σ. Of the infinitely many disjuncts on the right hand side, only one will have a true identity as its first conjunct. And its second conjunct will be the disquotation of the quote-name of σ. This disquotation may itself make reference to other denoting phrases (in our terms, these will be the members of its determination set), and may itself contain 'denotes' (or 'extension' or 'true' or 'true of'). Eliminate this occurrence of 'denotes' via the disquotational definition, selectively disquoting the members of σ's determination set. Keep going in this way. If we obtain an infinite regress, σ is ungrounded. Since the infinite regress is generated by repeated instantiations of the disquotational definition, it seems plausible to say that ungroundedness is captured disquotationally. On the present disquotational account, then, we characterize ungroundedness in disquotational terms, and then exclude pathological expressions from the scope of the disquotational definitions because they are ungrounded. (Or, alternatively, we exclude as axioms of the theory instances of the schemas associated with pathological expressions, because these expressions are ungrounded). Now we have a principled restriction on the disquotational definitions or the axioms, and the restrictive principle can be expressed in disquotational terms.

This response – a principled restriction on the disquotational definitions or schemas -- comes with concessions. The disquotationalist must give up the idea that 'denotes' or 'extension' or 'true' or 'true of' can in principle be eliminated via disquotation from every ordinary utterance

of English, or from the idiolect of every English speaker. A second related concession is that the disquotationalist must remain silent about uses of semantic terms in ungrounded utterances: they are beyond the reach of the disquotational account. Perhaps the disquotationalist will reply that this is not to concede much - these uses of semantic terms generate semantic pathology anyway, and it is not embarrassing to admit that a disquotational theory of truth cannot accommodate them. I think we should be unmoved by this reply: since the concept of truth in ordinary language gives rise to paradox, we might expect a theory of truth to account for paradox rather than set it to one side.

However, there is a still more powerful reason for finding this third response unacceptable: a version of revenge tailored to the present disquotationalist response. Consider the case of truth. Many items are excluded from the scope of *DefT* or *T* -- Julius Caesar, for example, since Julius Caesar is not true. The present disquotationalist will also exclude L. And so it follows that L is not true. But this is just L again. So a sentence excluded as ungrounded from the scope of the disquotationalist theory is a *consequence* of the theory. Similarly, since the disquotational theory excludes C from the scope of 'denotes', C does not denote – so A and B are the only expressions on the board that do denote. So the sum of the numbers denoted by expressions on the board is $\pi+6$. But here we have C again, and it does denote. C denotes just because the theory excludes it from the scope of 'denotes', on the grounds that it is paradoxical. Parallel reasoning goes through for P.

These revenge paradoxes are of course closely related to the repetition and rehabilitation discourses that motivate the contextual treatments of 'denotes', 'extension' and 'true'. Our present disquotationalist will say that a pathological expression like C or P or L is ungrounded,

and so is outside the scope of the disquotational definition (or, alternatively, the associated instance of D, E or T cannot serve as an axiom of the theory). So C fails to denote, P does not have an extension, and L does not have a truth value. But according to the singularity theory, when assessed from a reflective or neutral context, C *does* denote ($\pi+6$), P *does* have an extension (the unit extension whose member is the extension of 'moon of the Earth'), and L *does* have a truth value (it's true). C, for example, is a denoting expression, even though it contains an ineliminable occurrence of 'denotes', an occurrence that cannot be disquoted away. Indeed, C denotes $\pi+6$ just because it is ungrounded and does not denote when evaluated by the schema fixed by C's context of utterance, unlike the expressions A and B on the board. Once we accept that C denotes – and P has an extension, and L is true -- we must reject this third disquotationalist response.

With respect to truth, it is instructive to consider the correspondence conception here. According to the correspondence conception, L is true if and only if it corresponds to a state of affairs that obtains. And L does, according to the singularity approach. The state of affairs here is a semantic state, that of L's being untrue$_{cL}$. And this state obtains, because L is semantically pathological in its context of utterance. According to the correspondence conception, L is indeed true. The reflective evaluation of L as true is in line with the correspondence intuition.[42]

In correspondence terms, L is true in virtue of its correspondence to a state of affairs that obtains. The state of affairs is a *semantic* state of affairs. If we put things this way, I think we can see why the disquotationalist finds revenge reasoning so intractable. L's truth is grounded in a semantic state of affairs, and not a non-semantic state of the world. Recall the disquotational intuition that truth-talk can in principle always be eliminated in favour of direct talk about the

(non-semantic) world. If we accept the contextual analysis, revenge reasoning shows that this just isn't so for L or any ungrounded sentence. The reasoning establishes certain essentially semantic facts about L (that it is pathological, that it is untrue$_L$), and it's in virtue of these semantic facts that L is true. We cannot represent these facts or states of affairs in non-semantic terms. Truth is not eliminable. Parallel remarks hold for denotation, extension and satisfaction. For example, we arrive at a denotation for C only by recognizing certain *semantic* features of C. In parallel with revenge liar reasoning, the revenge reasoning about C establishes certain essentially semantic facts about C - that it is pathological, that the schema fixed by its context does not provide a denotation for it. It is in virtue of these semantic facts that C denotes, and these facts cannot be represented in non-semantic terms. The notion of denotation is not eliminable.

So if we accept the singularity theory -- or any contextual analysis of direct revenge reasoning -- it follows that we must reject this third response of the disquotationalist. Once the disquotationalist removes ungrounded expressions from the scope of 'denotes', 'extension' or 'true', they give up any assignment of semantic values to these expressions. But the contextual account shows that we can repeat expressions such as C, P and L, and then rehabilitate them, so that C denotes, P has an extension, and L is true.

10.5. The expressive role of truth

As we saw in 10.1, advocates of deflationism emphasize the *expressive* role of truth – the role that truth plays in expressing generalizations that could otherwise only be expressed by infinite conjunctions and disjunctions. A principal reason why Beall and Armour-Garb think that

deflationists should choose dialetheism is that any restriction on *DefT* or *T* will compromise the expressive role of truth. Suppose Joe says:

(J) Everything Kate says is true.

Suppose that among her utterances, Kate produces a Liar sentence L ('L is not true'). If L is barred from the scope of *DefT* or *T*, then it will not be part of the infinite conjunction generated from *DefT* or *T*.   But then 'true' is not playing its disquotational role with respect to one of Kate's utterances, and so it fails to play its expressive role.

But we don't need dialetheic disquotationalism to preserve truth's expressive function here. What exactly does a theory have to do to accommodate truth's expressive role? We take the sentence J to express the infinite conjunction of all instances of

> If Kate says 'p', then p.

This in turn is equivalent to the conjunction of all the things that Kate says (since for those things that Kate doesn't say, the antecedent is false, and the conditional is true). So a theory accommodates the expressive role of truth if J is equivalent to the conjunction of *all* the things Kate says.   And the singularity theory does this.

Consider first a simple case. Suppose Kate utters just two sentences 'Snow is white' and '7+5=12'. In general, the expressive role of truth depends on its disquotational role. J is equivalent to the infinite conjunction:

> If the sentence Kate says = 'Grass is green' then 'Grass is green' is true, and
> if the sentence Kate says = 'snow is white' then 'snow is white' is true, and
> **...** .

By the disquotational truth-schema, this is equivalent to

> If the sentence Kate says = '2+2=4' then 2+2=4, and
> if the sentence Kate says = 'snow is white' then snow is white, and

…

which is in turn equivalent to

> Snow is white and 7+5=12.

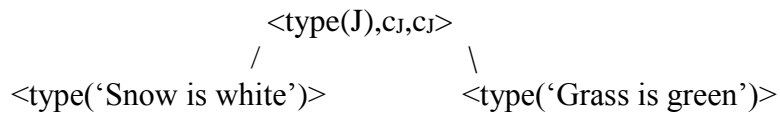Now, with the singularity theory in mind, let 'true$_{cJ}$' represent the occurrence of 'true' in

J. Then J is equivalent to the infinite conjunction

> If the sentence Kate says = 'Grass is green' then 'Grass is green' is true$_{cJ}$, and
> if the sentence Kate says = 'snow is white' then 'snow is white' is true$_{cJ}$, and
> **…** .

By the c_J-schema, and by removing all conditionals with false antecedents, this is equivalent to

> Snow is white and 7+5=12.

Now consider J's determination tree. It looks like this:

$$\langle type(J),c_J,c_J\rangle$$

$$\swarrow \qquad\qquad \searrow$$

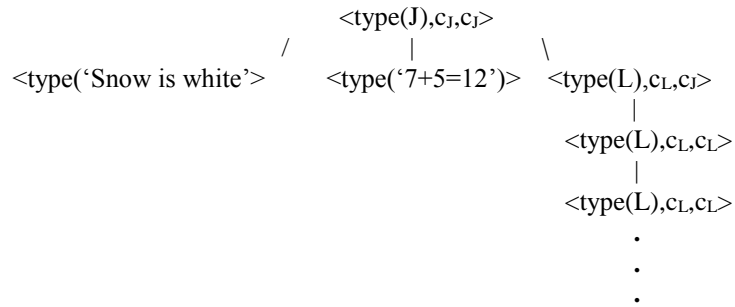$$\langle type(\text{'Snow is white'})\rangle \qquad\qquad \langle type(\text{'Grass is green'})\rangle$$

The equivalence between J and the conjunction 'Snow is white and 7+5=12' can be read off J's determination tree. The truth value of J depends on the values of the nodes at the second tier – these nodes represent the members of J's determination set. In this simple case, the members of J's determination set are straightforwardly true; they are free of the truth predicate. (In terms of the formal theory presented in Chapter 6, they are sentences of *L*. The c_J-schema has no distinctive role to play here -- any truth-schema will yield the same value for a sentence of *L*.) Since J is a universal generalization, it will be true if and only if each node is true. So, in the present case, J is true because each node at the second tier gets the value true. J's determination tree lays out all of Kate's utterances at the second tier – and J is equivalent to the conjunction of these utterances.

But this is so even if Kate produces a Liar sentence. Suppose that in addition to 'Snow is white' and '7+5=12', Kate produces the Liar sentence

(L) L is not true.

Now the primary tree for J looks like this:

$$<type(J),c_J,c_J>$$

$$/ \qquad | \qquad \backslash$$

$$<type(\text{'Snow is white'})> \quad <type(\text{'7+5=12'})> \quad <type(L),c_L,c_J>$$

$$|$$

$$<type(L),c_L,c_L>$$

$$|$$

$$<type(L),c_L,c_L>$$

$$.$$
$$.$$
$$.$$

The triple $<type(L),c_L,c_L>$ repeats on an infinite branch, indicating that L is a singularity of the occurrence of 'true' in L. The triple $<type(L),c_L,c_J>$ does not repeat, indicating that the context c_J is reflective with respect to L. The determination tree for J is:

$$<type(J),c_J,c_J>$$

$$/ \qquad | \qquad \backslash$$

$$type(\text{'Snow is white'})> \quad <type(\text{'7+5=12'})> \quad <type(L),c_L,c_J>$$
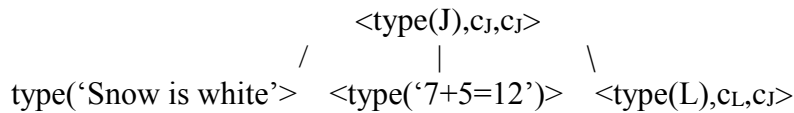
In line with the procedure presented in Chapter 6 (see in particular 6.6 and the discussion there of determination trees), we determine a value for J as follows. We determine a value for each node at the second tier. The values of the first two nodes are straightforwardly true. The value associated with the node $<type(L),c_L,c_J>$ is whatever the reflective c_J-schema yields for L. And L is true$_{c_J}$, since L is indeed not true$_L$ – L is a singularity of true$_L$. Since L is true$_{c_J}$, along with the other members of J's determination set, everything Kate says is true$_{c_J}$. So J is true$_{c_J}$, as is the conjunction of the members of its determination set. That is,

(J) Everything Kate says is true$_{cJ}$

 is equivalent to

        'Snow is white' is true$_{cJ}$ and '7+5=12' is true$_{cJ}$ and 'L is not true$_{cL}$' is true$_{cJ}$

which is equivalent to

        Snow is white and 7+5=12 and L is not true$_{cL}$.

So truth here is playing its expressive role: J is equivalent to the conjunction of *all* the things Kate says, even though one of them is a liar sentence.[43]

There are, however, limits to the expressive role of truth. As we just noted, the expressive role of truth depends on truth's disquotational role, which in turn depends on the applicability of the truth-schema. But sometimes the truth-schema breaks down.

Consider J again, but suppose that Kate says 'Snow is white' and '7+5=12' and also

(K) What Joe is saying is not true.

Then J and K are paradoxically looped – it's easy to check that we're led to contradiction, whether we suppose that J is true or that J is not true. So both J and K are pathological. Now suppose we reason as follows:

(1) Everything Kate says is true iff 'Snow is white' is true' and '7+5=12' is true and 'J is not true' is true.

So, by the truth-schema,

(2) Everything Kate says is true iff snow is white and 7+5=12 and J is not true.

The supposed expressive role of truth here is articulated by (2): J is equivalent to the conjunction of everything Kate says. But should we accept (2)?

On the one hand, we should *not* accept (2). The left hand side is false – one of the things Kate says, namely K, isn't true – it's pathological. But the right hand side is true: in particular, J

is indeed not true, because it's pathological. On the other hand, we *should* accept (2), because both sides are true. The left hand side is true because everything Kate says is true: in particular, K is true, because what Joe is saying is indeed not true, because it's pathological. And the right hand side is true, as before.

The singularity analysis makes sense of all this. It's easy to check that in the present case, the determination tree for J has an infinite branch on which the secondary representations $\langle type(J), c_J, c_K \rangle$ and $\langle type(K), c_K, c_J \rangle$ repeat. So J and K are each singularities of the occurrences of 'true' in both K and L, and neither can be assessed by the c_J-schema or the c_K-schema. (Compare the Fran-Grace loop in 6.5.) J is analysed as

(i) Everything Kate says is true$_{c_J}$.

And the conjunction of everything Kate says is analysed as

(ii) Snow is white and 7+5=12 and J is not true$_{c_K}$.

Now consider (2) again. Here is why we should reject (2). If (2) is to articulate the supposed expressive role of truth with respect to J, then the left hand side is (i) (that is, J) and the right hand side is (ii) (that is, the conjunction of everything Kate says). Now consider (i). Since K is pathological, and a singularity of 'true$_{c_J}$', it is not true$_{c_J}$. So we reflectively evaluate (i) as false. But (ii) is reflectively evaluated as true: in particular, J is a singularity of 'true$_{c_K}$', and so isn't true$_{c_K}$. So in this case of paradox, truth does not play its usual expressive role.

Why did the reasoning from (1) to (2) go wrong? Since the left hand side of (1) is J, (1) is analysed as:

(1) Everything Kate says is true$_{c_J}$ iff 'Snow is white' is true$_{c_J}$ and '7+5=12' is true$_{c_J}$ and 'J is not true$_{c_K}$' is true$_{c_J}$.

In moving from (1) to (2), we apply the c_J-schema to Kate's utterance K – but K is a singularity
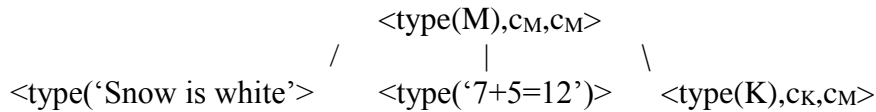
27

of 'true$_{cJ}$', and the application breaks down. And the expressive role of truth requires the applicability of the truth-schema.

Under what interpretation is (2) acceptable? Suppose we add Matt to the picture; Matt says

(M) Everything Kate says is true.

None of Kate's utterances are pathologically looped with M. It's easy to check that M's primary representation does not repeat on M's primary tree – M stands above the loop in which Joe and Kate are caught. (Compare the case of Fran and Grace with a third party Hugo added, in 6.6.) M's determination tree looks like this:

$$\text{<type(M)},c_M,c_M\text{>}$$

$$\diagup \qquad | \qquad \diagdown$$

$$\text{<type('Snow is white')>} \quad \text{<type('7+5=12')>} \quad \text{<type(K)},c_K,c_M\text{>}$$

In providing a value for M, we supply values via the c_M-schema to the nodes at the second tier. In particular, K is assessed in the light of its pathologicality by the reflective c_M-schema. So we have:

(1*) Everything Kate says is true$_{cM}$ iff 'Snow is white' is true$_{cM}$ and '7+5=12' is true$_{cM}$ and 'J is not true$_{cK}$' is true$_{cM}$.

And from this, we can use the c_M-schema to validly infer

(2*) Everything Kate says is true$_{cM}$ iff snow is white and 7+5=12 and J is not true$_{cK}$.

And so (2) read as (2*) *is* acceptable. And now truth's expressive role is restored. The same will hold for any reading of (2) where the evaluating schema is suitably reflective with respect to K.[44]

To sum up: it is an important feature of truth that it allows us to express generalizations

28

that we couldn't otherwise express. When Joe says 'Everything Kate says is true', what he says

is equivalent to a certain infinite conjunction of conditionals, which is in turn equivalent to the

conjunction of all of Kate's utterances. The singularity account captures truth's expressive role,

as it should. Even if the liar sentence L is one of Kate's utterances, still what Joe says is

equivalent to the conjunction of *all* of Kate's utterances. Now, if what Joe says is pathologically

looped with what Kate says, truth fails to play its expressive role, because the truth-schema

breaks down. But, according to the singularity theory, 'true' can still be used to express what the

infinite conjunction expresses – Matt can do it, and so can anyone who is not looped with Kate.

One further point. It is sometimes held that 'true' serves as a mere *endorsement*. This is

another sense in which truth is said to have an expressive role: truth expresses endorsement. To

say that a sentence is true is just to endorse what was said. If we took this kind of view of Joe's

utterance, then Joe is really just saying what Kate said, as evaluated in whatever context Kate

says it.

On this view, if, unbeknownst to me, the Pope just uttered the Liar sentence L ('L is not true'),

and I say 'What the Pope just said is true', then I am saying just whatever the Pope said – and so

I am producing a Liar sentence too.

It is possible to accommodate this view of truth formally. We can follow the procedure

presented in note 43 of this chapter: given a sentence S, evaluate the sentences of S's

determination set according to their *primary* trees, and then evaluate S accordingly.   (According

to the singularity theory, the members of S's determination set are evaluated in accordance with

their *secondary* representations, where the third member of each secondary representation is the

context of S.) In the case of the Pope and me, what I am saying is to be evaluated in the way

indicated by the triple $\langle type(L), c_L, c_L \rangle$. Since L cannot be assessed by the c_L-schema, what I am saying is pathological, just as the Pope's utterance is. From the perspective of the singularity theory, this is perhaps an account of the speech act of 'mere endorsement', but it's not an account of truth. The procedure has stopped too soon: the Liar sentence L is ungrounded, and a singularity of 'true' in L – and so it is to be (reflectively) evaluated as true.

It's rarely, if ever, noted that denotation also plays an expressive role. Consider an example of a blind ascription. I know that Joe referred to, or denoted, a number yesterday, and I remember thinking that the number is irrational – but I cannot remember which number it was. I may say:

The number Joe denoted yesterday is irrational.

Then what I say expresses what would otherwise take an infinite disjunction to express:

The expression Joe used = 'one', and one is irrational
or
The expression Joe used = 'pi', and pi is irrational
or
… .

Let E be the denoting expression that I use (viz., 'the number Joe denoted yesterday'). The determination set for E is the set of phrases to which E makes reference – in the case of E, the determination set has just one member, the expression that Joe used yesterday to denote a number. If Joe used the expression 'pi' yesterday, then the infinite disjunction is equivalent to the second disjunct. The determination tree for E looks like this:
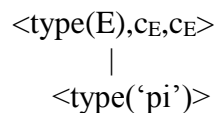
$$\langle type(E), c_E, c_E \rangle$$
$$|$$
$$\langle type(\text{'pi'}) \rangle$$

The tree indicates that the value of E depends on the value of 'pi' – and given that E denotes

whatever Joe's expression denotes, we can read off from the tree that E denotes π, and since π is irrational, that what I say is true. We can see the determination tree here as indicating the equivalence of what I say with the second disjunct, and so the equivalence of what I say with the infinite disjunction.

As with truth, the singularity theory accommodates the expressive role of denotation even if Joe's denoting expression is pathological. Consider a version of our simple paradox of denotation, where yesterday 'pi' and 'six' are written on the board in Caldwell 213. Joe, confused about his whereabouts, writes on the same board:

(C)   The sum of the numbers denoted by expressions on the board in Caldwell 213,

where the occurrence of denotes here is represented by 'denotes$_{cC}$'. In this case, if I say

>   The number Joe denoted yesterday is irrational,

this is equivalent to the same infinite disjunction as before – but this time is equivalent to a different disjunct:

>   The expression Joe used $= C$ and the sum of the numbers denoted$_{cC}$ by
>   expressions on the board in Caldwell 213 is irrational.

As we've seen, from any context reflective with respect to C, we can say that the sum of the numbers denoted$_{cC}$ by expressions on the board in Caldwell 213 is $\pi+6$. Now, in the present case, the primary tree for my denoting expression E is

$$
\begin{array}{c}
\langle type(E), c_E, c_E \rangle \\
| \\
\langle type(C), c_C, c_E \rangle \\
| \\
\langle type(C), c_C, c_C \rangle
\end{array}
$$

```
                <type(E),cE,cE>
                      |
                <type(C),cC,cE>
                      |
                <type(C),cC,cC>
              /       |       \
        type(A)    type(B)   <type(C),cC,cC>
                             /     |      .
```

$$\text{type(A)} \quad \text{type(B)} \qquad .$$

$$.$$

And the determination tree for E is:

$$<\text{type(E)},c_E,c_E>$$
$$|$$
$$<\text{type(C)},c_C,c_E>$$

The trees show that the context of E is (non-explicitly) reflective with respect to C. And from the determination tree, we can read off a value for E. The value of the node at the second tier is the reflectively established value for C, namely $\pi+6$, and E will have this same value. And so what I say is true, since $\pi+6$ is irrational. The determination tree helps to map out the equivalence between what I say and the appropriate disjunct – and hence helps to map out the equivalence between what I say and the infinite disjunction.

If my use of E should turn out to be pathologically looped with Joe's expression, then the associated denotation schema will break down, and so my use of 'denotes' cannot play its expressive role. But suppose a denotation for E can be reflectively established. Then someone, perhaps Matt again, can produce a token of the same type as mine which will stand above the loop in which Joe and I are caught -- and Matt's use of 'denotes' will play the expressive role that mine cannot.

'True of' and 'extension' can also play an expressive role. For example, we can imagine a blind ascription involving 'true of': "The word Joe used yesterday to describe Kate was true of her". Without a semantic predicate, this could only be expressed by an infinite disjunction. And if we know Joe is always talking about Kate, we might say "Kate was in the extension of every adjective Joe used yesterday". Without a semantic expression, this could only be expressed by an infinite conjunction. The singularity theory will treat the expressive roles played by 'true of'

32

and 'extension' along the same lines as 'true' and 'denotation'.

10.6 The prosentential theory and Horwich's minimalism

I've focused on disquotational theories of truth, but there are other deflationary theories, notably the prosentential theory and Horwich's minimalism, and we can ask of them how they deal with the paradoxes.

10.6.1 The prosentential theory

According to the prosentential theory, 'true' is used in forming prosentences. In the discourse:

Mary: Chicago is large

John: If that is true, it probably has a large airport

the expression 'that is true' is a prosentence, which shares its content with its antecedent, namely 'Chicago is large'. In a parallel fashion, the Liar sentence 'This is false' relies on its antecedent for its content,

> "but it is, unfortunately, its own antecedent and, as such, fails as an antecedent supplier of content."[45]

According to the prosentential account, Liar sentences fail to have content.

But consider the familiar revenge or strengthened reasoning about the Liar sentence

(L) L is not true.

We start with the Liar sentence L, reason to the conclusion that L is defective, and conclude:

(L*) L is not true.

In a discussion of strengthened reasoning, Grover remarks that "if 'true' is prosentential", then L* "fails to express a proposition"[46]. This conditional seems right. *If* we accept the prosentential

33

account, then we will accept that L is without content, and that L* relies for its content on L.

And consequently, L* will fail to express a proposition. But it has been a theme of this book that

we can assess semantically defective expressions – in particular, we can assert of a Liar sentence

that it isn't true. So L* not only has content, it is true. So if we accept the singularity theory, or

any theory that accommodates our ability to say that Liar sentences are not true, we must reject

the prosentential theory. Further, if we accept the singularity account, we will also accept that L

is true (when evaluated reflectively), and has the same content as L*. So even the basic Liar

sentence has content, contrary to the prosentential theory.[47]

Similar points can be made about the case of denotation. In her account of the Berry

paradox, Grover writes:

> "The phrase 'integer described in less than 19 syllables' should inherit from a set
> of antecedents, but because one of the antecedents happens to be the Berry
> description itself, there's an ungrounded inheritor. The Berry expression fails to
> refer."[48]

But the strengthened reasoning about the denoting phrase C shows that a phrase can have itself

as an antecedent, and yet still denote.

10.6.2 Horwich's miminalism

Horwich takes propositions, not sentences or utterances, to be the primary truth-bearers.

This, it seems to me, generates further difficulties peculiar to Horwich's version of deflationism.

According to Horwich [1990], the axioms of the minimal theory of propositions are all the

propositions whose structure is

E^* <<p> is true iff p>,

where '<p>' is written for 'the proposition that p'.[49] For example, one of the axioms is the

proposition

(1) <<snow is white> is true iff snow is white>.

According to Horwich, these axioms together constitute a complete theory of truth; no more

needs to be added.

We should note a difficulty with Horwich's presentation of the minimal theory in his

[1990]. According to Horwich, "[*E**] is a function from propositions to propositions."[50] I

cannot see how it is. If we are told that f(x) is a function from integers to integers, we expect to

replace the variable x in the expression 'f(x)' by the name of an integer, and thereby obtain an

expression that denotes an integer. But what do we put for each occurrence of 'p' in *E**?

Suppose we put a name of a proposition for each occurrence. Consider the first replacement: the

angle brackets will be around the name of a proposition, and the inappropriate result is the name

of a name of a proposition. Consider the second replacement - that will yield a name as the right

hand side of a biconditional.

In Horwich 1998, things are presented differently. The angle brackets are now

interpreted as follows:

> "I am employing the convention that surrounding any expression, *e*, with angled
> brackets, '<' and '>', produces an expression referring to *the propositional
> constituent expressed by e*." (Horwich 1998, p.18, fn.3)

The propositional structure E* remains the same, but Horwich's account of how we arrive at it is

different:

> "… we can begin with any one of the axioms and note that the sentence
> expressing it may be divided into two complex constituents. First there is a part
> that is itself a sentence and which appears twice.[51] In the case of (1), this is
>
> (3) 'snow is white'.

And second there is the remainder of the axiom-formulation – namely, the schema

(E) '<p> is true iff p'.

Now we assume that if a complex expression results from the application of a schema to a sequence of terms, then the meaning of the expression is the result of applying the meaning of the schema to the sequence of the terms' meanings. In particular, since, as we have just seen, the sentence

(1*) '<snow is white> is true iff snow is white'

is the result of applying (E) to (3), then the proposition expressed by (1*) is the result of applying what is expressed by the schema (E) to what is expressed by (3). That is to say, the axiom (1) is the result of applying the propositional structure (E*) to the proposition

(3*) <snow is white>."[52]

The difference between these two presentations is significant. In Horwich 1990, the presentation goes forward in terms of propositions only, as if in formulating the minimal theory we can set aside sentences and deal directly with propositions. But we can't. In Horwich 1998 the interpretation of the angled brackets and the account of the propositional structure (E*) are in terms of expressions, sentences, and the *expressing* relation -- and inevitably so.[53] Consider again Horwich's schema

<p> is true iff p.

Instances of this schematic generalization are obtained by replacing the two occurrences of 'p' by tokens of an English sentence. For example, if we put a token of 'snow is white' for each occurrence of 'p', we get (1*). [54] Clearly certain conditions must be placed on such an instantiation. We can list four:

(i) each 'p' is replaced with tokens of an (actual or possible) English sentence,

(ii) these tokens are given the same interpretation,

(iii) under that interpretation they express a proposition,

and

(iv) the terms "that" and "proposition" are given their English meanings.

And the notion of *proposition* is itself semantically loaded. Horwich has suggested that we should

> "characterize propositions by means of the principle
> 　　　*p* expresses the proposition that p"
> here the sentence type, *p*, is individuated semantically as well as physically,
> and where the two tokens of 'p' are understood in the same way".[55]

So the very statement of Horwich's minimal theory is not innocent of other semantical notions.

Horwich's formulation is shot through with semantical concepts – *express, interpretation*, *meaning* -- and talk of sentences.   This raises a question and a challenge. The question is this:

Why not work with sentences utterances all along, and adopt the more economical schema *'p' is true iff p*? The question is all the more pointed, given Horwich's own claim that the minimal theory of truth for propositions is easily inter-derivable with a minimal theory of truth for utterances.[56] And the challenge for the minimalist who takes propositions as the primary truth-bearers is this: explain the semantical concepts needed to articulate the theory without anywhere using the notion of truth in the explanations, on pain of circularity.

　　　Let us now bring the semantic paradoxes into the picture. According to Horwich, certain instances of *E** must be excluded, on pain of the Liar.[57] One might expect the instance

　　　　　The proposition that L is not true is true iff L is not true

to be one such instance. But on the right hand side, 'true' is predicated of a sentence and not a proposition, and so it is not obvious how the minimal theory will handle it.[58] Consider instead, then, a version of the Liar more clearly suited to the minimal theory. Following Horwich,[59] we

can take the locution *what Oscar said* to refer to the proposition Oscar asserted. And when we say: "What Oscar said is true", we are attributing truth to a proposition. Now suppose Oscar unwittingly lands in paradox, saying:

(Q)   What Oscar is saying is untrue.

Here, the predicate 'untrue' applies to the proposition expressed by the sentence Q that Oscar uttered. We obtain the relevant instance of the schema *E* by substituting the sentence Q for 'p':

(e) The proposition that what Oscar is saying is untrue is true iff what Oscar is saying is untrue.

What should we make of (e)? What is the proposition referred to on the left hand side? Well, it is the proposition expressed by the sentence Q. But what proposition is that? The answer will depend on one's theory of truth, and in particular on one's account of the Liar. For example, according to the singularity theory, and other contextual theories, the Liar proposition expressed by Q is true, when assessed from a suitably reflective context. So the left hand side of (e) is true, if the evaluating truth schema is suitably reflective. And we obtained the right hand side of (e) by putting the sentence Q for 'p' - and if Q is true, then the right hand side is true. So both sides of (e) are true. But then (e) itself is true - it is not a contradictory instance of *E* at all. However, another theory of truth, providing another account of the Liar, may produce very different results. Perhaps the theory will identify another proposition, or no proposition, as the subject of the left hand side of (e), and produce a different truth value, or no truth value at all, for the instance (e). Different theories of truth will generate different conclusions about the status of (e). The point is that we can only evaluate (e) if we are already in possession of a theory of truth. Given a Liar sentence, it is a highly non-trivial question as to what proposition, if any, it expresses.

In short, we should beware of an axiom schema for a theory of truth that is couched in terms of the schematic phrase 'the proposition that p'. For it takes a theory of truth to determine the reference of 'the proposition that p', at least when we put a Liar sentence for 'p'.[60]   And so it takes a theory of truth to determine whether certain instances of *E* are true or not. But then we cannot in general regard instances of Horwich's schema *E* as *axioms* of a theory of truth - for there may be instances whose truth is established by a *prior* theory of truth.   Such instances will be theorems of the prior theory - they will not have the axiomatic status that Horwich's minimal theory accords them. And it will not do to simply exclude pathological sentences from the scope of *E*: that was the moral of 10.4.3. We have good reason, I have argued, to attach truth values to pathological sentences. And if that is right, pathological sentences are in the scope of the truth-schema. And they yield true instances of *E* that cannot be regarded as axioms of a minimal theory of truth. For, again, the truth of these instances is a consequence of a prior theory of truth.

Parallel points can be made about denotation, satisfaction and extension. For Horwich, propositions (rather than sentences) are the truth-bearers, and *propositional constituents* (rather than linguistic expressions) denote and are satisfied. So, for example,

> "The propositional constituent expressed by the word, 'Aristotle', refers (if at all) to Aristotle."[61]

and

> "The propositional constituent associated with the predicate 'F' is satisfied by, and only by, things that are F."[62]

This account presupposes a way of moving from a denoting expression or a predicate to its associated propositional constituent. But what is the propositional constituent associated with, say, the pathological denoting expression C, or the pathological predicate P? This is surely a

39

substantive question, one that can only be answered by a substantive theory of denotation or satisfaction. Even if, like Horwich, we leave open the question of what a proposition is and what its propositional constituents are, we can see that different theories will yield different answers. If we adopt the singularity theory, or any contextual theory, then the associated propositional constituent will reflect the presence of a context-sensitive term in C (or P). And the theory will provide good reason for supposing that the propositional constituent associated with C *does* denote (or that the propositional constituent associated with P does have an extension). Another theory may yield a propositional constituent tied to a level of language. Another may deny that there is any associated propositional constituent, because C is pathological. And so on.

Now consider Horwich's schema for denotation and satisfaction:

Ref   (x)(<d> refers to x <-> d=x)

and

Sat   (x)(s satisfies <F> <-> Fx),

where '<d>' abbreviates 'the propositional constituent associated with the denoting expression d' and '<F>' abbreviates 'the propositional constituent associated with the predicate F'. Horwich's minimalist theory of denotation comprises the instances of *Ref*, and the minimalist theory of satisfaction, the instances of *Sat*. But, focusing on *Ref* for the moment, we cannot regard the instances of *Ref* as *axioms* of a theory of denotation. For the very statement of *Ref* presupposes an account of the relation between denoting expressions and their associated propositional constituents. And for that we need a prior theory of denotation. This is true in particular for pathological denoting expressions such as C. And again, since we have good reason to say that the propositional constituent associated with C *does* denote, we cannot simply withhold C from

40

the scope of the denotation schema. Pathological denoting expressions provide true instances of *Ref*, where the truth of these instances is established by a prior theory of denotation. And so the instances cannot be regarded as axioms. Similar remarks can be made about the instances of *Sat*, and about the instances of the extension schema too.[63]

10.7 Concluding remarks

Our present discussion lies at the intersection of two contemporary debates about semantic concepts. On the one hand there is the debate between deflationists and substantivists (principally, correspondence theorists) about the nature of our semantic concepts. On the other hand, there is the debate about the proper treatment of the semantic paradoxes. It might be thought that since the paradoxes are a problem for everyone, it can little impact on the debate between deflationists and substantivists. Marian David puts it this way:

> I have completely neglected the paradox of the liar in my discussion of disquotationalism and the correspondence theory. The reason is simple. Since it afflicts both candidates, considerations concerning the paradox will not be of much help in advancing the debate between substantive and deflationary theories of truth.[64]

And it is not just that the Liar is a shared affliction. It is also claimed that the possible cures - the various proposed solutions to the Liar - will be equally available to the deflationist and the correspondence theorist alike. Horwich writes:

> "There is no reason to suppose that the minimalist answers that are advanced in this essay could be undermined by any particular constructive solution to the paradoxes - so we can temporarily set those problems aside."[65]

In my view, the Liar, and the semantic paradoxes generally, cannot be set aside in this way.[66] When it comes to questions about the nature of truth - and about the nature of reference, satisfaction and extension - solutions to the paradoxes are not neutral. In particular, revenge or

strengthened liars present serious difficulties for those who take truth to be disquotational truth.

For example, Field takes truth to be disquotational truth, to be captured by a Kripke-style theory.

Revenge challenges, generated by our ability to say that Liar sentences are not true, take us

beyond disquotational truth to the series of stronger and stronger truth concepts in Field's

determinacy hierarchy. I've argued in Chapter 8 that this determinacy hierarchy carries us

further and further away from natural language. And yet revenge reasoning – whether in terms

of denotation, extension or truth – is intuitive, natural reasoning (as we saw with C, P, and L, for

example). And since this kind of reasoning takes us away from disquotational truth, it follows

that disquotationalism never really pinned down *our* notion of truth in the first place.

To approach things from the opposite direction: if we accept the singularity theory, then

we cannot be deflationists. According to the singularity theory, we can evaluate pathological

expressions, despite their pathology. Ungrounded denoting expressions denote; ungrounded

predicates have an extension; ungrounded sentences have a truth value. And, borrowing a term

from the correspondence theorist, we can say that the values of these pathological expressions

depend on essentially semantic *facts*, including the fact of their own ungroundedness. These

facts can be described only in semantic terms. If the singularity theory is right, there is no

getting rid of 'denotes', 'extension', 'true' or 'true of'. Our semantical concepts cannot be

disquoted away.

Notes to chapter 10

1. The denial that truth has no substantial nature can take more than one form. It might be the claim that truth has a nature, but a thin or trivial one. For example, one might think that truth is a property because 'true' is a predicate. Versions of disquotationalism might be taken this way. Or it might be the claim that truth is not a property at all; consider, for example, the prosentential theory, according to which 'true' is not even a predicate.

2. Tarski 1944, p. 50. See also Tarski 1930-1, p. 155. I should note that although some deflationists have drawn heavily from Tarski's work, it is far from clear that Tarski himself is a deflationist. For more on this, see Simmons 2009, pp.541-559.

3. Quine 1970, p. 11.

4. ibid.

5. Tarski 1930-1, p.159.

6. The disjunction might be finite, if the language or idiolect to which 's_1', 's_2', ... belong contains only finitely many sentences. But here, and throughout, the presumption will be that the language in question contains infinitely many sentences, and the corresponding disjunction is infinite.

7. The derivation of the infinite conjunction
   (s_1 or not-s_1) and (s_2 or not-s_2) and ...
from the analysans is straightforward, given an infinitary logic.

8. Quine, op. cit. p.12.

9. Michael Williams 1999, p.547 (original emphases).

10. Such a characterization is suggested by remarks in Leeds 1978, pp.121-1 and fn.10; and versions of it is presented explicitly in Field 1986, p.58, Resnik 1990, p.412, and David 1994, Chapter 4 and p.107.

11. For example, put 'snow is white' for x. We will obtain just one true disjunct on the right hand side. If we eliminate the false disjuncts, the definition yields:
'snow is white' is true iff 'snow is white'='snow is white' and snow is white.
Dropping the true conjunct of the right hand side, we obtain the T-sentence:
   'snow is white' is true iff snow is white.

12. Can we find a disquotational definition of truth that is finitely stated? Tarski considers this schematic definition:
   x is a true sentence iff $\exists p(x='p'\&p)$. (Tarski 1930-1, p.159)

Obvious problems arise if we interpret the quantifier objectually. There is the problem of quantifying into quotes. And the open formula "x='p'&p" is grammatically ill-formed, since what follows the conjunction sign is not a sentence but a name. Moreover, the component "x='p'" is problematic: put names of sentences for 'x' and 'p', and the component will identify a sentence with a name of a sentence.

Instead, then, we may take the quantifier substitutionally. In the standard terminology, we write:

$$x \text{ is a true sentence iff } \Sigma p(x='p'\&p),$$

where we associate with the variable 'p' a set of expressions that are acceptable substituends (here, sentences of the language for which 'true' is defined). But how should we read the definiens? We can provide metalinguistic truth conditions: that there si a substitution instance of the open sentence "x='p'&p" which is true. But the specification of these truth conditions is in terms of truth, the very notion we are trying to define. So we need another reading. And it seems that the disquotationalist can do no better than to revert to DefT. The idea is to take "$\Sigma p(x='p'\&p)$" as abbreviating an infinite disjunction, namely, the right-hand side of DefT (see Field 1986, pp57-59, and David 1994, p.98).

A disquotationalist might abandon a direct definition of truth in favour of a recursive account, according to which 'true' is defined Tarski-style in terms of the more basic notions of reference and satisfaction. Given a language with a finite stock of names and predicates, reference may be disquotationally defined by a finite list of sentences of the form "'a' refers to a", and satisfaction by a finite list of sentences of the form "x satisfies 'F' iff x is F". In this way, reference and satisfaction are finitely defined - and so truth is finitely defined. But such a recursive disquotationalist is restricted to languages whose sentences have the appropriate kind of logical form. And there is an array of truths that are notoriously hard to fit into the Tarskian mould: belief attributions, counterfactuals, modal asssertions, statements of probability, and so on. (Problems for the recursive disquotationalist are discussed in David 1994, pp.117-9; Field 1994, p.269; and Horwich 1990, pp.28-31 and Horwich 1999, pp.253-254.) In the face of these difficulties, the disquotationalist might well prefer a direct definition of truth for sentences, even if it is infinitary.

13.   According to Horwich, for example, our understanding of 'true' is a matter of our inclination to accept instantiations of the truth-schema (see Horwich 1998, p.35).

14. Of course, we will expect the T-sentences to be theorems of a correspondence theory of truth. But the correspondence theorist will do more than just affirm the T-sentences - she will also make substantial claims about the nature of truth.

15. Horwich 1990, pp.103-109.

16. Field 1994, p.261.

17. Horwich 1999, p.248.

18. Horwich 1990, p52, fn.3. See also Horwich 1990, pp.122-4.

19. See Resnik 1990, pp.414-415. My formulations differ slightly from Resnik's in that I have worked in terms of 'true of' rather than 'satisfies', and I have used the term 'denotes' instead of 'designates'. For another disquotational version of the schema for 'refers', see Field 1994, p.261. Horwich formulates the schemas in terms of proposition parts, since he takes propositions to be the primary truth bearers (see Horwich 1990, pp.122-4). But presumably Horwich would find the disquotational versions agreeable too, since he takes his minimal theory of propositions to be equivalent to a disquotational theory for utterances, *modulo* two principles that he regards as uncontroversial (see Horwich 1990, p.107).

20.   There are also connections here to Kripke's theory of truth, and a suggestion of Horwich's (couched in terms of propositions) in Beall 2005, pp.81-2.

21.   See Jc Beall 2001 for an articulation of this proposal.

22**.** See Beall and Armour-Garb 2003.

23.   For example, Horwich claims that the only acceptable solution to the Liar is to restrict the equivalence schema: "only certain instances of the equivalence schema are correct" (Horwich 1998, p.41).

24.   See Armour-Garb and Beall 2001.

25.   See *op. cit*.

26   Bar-On and Simmons 2007 argues that truth is needed to explain assertion, and Bar-On, Horisk and Lycan 2005 argues that truth is needed to explain meaning.

27.   Beall and Armour-Garb point out that the claim that L is both true and false is equivalent to the claim that L is neither true nor false, given the schema *T* and de Morgan principles (which are valid in Priest's dialetheic logic LP, in Priest 1979) – see Beall and Armour-Garb 2003, p.318.

28.   Indeed, dialetheists are likely to dissolve the object language/metalanguage distinction altogether. For more critical discussion of this feature of dialetheism, see Littmann and Simmons 2004.

29.   Horwich 1998, pp.40-1.

30.   Horwich 1998, pp.41-2.

31.   See for example Horwich 1998, fn 3 on pp.18-19, and pp.100-1.

32.   See Resnik 1990, pp.413-4; in Resnik 1997, this restriction is eased.

33.   See Field 1994, p.250, where the disquotational conception is characterized in terms of the

sentences that a given speaker understands: "... for a person to call an utterance true in this pure disquotational sense is to say that it is true-as-he-understands-it. ... a person can meaningfully apply "true" in the pure disquotational sense only to utterances that he has some understanding of ...".

34. See Horwich 1998, fn 3, pp.18-19, and pp.100-1.

35. Syntacticism is discussed in Jackson *et al*. 1994, pp.291-3.

36. Wright 1992, p.29. As Boghossian puts it, the sentence must be "significant", or, more fully, must "possess a role within the language: its use must be appropriately disciplined by norms of correct utterance" (1990: 163)

37. And it should not be thought that a theory of truth in particular can be guided solely by the desire to maximize the scope of the truth definition, or to maximize the number of T-sentences we accept as axioms. Take a theory in which the diagonal lemma can be proved. Vann McGee has shown that, in the context of such a theory, there are many consistent sets of T-sentences that are maximal and mutually incompatible (see McGee 1992).

38. Beall and Armour-Garb 2003, pp.308-311.

39   Jackson, Oppy and Smith 1994, p.287.

40. Take a gappy sentence, and suppose, towards a contradiction, that it is a member 's_k' of the substitution class. Given that s_k is not true, we obtain the negation of the infinite disjunction that forms the right hand side of DefT. And from this we obtain $\sim s_k$. Similarly, given that s_k is not false, we obtain $\sim\sim s_k$. So we have $\sim s_k \& \sim\sim s_k$, a contradiction.

41. Beall has argued that the deflationism is compatible with truth value gaps – see Beall 2002. This might suggest that, on the assumption that pathological expressions are gappy, the disquotationalist can accommodate the Liar and related sentences.   But it is only a restricted disquotationalism that is compatible with gaps. Beall assumes that the central deflationary notion of truth is *weak* truth, for which
              'p' is true iff p
holds whether p is true, false, or gappy. In contrast, if truth is *strong* truth, and 'p' is gappy, then the biconditional fails to hold, since the left hand side is false, and the right hand side is gappy. So the disquotationalist is not offering a disquotational account of strong truth.

        Beall points out that strong truth can be defined in terms of weak truth and exclusion negation (where strong negation takes a false sentence or a gappy sentence to a true sentence). But if the disquotationalist admits exclusion negation, then, as Beall acknowledges, paradox returns in the form of the sentence 'This sentence is not true', where negation is exclusion negation. The disquotationalist is now set on a path back to dialetheism.   As we saw in Chapter 8, Field is one disquotationalist who rejects exclusion negation.

42. The strengthened liar will pack some surprise for the correspondence theorist (as it will for

any truth theorist). On the correspondence account, L does correspond to a fact - the fact that it is not true$_{cL}$. But we cannot relate L to this fact via 'true$_{cL}$' - and here is the surprise. But the basic correspondence intuition remains intact - L is true (when evaluated reflectively) in virtue of its correspondence to a fact.

43. We might have the following intuition: given that Joe says 'Everything Kate says is true', and given that it turns out that one of the things Kate said is pathological, and so not true, it follows that what Joe says is false. We could capture this intuition formally by the following procedure. Consider the primary trees for each of the sentences that Kate says (these are the members of J's determination set). If any of these trees show that the sentence is pathological, take that to be its final evaluation, and evaluate what Joe says accordingly. Since the primary tree for L indicates that L is pathological, the procedure yields the result that J is false.

   According to the singularity theory, this procedure stops too soon – we can reason through pathology and reflectively determine a value for L. L is indeed not true$_{cL}$. But Joe's use of 'true' is tied to Joe's context of utterance, not to the context c$_L$. And Joe's context of utterance is reflective with respect to L (though not explicitly so), and L is true$_{cJ}$. This is a final reflective evaluation of L, and so what Joe says is evaluated accordingly. According to the singularity theory, each member of J's determination set is to be evaluated by the c$_J$-schema, not by the schema associated with that member's context of use.

44. It's perhaps worth noting here a contrast with hierarchical approaches to the liar. Consider Joe's utterance: 'Everything Kate says is true'. A hierarchical account has to give some account of how the level of Joe's utterance gets fixed. Burge, for example, appeals to a pragmatic principle of Verity ("subscripts on 'true' are assigned ceteris paribus so as to maximize the interpreter's ability to give a sentence truth conditions by way of a truth schema" (Burge 1979, in Martin 1979, p.109.)   The singularity theory doesn't need to appeal to a pragmatic principle of interpretation here. Matt's evaluation of Kate's utterances is suitably reflective in virtue of the semantic network mapped out by M's primary tree.

45. Grover 1992, p.124.

46. *op. cit.* p.203.

47. Grover 2005 argues that while liar sentences are syntactically well-constructed, they do not have "operative meaning" – they do not have a communicatively significant *use*, any more than the term '6÷0'. As Grover notes, this encourages a strengthened version of the liar: (S) S is false or lacks operative meaning (pp.191-3). Now a contradiction seems derivable from Grover's key claim that S lacks operative meaning. From this claim it follows that S is false or lacks operative meaning. And substituting in the truth schema
S is true iff S is false or lacks operative meaning,
we obtain that S is true, and from this a contradiction follows. Grover responds that since S lacks operative meaning, it cannot be substituted into the truth schema, and the reasoning is blocked. However, it's part of Grover's theory that S lacks operative meaning, and so a consequence of her theory that S is false or lacks operative meaning -- yet we cannot infer the truth of this consequence of Grover's own theory. And certainly there is a use for the sentence

'S lacks operative meaning' – Grover uses it to articulate her theory. Thanks here to Jamin Asay.

48. Grover 1992, p.

49. Horwich 1990, p.18.

50. *op. cit.*, p.42.

51. This sentence also appears in Horwich 1990, except that the word 'proposition' appears instead of 'sentence' (Horwich 1990, p.19).

52. Horwich 1998, p.18.

53   In both his 1990 and 1998, Horwich offers an alternative characterization of the axioms of the minimal theory (1990, p.19, fn.4; 1998, p.18, fn.3.) The axioms are anything that is expressed by instances of the *sentence schema*
            <p> is true iff p.
And this schema "is instantiated by sentences in any possible extension of English" (1990, p20, fn.4). Here the suggestion seems to be that we substitute a *sentence* for each occurrence of 'p' - whether it occurs within angled brackets or not. Again, in formulating the minimal theory for propositions, we cannot sidestep sentences.

54. We may feel some discomfort here: the tokens of 'aardvarks amble' are placed in two quite different contexts. The first token forms part of a referring term, the term '<p>' which abbreviates 'the proposition that p'. The second constitutes the right hand side of the biconditional. With Davidson, we may wonder how these two appearances of 'aardvarks amble' are connected (Davidson 1996, pp. 273-4).

55. Horwich 1999 'The Minimalist Conception of Truth', in Blackburn and Simmons 1999, p.251.

56. Horwich 1998, pp.103-108.

57. Horwich 1998, pp.40-42.

58.   Horwich himself presents the liar by letting '#' abbreviate 'THE PROPOSITION FORMULATED IN CAPITAL LETTERS IS NOT TRUE', and considering the results of assuming # to be true, and assuming # to be not true (Horwich 1998, p.40). Here again it is not easy to disentangle sentences and propositions. It seems natural to suppose that what is formulated in capital letters is a sentence rather than a proposition, so that # should be replaced by, say, 'THE PROPOSITION EXPRESSED BY THE SENTENCE FORMULATED IN CAPITAL LETTERS IS NOT TRUE'. Notice that here in the very articulation of the liar we do not escape sentences and the expressing relation.

59. See for example, *op. cit.*, p.16.

60. Davidson 1996 raises a related question for Horwich's theory:
   "How are we to understand phrases like 'the proposition that Socrates is wise'? In giving a standard account of the semantics of the sentence 'Socrates is wise', we make use of what the name 'Socrates' names, and of the entities of which the predicate 'is wise' is true. But how can we use these semantic features of the sentence 'Socrates is wise' to yield the reference of 'the proposition that Socrates is wise'. Horwich does not give us any guidance here."
Horwich replies to Davidson's concerns in Blackburn and Simmons 1999, p.250, and in Zeglen 1991. Whether or not there is, as Davidson suggests, a general problem for Horwich concerning the reference of 'the proposition that p', there is at least an acute problem when we put a Liar sentence for 'p'.

61. Horwich 1998, p.27.

62. *op. cit.*, p.28.

63. For further discussion of Horwich's minimalism, see Simmons 2016.

64. David 1994, p.191; see also p.7 and p.70.

65. Horwich 1998, p.42.

66. In Simmons 1999, I argue that the deflationary conception of truth is compromised by the Liar in ways that the correspondence theory is not.