Forthcoming in Synthese

Minimalism and Paradoxes *

Michael Glanzberg
Massachusetts Institute of Technology

Abstract. This paper argues against minimalism about truth. It does so by way of a comparison of the theory of truth with the theory of sets, and consideration of where paradoxes may arise in each. The paper proceeds by asking two seemingly unrelated questions. First, what is the theory of truth about? Answering this question shows that minimalism bears important similarities to naive set theory. Second, why is there no strengthened version of Russell's paradox, as there is a strengthened Liar paradox? Answering this question shows that like naive set theory, minimalism is unable to make adequate progress in resolving the paradoxes, and must be replaced by a drastically different sort of theory. Such a theory, it is shown, must be fundamentally non-minimalist.

Why is there no strengthened version of Russell's paradox, as there is a Strengthened Liar paradox? This question is rarely asked. It does have a fairly standard answer, which I shall not challenge. But I shall argue that asking the question helps to point out something important about the theory of truth. In particular, it raises a serious challenge to an increasingly popular version of *deflationism* about truth.

To see what the problem is, it is useful to ask another slightly off-beat question. What is the theory of truth about? There is, of course, an obvious answer to this question: the theory of truth is about truth bearers and what makes them true. Taken at face value, this answer seems to bring with it a commitment to a substantial notion of truth. A deflationist about truth might well wish to give a very different answer: the theory of truth is not really about anything. There is no substantial property of truth, so there is nothing—no domain of objects, or properties, or phenomena—which the theory of truth describes. Not all positions called 'deflationism' subscribe to this view, but the important class of so-called *minimalist* views do.

I shall argue that this sort of deflationist answer is untenable, and thus argue in broad strokes against minimalism. I shall argue by way of a comparison of the theory of truth with the theory of sets, and consideration of where paradoxes may arise in each. I shall show that deflationist positions that accept the idea that truth is not a real or substantial property are too much like naive set theory. Like naive set

* Thanks to Otavio Bueno, Alex Byrne, Ned Hall, Richard Heck, Jim Pryor, Susanna Siegel, Ralph Wedgwood, and Steve Yablo for valuable comments and discussions. Thanks also to two anonymous referees for this journal. A version of this material was presented at the Society for Exact Philosophy Annual Meeting, 2001. I am grateful to my audience there for many useful comments and questions.

theory, they are unable to make any progress in resolving the paradoxes, and must be replaced by a drastically different sort of theory. Such a theory, I shall show, must be fundamentally non-minimalist.

This paper is divided into six sections. Section (1) makes the comparison between naive set theory and the minimalist version of deflationism, and explains the sense in which both theories can be said not to be about anything. Section (2) points out that both theories suffer from nearly identical problems, as Russell's paradox and the Liar paradox may be seen to be extremely similar in important respects. Section (3) shows where these parallels break down. The sort of response to the Liar which might be offered by minimalism proves to be unstable, in that it is vulnerable to the Strengthened Liar paradox. The standard response to Russell's paradox in set theory is not so unstable. Section (4) investigates the source of this difference. It argues that any theory of truth able to evade the Strengthened Liar must at least be about some domain in the way that standard set theory is about the domain of sets. Section (5) argues that the usual ways of avoiding the Strengthened Liar meet this condition by abandoning minimalism for a more correspondence-like notion of truth. Section (6) then goes on to show that no minimalist position can meet the condition, and so none is tenable. The paper concludes by briefly considering whether any other version of deflationism might fare better.

1. Minimalism and Naive Set Theory

It is well-known that the banner of deflationism is flown by a great diversity of philosophical views. The class of views with which I shall be concerned here is marked first by the idea that in some appropriate sense there is no substantial or genuine property of truth. For instance, Paul Horwich writes:

> Unlike most other predicates, 'is true' is not used to attribute to certain entities (i.e. statements, beliefs, etc.) an ordinary sort of property—a characteristic whose underlying nature will account for its relation to other ingredients of reality. Therefore, unlike most other predicates, 'is true' should not be expected to participate in some deep theory of that to which it refers . . . (Horwich, 1990, p. 2.)

Horwich does not quite claim that there is no such property as truth, but the sense in which there is a property seems to amount to little more than there being a predicate of truth in our language. There is no genuine phenomenon of being true which this predicate describes.

There are two other marks of the sort of view I shall consider, both of which center around the *T-schema*:

's' is true iff s.

First, those who maintain that there is no substantial property of truth cannot maintain that instances of this schema hold because of the nature of the property—because of the way truth is. Instead, they must say that the schema holds analytically, or perhaps by definition, or perhaps by stipulation. There are, of course, important differences between these specific proposals, but they will not matter for our purposes here. What is important is that a status for the T-schema is provided which ensures its truth without looking to the nature of the property of truth ('underlying nature' as Horwich puts it). In what follows, I shall compare this to the status of logical truth. Second, those who advocate this sort of position will point out that rather than describing a feature of truth, the T-schema provides us with a device of disquotation. This device is useful, for instance, as it allows us to make infinitary generalizations. Hence, the stipulative or analytic T-schema provides us with a useful linguistic device, rather than describing a genuine property, and there is nothing to truth beyond this.

It is more or less standard terminology to call the view that subscribes to these three ideas (no substantial property of truth, analyticity of the T-schema, and truth as a device of disquotation) *minimalism*. It will be useful to identify a particularly straightforward version of minimalism, which which I shall call *pure minimalism*. Pure minimalism is marked by taking the instances of the T-schema to hold for any well-formed declarative sentence. Beyond that, it insists that the commitments of minimalism are all there is to say about truth.[1]

We may think of pure minimalism as factoring into two components. One is a theory in the logician's sense. The core of the theory is the T-schema, taken now as an axiom schema:

(T) $$Tr(\ulcorner\phi\urcorner) \leftrightarrow \phi.$$

[1] Many current minimalist positions depart from pure minimalism in some ways. I shall return to other versions of minimalism in Section (6). It is a somewhat delicate question whether anyone has actually held pure minimalism, but I believe it encapsulates an important idea, which is reflected in the positions of a number of authors. It is often attributed to Ayer (1946). Ayer defines truth for propositions rather than sentences, which seems technically to make him something other than a pure minimalist. However, as I shall discuss more in Section (6), his definition of proposition is sufficiently closely tied to sentences that this may not be a significant difference. Horwich (1990) likewise holds most of the theses of pure minimalism, but construes truth as applying to propositions. Especially some remarks he makes in Horwich (1994) make me wonder if his departure from pure minimalism is equally minimal. Quine (1986) comes quite close to pure minimalism, but of course, he would never stand for an inconsistent theory like M.

We need to construe this as added to a theory strong enough to do some elementary syntax. It must have a name $\ulcorner\ulcorner\phi\urcorner\urcorner$ for each sentence $\ulcorner\phi\urcorner$, and I shall assume it can carry out the proof of the diagonal lemma. Let us call this theory M.

The other component consists of the philosophical commitments of pure minimalism. In this case, it is odd to describe them as explaining the intended interpretation of M, but they do explain how M is to be understood.

Pure minimalism has the additional feature of minimalism about *truth bearers*. Truth bearers are appropriate candidates for truth, and have a truth status. They are true or false.[2] Truth bearers need not be true, but they must be *truth apt*. To be a truth bearer, according to pure minimalism, is nothing but to figure into predications of $\ulcorner Tr\urcorner$ or $\ulcorner\neg Tr\urcorner$. In the presence of the T-schema or the axiom schema (T), this is simply to be a well-formed declarative sentence. Pure minimalism is thus minimalist about truth bearers in that it says no more about what makes something a truth bearer than it does about truth. We may reasonably suppose the syntax component of the theory will be able to delineate the class of well-formed declarative sentences, but notice this tells us nothing about their status as truth bearers. When it comes to this status, all that the theory tells us derives from the analytic (or whatever other appropriate status) schema (T). There is thus no underlying property that makes it the case that declarative sentences are all truth bearers, as there is no underlying property that makes the instances of (T) hold.[3]

I shall compare pure minimalism to *naive set theory*. Like pure minimalism, this theory factors into two components. One is again a formal theory. Again it is captured primarily by a single axiom schema, the *naive comprehension schema*:

$$\text{(COMP)} \qquad y \in \{x \mid \phi(x)\} \leftrightarrow \phi(y).$$

Like (T), we must think of this as added to an appropriate base theory, which is able to construct a name $\ulcorner\{x \mid \phi(x)\}\urcorner$ for the set determined by $\ulcorner\phi\urcorner$. We might also assume a principle of extensionality, but it will not matter for the discussion to follow. Let us call this theory N.

[2] On some views, which rely on many-valued logics, truth bearers are true, false, or any of the other truth values.

[3] A well-know argument of Jackson et al. (1994) attempts to show that minimalism about truth does not lead to minimalism about truth bearers. I do think that the points I am making here reveal a genuine commitment of pure minimalism. I also doubt they really conflict with the substance the points of Jackson et al., though this is a rather delicate issue. I shall return to it in Section (6), when discussing departures from pure minimalism. For the moment, it would be all right with me if we counted truth-bearer minimalism as a component of pure minimalism by stipulation.

Like pure minimalism, naive set theory comes with a philosophical component as well. I have in mind naive set theory as it would have been understood by someone who really held it: Frege, or in some form perhaps a traditional logician.[4] Such a theorist would hold that (COMP) is in some way a *logical* principle. Now, there has been a great variety of ideas about what makes something a logical principle. But common to them is the thought that insofar as logical principles might be taken as being about anything, they are about absolutely everything. They do not, in particular, hold because of the nature of a *specific* range of objects or properties or phenomena. In this regard, naive set theory is remarkably similar to pure minimalism. Those who hold these theories can agree that there is no underlying nature of anything in particular which makes the fundamental principles of their respective theories true. Both agree that there are no specific properties that the fundamental principles of their theories describe. Rather, these principles are in the general class of the logical, or the analytic, or the definitional. I do not want to go so far as to assimilate logical truth to analytic or definitional truth; I only need to note that they share the feature of there being no underlying natures or particular properties to which such principles owe their truth. There is thus a crucial similarity between the status ascribed to the fundamental principles of pure minimalism and of naive set theory.

In light of this, both theories may be described as not being about anything in particular. Now, it may be noted that the truth predicate occurs in (T) and set abstracts in (COMP), so we could say that one is about truth and the other sets. But we say this in an entirely minimal way. There are no special domains of objects, properties, events, or any other phenomena that make the axioms true. In the case of naive set theory, this is reflected both by the philosophical gloss on the theory, and by the unrestricted nature of set abstraction. As a matter of logic, any objects of any kind may be collected into a set. There is no special domain of the theory. In the case of pure minimalism, we can likewise observe that the theory is neither supposed to reveal a basic feature of the property of truth, nor does it provide any more substantial an account of what makes something a truth bearer. There is no more a special domain of this theory than there is of the naive theory of sets. As I mentioned above, the truth bearers are the syntactically well-formed declarative sentences. It is thus tempting to say that the theory is about these sentences. But it is so only in a trivial way. The theory appropriates some syntax, but this tells us nothing about truth. The

[4] Both Frege (at some moments) and, say, the *Port-Royal* logicians, would have preferred 'extension' to 'set'.

principles that are supposed to tell us something about truth fall into a different category, and these hold of well-formed sentences only because this the only way the stipulations themselves wind up syntactically well-formed. Truth is thus predicated as widely as makes syntactic sense, not on the basis of the nature of any particular domain. Though when we write the theory down we rely on some syntax to do so, a far as the basic commitments of minimalism go, there is in no substantial sense a special domain of the theory of truth.

The two theories differ somewhat in the way they describe this situation: pure minimalism does not quite claim to be a matter of logic, and naive set theory offers nothing like semantic ascent. But both rely on principles which hold in some other way than by accurately describing a domain, and as a result both are in similar ways not genuinely about anything.

2. Paradoxes

Both M and N are inconsistent, as is well-known. Russell's paradox shows N to be inconsistent, and the Liar paradox does the same for M. But the response to its paradox has been quite different for each. Naive set theory is usually taken to be a disaster, while minimalism is often taken to be in need of modification but still viable. Given how similar the two theories are, this is odd. It appears even more odd when we reflect on how similar the paradoxes really are. This will lead us to consider a crucial difference between responses to the paradoxes, which will in turn show us something about the viability of minimalism.

Let us first consider the paradoxes. For the Liar, we appeal to the diagonal lemma to find a sentence $\lceil\lambda\rceil$ such that:

$$M \vdash \lambda \leftrightarrow \neg Tr(\ulcorner\lambda\urcorner).$$

Combining this with (T) gives the contradiction:

$$M \vdash Tr(\ulcorner\lambda\urcorner) \leftrightarrow \lambda \leftrightarrow \neg Tr(\ulcorner\lambda\urcorner).$$

The diagonal lemma hides the diagonal procedure for producing $\lceil\lambda\rceil$, but it is clear that $\lceil\lambda\rceil$ 'says of itself' that it is not true. We then ask about the truth of this sentence, and see that it is true just in case it is not true.[5]

We do the same for Russell's paradox. With the Liar, we found a sentence that says of itself that it is not true. Here we need a predicate

[5] Many minimalists add a clause saying something like 'only non-problematic instances of (T)'. The success of this has been discussed by McGee (1992) and Simmons (1999).

that says something is not in itself, i.e. $\ulcorner \neg x \in x \urcorner$. With the Liar, we asked about the truth of that very sentence. Here we ask about this predicate applying to its own extension. Let its extension be $R = \{x \mid \neg x \in x\}$. From (COMP) we have:

$$N \vdash R \in R \leftrightarrow \neg R \in R.$$

As with the Liar, we have a contradiction.

The two paradoxes differ in that Russell's paradox involves predication, and the Liar paradox truth, but otherwise, we do basically the same thing in both. Indeed, we may bring out the similarity even more, by replacing (T) and (COMP) with a single principle. Consider a family of predicates $\ulcorner Sat_n(x, y_1, \ldots, y_n) \urcorner$, and corresponding axioms:

(SAT) $$Sat_n(\ulcorner \phi \urcorner, y_1, \ldots, y_n) \leftrightarrow \phi(y_1, \ldots, y_n).$$

If we replace $\ulcorner Sat_0(\ulcorner \phi \urcorner) \urcorner$ by $\ulcorner Tr(\ulcorner \phi \urcorner) \urcorner$, we have (T). If we replace $\ulcorner Sat_1(\ulcorner \phi \urcorner, y) \urcorner$ by $\ulcorner y \in \{x \mid \phi(x)\} \urcorner$ we have (COMP).

A more general diagonal construction yields the inconsistency of (SAT). We need only be able to prove for any predicate $\ulcorner F(x, y_1, \ldots, y_n) \urcorner$ there is a $\ulcorner Q(y_1, \ldots, y_n) \urcorner$ such that:

$$Q(y_1, \ldots, y_n) \leftrightarrow F(\ulcorner Q \urcorner, y_1, \ldots, y_n).$$

This is a straightforward modification of the more familiar diagonal lemma. (See Boolos, 1993.) Let S be a theory that contains (SAT) and can prove the general diagonal lemma.

The proof that S is inconsistent is the generalization of both the Liar and Russell arguments. Consider the predicate $\ulcorner \neg Sat_n \urcorner$ for any n. Using the general diagonal construction, we may find predicates $\ulcorner \sigma_n(y_1, \ldots, y_n) \urcorner$ such that:

$$S \vdash \sigma_n(y_1, \ldots, y_n) \leftrightarrow \neg Sat_n(\ulcorner \sigma_n \urcorner, y_1, \ldots, y_n).$$

Combining this with (SAT), we have:

$$S \vdash Sat_n(\ulcorner \sigma_n \urcorner, y_1, \ldots, y_n) \leftrightarrow \sigma_n(y_1, \ldots, y_n) \leftrightarrow \neg Sat_n(\ulcorner \sigma_n \urcorner, y_1, \ldots, y_n).$$

For each n, the schema (SAT) produces inconsistency.

The argument here is the same as that used in both the Liar and Russell's paradoxes. For the case of $n = 0$, we have the Liar. The $\ulcorner Sat_0 \urcorner$ instances of (SAT) are just (T), and the diagonal lemma yields $\ulcorner \lambda \urcorner$. For the case of $n = 1$, we have a version of Russell's paradox. The $\ulcorner Sat_1 \urcorner$ instances of (SAT) yield a version of (COMP). The diagonal lemma give us a formula $\ulcorner \rho(y) \urcorner$ such that:

$$\rho(y) \leftrightarrow \neg Sat_1(\ulcorner \rho \urcorner, y).$$

This is essentially the Russell predicate. Applying it to itself we see:

$$\rho(\ulcorner\rho\urcorner) \leftrightarrow \neg Sat_1(\ulcorner\rho\urcorner, \ulcorner\rho\urcorner) \leftrightarrow \neg\rho(\ulcorner\rho\urcorner).$$

This makes all the more clear that the formal difference between the Liar paradox and Russell's paradox is just one of a parameter, which has little effect on the way the paradoxes are generated.[6]

We now have seen two theories, pure minimalism and naive set theory, that are strikingly similar. We have also seen two paradoxes, or rather two versions of basically the same paradox, which show the two theories to have inconsistent formal components. From here on, however, the situations with truth and sets diverge rather drastically, as we shall see in the next section.

3. Strengthened Paradoxes

Let us briefly consider how to respond to the paradoxes. The pure minimalist wants to hold on to the philosophical account of truth, as much as is possible, but avoid what might be seen as a merely technical failure of M. The usual approach is to say that though the basic idea behind (T) is right, it is technically misstated.

One leading idea for revising (T) is to make the truth predicate somehow *partial*, so that problematic sentences like $\ulcorner\lambda\urcorner$ come out neither true nor false. There are many different ways to implement this idea, but for discussion purposes, it will be useful to concentrate on one that remains in some ways close in spirit to M. The idea is to replace the axiom schema (T) with the following collection of inference rules:

$$\text{(INF)} \qquad \frac{P \vdash Tr(\ulcorner\phi\urcorner)}{P \vdash \phi} \qquad \frac{P \vdash \neg Tr(\ulcorner\phi\urcorner)}{P \vdash \neg\phi}$$
$$\frac{P \vdash \phi}{P \vdash Tr(\ulcorner\phi\urcorner)} \qquad \frac{P \vdash \neg\phi}{P \vdash \neg Tr(\ulcorner\phi\urcorner)}.$$

Call the resulting theory P. A theory like P can be modified or extended in many ways, but it will suffice to illustrate the point as it stands.[7] P

[6] My presentation of M and N, and of the paradoxes, draws heavily on Feferman (1984). For further discussion of $\ulcorner Sat_n\urcorner$, and its relation to set theory, see Parsons (1974b).

I should mention that in pointing out the similarities between the Liar paradox and Russell's paradox, I am not particularly taking issue with the original distinction between semantic and logical paradoxes of Ramsey (1926). The issues he raised are somewhat different than those that bear here.

[7] There are other ways to implement the idea of partiality. One, for instance, is to construct a partial interpretation for $\ulcorner Tr\urcorner$, along the lines of Kripke (1975).

makes truth partial in the following sense. For some sentences $\ulcorner\phi\urcorner$, we have $P \vdash Tr(\ulcorner\phi\urcorner)$, so according to P, $\ulcorner\phi\urcorner$ is true. For some sentences $\ulcorner\phi\urcorner$ we have $P \vdash \neg Tr(\ulcorner\phi\urcorner)$, so according to P, $\ulcorner\phi\urcorner$ is false. (Observe if $P \vdash \neg Tr(\ulcorner\phi\urcorner)$, then $P \vdash Tr(\ulcorner\neg\phi\urcorner)$.) But for some sentences, like $\ulcorner\lambda\urcorner$, we have neither, so P assigns such sentences neither the value true nor the value false. (In many cases, we expect to get results like $\ulcorner Tr(\ulcorner\phi\urcorner)\urcorner$ from P together with some other theory, which tells us the facts of some special science. This does not matter for a sentence like $\ulcorner\lambda\urcorner$, which contains no terms from any special science. All that appear in $\ulcorner\lambda\urcorner$ are a sentence name, $\ulcorner Tr\urcorner$, and the negation operator.)

Philosophically, it might be argued that the move from M to P does not change the commitments of pure minimalism. The rules in (INF) might be glossed as having the same analytic or definitional status as (T), and they do substantially the same job of introducing a device of disquotation. The minimalist will have to accept some partiality, which raises a great many issues, but let us grant P to the pure minimalist for argument's sake.

P is consistent; yet the Liar paradox makes trouble for it nonetheless. This is because of what is known as the *Strengthened Liar paradox*. We reason as follows. P is designed to ensure that $\ulcorner\lambda\urcorner$ does not come out true, i.e. $P \nvdash Tr(\ulcorner\lambda\urcorner)$. So it seems, using P as a guide, we have come to conclude $\neg Tr(\ulcorner\lambda\urcorner)$. But $\ulcorner\lambda\urcorner$ just 'says' $\neg Tr(\ulcorner\lambda\urcorner)$. P itself tells us this, as $P \vdash \lambda \leftrightarrow \neg Tr(\ulcorner\lambda\urcorner)$. Thus, it appears that just relying on P, we have come to conclude λ. We are now back in paradox.

Now, this inference cannot be carried out in P, so P remains consistent. But it still poses a problem. The conclusion we draw seems to be entirely correct, whether it can be carried out in P or not. P is purpose-built to make sure $\ulcorner\lambda\urcorner$ does not come out true. That is how consistency is achieved. So, we simply rely on P to come to the conclusion that $\ulcorner\lambda\urcorner$ is not true. Insofar as P is supposed to capture the notion of truth, it appears this is just the conclusion $\neg Tr(\ulcorner\lambda\urcorner)$. Opinions differ on just how serious a problem this is, and how it may be solved.[8] For our purposes here, all I need to note is that the inference is intuitively compelling, and poses a problem that requires a solution one way or another.

This is often done in the setting of a three-valued logic. A theory like P eases the comparison with a theory like naive set theory, so I have chosen this route. Most of what I say applies equally to other approaches to partiality.

The rules of (INF) appear in McGee (1991), though McGee has much more to say about the issue. For proof-theoretic investigation of similar systems, see Friedman and Sheard (1987). Much stronger systems are developed in Feferman (1991).

[8] I have tried to address these issues in my (2001).

In contrast, let us look at the response to Russell's paradox. There is, I believe, a standard response. It has two components, corresponding to the two components of the naive theory. First, the formal theory is replaced. There are a few candidates to replace it, but for illustration, let us take the Bernays-Gödel theory of sets and classes *BGC*. Second, the philosophical gloss on the naive theory is replaced by an account of the domains of sets and classes. Again there are a few competitors, but for argument's sake, let us assume some version of the iterative conception of set, together with the idea that (proper) classes are the extensions of predicates of sets.[9] Let us call the combination of these the *standard theory*. Its components are by no means beyond controversy. The iterative conception of set is still a matter of philosophical investigation and dispute. The continuum problem looms large as a difficulty of both components. Nonetheless, both components are standard in that they are to be found in introductory set theory texts, and they enjoy reasonably wide endorsement. Let us take the standard theory for granted.

The standard theory provides the *standard solution* to Russell's paradox. (COMP) has been dropped, and the revised theory *BGC* is presumably consistent. The question that needs to be asked, given the remarkable similarity between the Liar paradox and Russell's paradox, is if we can do the same for the standard solution to Russell's paradox as we did for the response to the Liar. Can we re-run Russell's paradox to get a strengthened version of it that poses a problem for this revised theory of sets?

The answer is that we cannot. First, recall how the standard solution resolves Russell's paradox. According to it, the Russell class R is simply not a set. There is, according to the Bernays-Gödel theory, a proper class R, which we may think of as the extension of the predicate $\ulcorner \neg x \in$

[9] *BGC* is a two-sorted theory, with variables $\ulcorner x \urcorner$ for sets and $\ulcorner X \urcorner$ for classes. The crucial axioms are restricted class comprehension $\ulcorner \forall X_1, \ldots, \forall X_n \exists Y (Y = \{x \mid \phi(x, X_1, \ldots, X_n)\}) \urcorner$ where only set variables are quantified in $\ulcorner \phi \urcorner$, and axioms that say every set is a class and if $X \in Y$, then X is set. *BGC* also has the usual pairing, infinity, union, powerset, replacement, foundation, and choice axioms, as well as a class form of extensionality. (See Jech, 1978.)

For those unfamiliar with the iterative conception of set, it is roughly the idea that the sets are build up in stages. The process starts with the empty set $\emptyset$, then forms all the sets that can be formed out of those, i.e. $\emptyset$ and $\{\emptyset\}$, then all sets that can be formed out of those, i.e. $\emptyset$, $\{\emptyset\}$, and $\{\emptyset, \{\emptyset\}\}$, and so on. (For more thorough discussion, see Boolos 1971, 1989.)

BGC is convenient for this discussion because it talks about classes explicitly, which will be useful when we return to Russell's paradox. However, everything I say could be expressed perfectly well if we chose a formal theory like *ZFC* that describes only the domain of sets. We can always, on the informal side, introduce (predicative) classes as the extensions of predicates of sets.

$x^\urcorner$ where $\ulcorner x \urcorner$ ranges over sets. From the axiom of foundation, in fact, we know that R is coextensive with the class V of all sets. Though there is a class R, it is not a set. Only sets are members of classes. Indeed, only set terms can occur on the left of the membership sign $\ulcorner \in \urcorner$, so we cannot even ask if $R \in R$ or $\neg R \in R$.

Pursuing the parallel between the paradoxes, we might attempt to reinstate Russell's paradox as we did the Strengthened Liar. We get the Strengthened Liar by noting that we still have the Liar sentence $\ulcorner \lambda \urcorner$, and asking what the theory in question tells us about its truth status. Similarly, we still have the Russell *predicate* $\ulcorner \neg x \in x \urcorner$. In parallel with asking about the truth status of $\ulcorner \lambda \urcorner$, we might ask what falls in the extension of this predicate. In particular, we might ask if the object R falls in its extension. With the Liar, we got the answer that $\ulcorner \lambda \urcorner$ is not true. Likewise, here we get the answer that R does not fall within the extension of the predicate $\ulcorner \neg x \in x \urcorner$, because it is not a set. With the Liar, this led back to paradox, as we seemed to have reached exactly the conclusion $\neg Tr(\ulcorner \lambda \urcorner)$. But here the parallel ends. There is no such problem with R. There would have been, if the predicate $\ulcorner \neg x \in x \urcorner$ said that x does not fall in the extension of x. If so, we would be forced to conclude $R \in R$, leading to paradox. But that is not what the predicate says at all. Rather, it says that the set x is not a member of the set x. We know that R is not a set, whereas the extension of $\ulcorner \neg x \in x \urcorner$ is a collection of sets, so R is not among them. This does not produce any paradox. The invitation to conclude that as R does not fall within the extension, then it does after all, is simply a confusion of sets with classes, and of set membership with falling within a class. The extension of the Russell Predicate is determined only by the facts about set membership.

There is no strengthened version of Russell's paradox for the standard theory, while there is a Strengthened Liar paradox for pure minimalism with P. There is thus a crucial difference between our responses to the two paradoxes. Though the paradoxes are basically the same, the response for truth is unstable. It is vulnerable to continued problems from the Liar. In contrast, the response for sets is stable. No problem from Russell's paradox remains. The question we need to consider is why. Answering this question will help us to see what a viable theory of truth must look like. We will then be able to consider whether minimalism can provide one.

4. Stability and Divisiveness

Recall the point from Section (1) that both naive set theory and the pure minimalist theory of truth are in an important sense not about anything. The standard set theory is entirely different. It is genuinely about something: sets and classes. This is so in two ways. First of all, the formal theory *BGC* itself makes claims about the scope and nature of the domain of sets that are true specifically of it, and do not hold of other domains. It provides some existence principles, and some generation principles that show how sets are generated from other sets. The iterative conception of set works with the formal theory, to help make clear what the intended interpretation of the theory is. Together, they go some way towards describing the domains of sets and classes. Of course, they have some well-know failings. They do not by any means complete the task as they stand. But anyone who understands the two components of the standard solution can reasonably claim to understand something of what sets are and something of how they behave; understand well enough, at least, to understand the difference between sets and classes.

This is crucial to the stability of the standard solution to Russell's paradox. The formal and philosophical components of the theory come together to allow us to conclude that the Russell class R is not a set. Once this distinction is in place, we can rely on it to decline the invitation to draw paradoxical conclusions. Once we see the difference between set membership and falling within a class, and understand the Russell predicate as a predicate of sets, the invitation may be seen clearly to be a gross mistake. We would like to make the same sort of reply to the Strengthened Liar. The question is what we need from the theory of truth to be able to do so.

In describing the domains of sets and classes, the standard theory of sets behaves as we expect of most theories. Most theories some way or another divide off their subject-matter from the rest of the world. It is the correctness of the description of the subject-matter provided by the theory that makes the theory true. Let us label this feature *divisiveness*.

Theories may be divisive in different ways and to different degrees. Perhaps the most striking case is the second-order theory of arithmetic. In this case, the formal theory itself fully determines its domain of application, by being categorical. Few theories live up to this rather demanding standard. Our standard set theory certainly does not. But together with the philosophical explanation of its intended interpretation—the iterative conception—it does provide a substantial division between sets and classes and other things, even if the precise extend of the domain of sets remains elusive. The standard theory is sufficiently divisive to at least partially describe the domain it is about, to which it is responsible for its own truth. Perhaps most importantly for our purposes, it is divisive enough to draw some important conclusions about what does *not* fall within the domain of sets.

We expect empirical theories to be divisive to roughly this degree as well. An example much like the case of set theory is to be had from quantum mechanics. My friends in physics assure me that the domain of application of this theory is phenomena of very small scale. Just what is small scale is explained in part by the more informal gloss given to the theory, but also in part by the value of Planck's constant. More generally, it is no surprise that any decent theory should describe whatever it is about well enough to give some indication of what that domain is. Such an indication had better enable us to conclude, at least in obvious cases, that something is not in the domain. For the most part, any decent theory should be divisive.

What is striking about both naive set theory and pure minimalism is that they are as non-divisive as can be. Both are so by design: it is a reflection of philosophical commitments of both. This is a consequence of the point of Section (1) that neither theory is properly about anything. Insofar as naive set theory is supposed to be a matter of logic, it is only about anything in being about absolutely everything. There is no particular domain that it describes. Likewise pure minimalism precisely attempts not to describe any domain. The range of instances of (T) or (INF) is limited by syntax; but not because that is the limit of the domain these principles describe, but rather only because that is the limit of what can be written down. Pure minimalism still provides no real divisive content. It has no principles that reveal the nature of truth or the things to which truth applies—no principles that explain the nature of truth bearers and demarcate their domain. Pure minimalism holds there is no such thing as a nature of truth or truth bearers!

To get a sense of how non-divisive even our revised pure minimalism is, consider again the theory P. What does P tell us about the domain of truths or truth bearers? It does prove some facts about truth, like $P \vdash Tr(\ulcorner 1+1=2 \urcorner)$ (assuming that the theory is based on arithmetic). It even makes some existence claims, as we know $P \vdash \exists x Tr(x)$. But the theory is still as minimally divisive as can be. When the theory does prove something, it is only because something else having nothing to do with truth—nothing to do with its subject-matter—does most of the work. Once something else about the theory proves, for instance, $\ulcorner 1+1=2 \urcorner$, then the theory is able to deduce $\ulcorner Tr(\ulcorner 1+1=2 \urcorner) \urcorner$. From there, it can perform an existential generalization to get $\ulcorner \exists x Tr(x) \urcorner$. But the principles governing truth that are the heart of the theory of truth only play a role in deducing these facts in the step from $\ulcorner 1+1=2 \urcorner$ to $\ulcorner Tr(\ulcorner 1+1=2 \urcorner) \urcorner$. The rest is a completely independent matter of arithmetic or logic. The theory can determine that something is the case, and then add that it is a truth, and then extract some logical consequences from this fact. But it cannot say anything about truths, the purported objects the theory is describing, more directly. None of the principles of truth in P by themselves make any substantial claims about truths in general, but only relate truth to specific sentences whose correctness has been independently decided. P states no general principles which can help us to understand the extent of the domain of objects to which truth applies. It is only divisive where something unrelated to truth makes it so. This is just as the philosophical principles of pure minimalism would have it, so the situation does not change when we supplement the formal theory with these principles.

To compare this situation with that of set theory, imagine a theory of sets more like our partial theory of truth P. It would have some principles which, once you concluded something else, could be used to conclude that some set exists, or has certain members. As a result, the theory could generate a list of statements that say that something is a member of something else ($\ulcorner a \in b \urcorner$), or not a member of something else ($\ulcorner a \notin b \urcorner$); but the only generalizations it could make would be those that followed from elements of the list by logic (by analogy with $P \vdash \exists x Tr(x)$).

Unlike the standard theory, the theory we are now imagining is not divisive enough to enable us to reply to the attempt at a strengthened Russell's paradox. If we found a pair of objects a and b such that we determined somehow that the theory could not have on the list $\ulcorner a \in b \urcorner$, we would be able to conclude only that, as far as the theory tell us, $a \notin b$. We would not be able to draw any more subtle conclusions. Crucially, we would not be able to draw the conclusion that $\ulcorner a \in b \urcorner$ is not on the list because b is outside of the range of objects the theory is attempting to describe, or because a is the kind of object that cannot be a member of anything. Hence, if were were to observe $\ulcorner R \in R \urcorner$ is not on the list, we would indeed fall into a strengthened Russell's paradox.

A theory like this is not divisive enough to avoid the strengthened paradox, just as pure minimalism is not. When we encounter $\ulcorner \lambda \urcorner$, pure minimalism, even modified by P, allows us nothing to say except that according to P, $\ulcorner \lambda \urcorner$ is not true. We cannot go on to make any substantial claim about why this is so. We can cannot observe that it is so because $\ulcorner \lambda \urcorner$ falls outside the domain of the theory, or falls under a special category within the theory, as we say with the Russell class on the standard set theory; nor can we say anything else about why.

In Section (1), I pointed out that pure minimalism, like naive set theory, is in a sense not about anything. We have seen in this section that this makes pure minimalism highly non-divisive. Failing to be reasonably divisive renders pure minimalism unable to respond to the Strengthened Liar, even when modified to use a theory like P. We have seen that a sufficiently divisive theory, like the standard set theory, easily dismisses the attempt at a strengthened Russell's paradox, even though the two paradoxes are themselves virtually alike. The moral is that to have any prayer of avoiding the Strengthened Liar, we must look for a more divisive theory of truth.

5. A Divisive Theory of Truth?

What would a more divisive theory of truth look like? One that is more like standard set theory in describing a domain of objects? As is often pointed out, it is too much to ask of a theory to characterize the domain of all truths. This would be impossibly demanding, as it would make the theory the complete theory of absolutely everything; containing all sorts of facts about all sorts of subjects.[10] But it is reasonable to ask the theory to be divisive about *truth bearers*. This could make the theory divisive about truth in the right way, as delineating a domain of truth bearers appropriately delineates the range of application of the truth predicate.

To indicate what such a divisive theory might look like, we should return once more to the debate over deflationism. Deflationism is often contrasted with the idea of a correspondence theory of truth. Outside of its classical form, such as in works of Russell, it is notoriously difficult to state clearly what the correspondence theory of truth is. Nonetheless, there is a core idea behind talk of correspondence which points towards appropriately divisive theories.

The core idea is that truth bearers are *representational*. They describe the world as being some way, and are true if the worlds is that way. There is some leeway in just how we characterize truth bearers along these lines. We might say that truth bearers are propositions, where propositions are objects that encapsulate collections of truth conditions (to borrow a phrase from Hartry Field). Truth then obtains

[10] There is a significant question of whether there is really a coherent notion of absolutely all truths. Some reasons to be skeptical may be found in Grim (1991). Related issues are discussed in Parsons (1974a) and my (MS). However, my worry here is much more pedestrian. It is already too much to ask of the theory of truth to contain all our current knowledge, whether or not a single complete theory of absolutely everything makes sense.

when the actual circumstance is among a proposition's truth conditions. On a more extensional approach, we might say that truth bearers are interpreted sentences, where the interpretations provide truth values for sentences (in contexts, where appropriate). A similar theory could be given based on utterances. The idea of correspondence comes up naturally in explaining how a sentence or utterance could wind up being a representation or expressing a proposition. This is naturally to be explained by some sort of appeal to reference, for instance. Regardless of exactly how it is explained, so long as being representational is itself construed as a substantial property, it implicates substantial properties of being a truth bearer and of truth.[11]

A theory along these lines could wind up being sufficiently divisive. An account of how truth bearers are representational could provide a suitably divisive picture of the domain of truth bearers, so long as it is able to make a principled distinction between sentences that genuinely provide truth bearers—express propositions or are otherwise representational—and sentences that may look like they provide truth bearers but do not. From this, we could distinguish two ways of failing to be true. One is to be a truth bearer that is not true, and the other is to fail to be a truth bearer at all. The partial theory P attempted to implement a distinction like this, by making the truth predicate partial. But in lacking a divisive theory of truth bearers, it failed to do so in way that is sufficiently robust to resist the Strengthened Liar.

If we had a sufficiently robust, sufficiently divisive theory of truth bearers, we could begin to resist the Strengthened Liar. This claim needs to be made with some care. It is well-known that appealing to propositions or truth-value gaps does *not* by itself suffice to solve the Strengthened Liar. My point is rather that a divisive theory of truth bearers is required to even begin to make progress towards addressing it. If we had such a theory, we could go this far in responding to the Strengthened Liar: when we conclude that $\ulcorner\lambda\urcorner$ is not true, we could come to two different conclusions. Either $\ulcorner\lambda\urcorner$ is (or expresses) a truth

[11] I should note that in asking for a theory which is divisive about truth bearers, I am not raising the classic question of what the *primary* bearers of truth are. Rather, the issue here is one of distinguishing truth bearers, be they primary or otherwise, from non-truth bearers in a sufficiently divisive way.

I should also note that some of the ideas for representational accounts of truth bearers I have canvassed rely upon a *relational* notion of truth or truth conditions. This again raises a classic question about the nature of truth; though again, one I think is best set aside for this discussion. As the extensional approach shows, a relational account of truth may allow for an ontologically weak account of truth bearers. This may still provide a theory which is divisive about truth bearers, as it will distinguish sentences which stand in the right relations to be truth bearers from those which do not.

bearer, but is not true, or $\ulcorner\lambda\urcorner$ is not (or does not express) a truth bearer. The latter is very much analogous to concluding that the Russell class R is not a set. If we could reach this conclusion, we could resist the invitation to infer that we have concluded λ, just as when we say that R is not in the extension of the Russell predicate, we have not concluded $R \in R$. It is in not providing a principled way to draw this sort of distinction that P and pure minimalism leave us no where to turn to avoid the Strengthened Liar.

Drawing a stable distinction between different ways of failing to be true is the first step in solving the Strengthened Liar. It is not the last. Crucially, we still need to explain what sense is to be made of the conclusion that $\ulcorner\lambda\urcorner$ is not true, even if this is because it is not a truth bearer. Many have taken this to require that we invoke a hierarchy of truth predicates. My own preference is for an approach relying more heavily on ideas about context dependence. I shall not advocate for any particular approach here. I only claim that we need our theory of truth to draw the distinction between different ways of not being true to proceed at all. Drawing it requires being divisive about the domain of truth bearers, in just the way the standard theory is about sets, and in just the way the pure minimalist approach is not. Just as having *some* set/class distinction is not by itself enough to solve Russell's paradox, having some characterization of the domain of truth bearers is not by itself enough to solve the Strengthened Liar; but it is a necessary precondition.

I suggested a moment ago that the required divisiveness might be found in a correspondence-based theory of truth. Unlike the minimalist approach, this view maintains that there is some underlying nature to the property of truth, to be found in the ideas of representation and the world fitting a representation. If this yields substantial principles about what makes a well-formed sentence or utterance a genuine truth bearer, it could provide just the divisiveness we need to begin building a stable response to the Liar. There have been many attempts over the years to articulate the ideas of correspondence and of representation. Rather than rehearse them here, I shall restrict myself to pointing out that many of the leading ways of dealing with the Strengthened Liar rely on them.

Most approaches to the Strengthened Liar ultimately rely on some sort of hierarchy. To see how this relates to correspondence and representation, we should start with the idea of *grounding* developed by Kripke (1975).[12] Sentences are divided into two classes: grounded and ungrounded. Grounded sentences are naturally assigned the values true

[12] The term first appears in Herzberger (1970), which also makes some comparisons with paradoxes in set theory. Kripke's work provides an extensive development

or false, while ungrounded ones are not. An account of grounding can thus be used to provides a divisive account of truth bearers. The account Kripke gives of grounding is, informally speaking, one of starting by describing the world in non-semantic terms, and then building up successively more complex descriptions involving semantic expressions. This is naturally taken to provide just the sort of account of representation to which I alluded a moment ago. We start with the idea of reference for non-semantic terms, and on the basis of it assign truth values to sentences involving only these terms. We then progressively build up truth assignments for sentences containing semantic terms—namely the truth predicate. As is well-known, the presence of self-reference and other semantic complexity makes this process transfinite, and it assigns truth values to some but not all sentences. Those that are assigned a truth value—either true or false—by this process are the grounded ones. The picture that emerges is one of truth bearers being those sentences that ultimately describe the world through this iterative process. We may well say that these are the sentences that express propositions, or are representational.

It is well-known that in the face of the Strengthened Liar, Kripke ultimately appeals to "the ghost of the Tarski hierarchy" (1975, p. 80). This is a hierarchy based on grounding rather than on syntax. At the first level is the truth predicate produced by the process of generating grounded sentences just described. From the first level, we can ascend to the next level in the hierarchy by reflecting on the entire process, and noting that on the basis of it certain sentences are not true. On this view, the basis for the hierarchy is the idea of grounding. Grounding generates the first-level truths and falsehoods, and then provides the material for the construction of the next level through some sort of reflection on the process, and so on.[13]

With this in mind, we should consider the more traditional hierarchical approach stemming from Tarski (1935), which imposes a hierarchy of indexed truth predicates and syntactic restrictions on how they may appear in a sentence. We can see the syntactic requirements Tarski imposes as requiring an explicit syntactic representation of much the same kind of grounding process as Kripke describes (though in Tarski's

of the idea. Technically, in Kripke's framework, grounded sentences are those that are true or false in the least fixed point, and the informal sketch I give below echoes this idea. However, it would not change the basic points I make about grounding if one were to offer some reason for starting with a more extensive assignment of truth values and generating a larger fixed point.

[13] The idea of such reflection is fundamental to the approach of Parsons (1974a). Both Parsons and Burge (1979) pursue these sorts of ideas with much more explicit attention to the role of context.

work, lacking the transfinite and level-merging aspects). Hence, we can see Tarski's approach as based on the same kinds of ideas about representation or correspondence as Kripke's. It is hotly debated just how much of a deflationist Tarski is, and I do not want to take a stand on Tarski interpretation here so much as to point out how a natural understanding of his theory draws on correspondence ideas. Moreover, it is important to note that irrespective of this, his theory is divisive in one vital sense. Surface well-formedness is not sufficient to make a sentence a truth bearer. A sentence must also be, we might say, *well-indexed*: its truth predicates must be indexed so as to make it a genuine sentence of the Tarski hierarchy of languages. Insofar as a characterization of this property must be part of the full theory of truth, we have a divisive principle. I have merely suggested that one source of such a principle is in correspondence-based ideas.

One other important example is the theory of Barwise and Etchemendy (1987). Their work is based on a situation theory which is explicitly modeled on ideas about the correspondence theory of truth as seen by Austin (1950).

I am not here advocating a particular solution to the Strengthened Liar; rather, I am claiming that a theory which can provide one must be divisive about truth bearers. Moreover, I claim, the general outlook of a representational approach to truth—embodying some aspects of the idea of correspondence—provides a basis for the development of such a divisive theory. We have seen how some of the leading theories that have been developed to solve the Strengthened Liar rely on this outlook.

6. Deflationism Revisited

I have stressed the parallel between pure minimalism and naive set theory on the one hand, and the divisiveness of standard set theory and a theory of truth able to address the Strengthened Liar on the other. In the face the paradoxes, a viable theory of truth, like a viable theory of sets, must be sufficiently divisive. I have noted that the general idea of the correspondence approach to truth provides a basis for such divisiveness. The question remains whether a deflationist approach to truth could do so. I shall argue in this section that a minimalist approach cannot. I shall then briefly raise the question of whether any other sort of deflationism could.

Pure minimalism, as I have described it, is clearly not able to provide a theory which is divisive on truth bearers. As we saw in the discussion of Section (4), the move from M to P does not help matters. This is striking, as P was designed to avoid the Liar. Its failure shows that just weakening the formal theory does not make the view divisive. And clearly, nothing in the philosophical component can help. The position there is precisely that the theory is not a theory about any domain, so it cannot be divisive.

There are a number of positions that depart from pure minimalism in ways that do not seem to affect the issue of divisiveness about truth bearers. Some views that hold to the basic principles of minimalism take the truth predicate to apply to propositions. However, these views remain minimalist by saying that for each well-formed declarative sentences, there is a proposition expressed by it. There is no more an underlying fact about what makes a sentence express a proposition than there is an underlying nature of truth. Hence, these theories are no more divisive than pure minimalism on truth bearers.[14]

A more significant departure from pure minimalism on the issue of truth bearers is derived from Wright (1992). He proposes a criterion for the truth aptness of sentences based on "surface constraints of syntax and discipline" (1992, p. 35). Truth bearers are then the truth-apt sentences. The basic idea of "discipline" is that of norms of *use* which are typical of assertion, including the appropriateness of embedding in antecedents of conditionals, for instance. This might allow for a minimalist theory to be a little more divisive than pure minimalism.[15]

However, I do not see how anything along the lines of surface constraints of syntax and norms of use can provide a criterion that will help in the face of the Liar. The problem there is precisely that we do seem to have a sentence that meets the constrains of syntax and of discipline as well, but cannot be a truth bearer. Compare this once again with the set theory case. Proper classes bear all the surface marks of sets. They have members, are extensional, and so on. By these lights, the Russell class is as good a set as any other, and hence the paradox. What avoids the paradox is a much more substantial ontological distinction, between set-like objects that are sets, and those that are not. In the truth case, a criterion like Wright's still seeks to make anything that looks like a truth bearer be a truth bearer. This is the force of the constraints of syntax and discipline being *surface constraints*. The addition of discipline refines the notion of looking like a truth bearer, but the problem with the Liar is that we have something that *does* look like a truth bearer by any of these standards, and we need a theory that provides a stable answer that nonetheless it is not one. We need a more divisive theory than the combination of syntax and discipline can give.

[14] This is quite close to the view held by Ayer (1946).

[15] Wright's use of the term 'minimalism' is in some ways slightly different from mine, but his criterion is still a natural one for a minimalist in my sense to employ.

The failure of this departure from pure minimalism to be divisive enough shows a general reason why anything that counts as minimalism will fail to be sufficiently divisive. A minimalist approach, however refined, is committed to there being nothing but overt surface properties that determine whether something is a truth bearer or not. For there to be anything else would *eo ipso* provide an 'underlying nature' of truth. This would make truth 'substantial' in just the way the minimalist says it is not. The Liar is such a problem in part because the Liar sentence really does appear to be a perfectly good truth bearer. It meets all the overt or surface criteria for being one. Because of this, minimalist solutions to the Liar tend to be unstable in the face of the Strengthened Liar. When we try to make the Liar sentence not a truth bearer, we then find reasons to reinstate it as a truth bearer after all. Overt or surface criteria give us no reason to which we might appeal to reject this conclusion. Hence, the kind of divisiveness we need to begin to address the Strengthened Liar is precisely one that is not based on overt surface features. This is not something any brand of minimalism can provide.

What about a theory that somehow strives to retain its minimalism but say less, or even nothing, about truth bearers? The discussion in Section (4) of why P fails to provide enough divisiveness shows why this cannot help. A theory like P is already formally silent on truth bearers, in that it has sentences for which it proves neither $\ulcorner Tr(\ulcorner\phi\urcorner)\urcorner$ nor $\ulcorner\neg Tr(\ulcorner\phi\urcorner)\urcorner$, yet it is vulnerable to the Strengthened Liar. It is so vulnerable because it does show how to predicate truth of some sentences, and then (appears to) show that the Liar sentence is not among them. It lacks an explanation of why this is not simply the conclusion $\neg Tr(\ulcorner\lambda\urcorner)$, and this in turn leads to paradox. When it comes to the puzzle of the Liar, having a theory that is simply silent on truth bearers does not give a sufficiently divisive theory. To the contrary, what we need is a theory that gives us a principled distinction between truth bearers and things that look on surface like truth bearers but are not, to which we can appeal in responding to the Strengthened Liar. If we somehow excise commitments on truth aptness from minimalism, we get no such thing. Instead, we get a theory which makes no substantial claims about truth bearers, but still allows the drawing of conclusions about applications of the truth predicate. This allows for the Strengthened Liar, and does not give us anything like the resources needed to resolve it.[16]

[16] The argument of Jackson et al. (1994) that minimalism about truth does not by itself lead to minimalism about truth aptness is thus not sufficient to save minimalism. They note that an instance of the T-schema, such as ''torture is wrong' is true iff torture is wrong', only confers truth conditions if the right-hand side of

I conclude that no minimalism can be divisive enough to respond to the Strengthened Liar, and so no minimalism is tenable. This raises the question of whether one can be a deflationist of some other sort and evade the problem I have raised for minimalism. The vague nature of the category of deflationism makes this question rather hard to answer, but the problem is one that is very difficult for most deflationist positions to overcome.

The most likely deflationist views to consider are those that, unlike minimalism, do include a substantial theory of propositions or of meaning, but still remain committed to a strong form of (T) and hold it to be analytic or necessary. Ramsey (1927) is sometimes seen this way, but I am inclined to side with Field (1986) in saying that Ramsey does not appear to be a genuine deflationist. Ramsey interpretation aside, the point is that any view that has a substantial account of propositions as encapsulating *truth* conditions can certainly have a strong version of (T), but is no more or less deflationary than its account of truth conditions.[17] The only way to pursue this line as a deflationist seems to be to offer a *non-truth-conditional* account of propositions, and yet hold that propositions are truth bearers.[18]

---

the biconditional is itself truth apt. This claim appears to be entirely correct, but it does not help with the issue at hand. At best, it provides a way for a minimalist to be silent about issues of truth aptness. As we have seen, this will not suffice. (I do think that their view of a plausible account of truth aptness leaves less room for minimalism about truth than they suggest. I shall return to this in Footnote (18), in the discussion of other forms of deflationism.)

The notion of truth aptness at issue here is that of the application of the truth predicate making sense. So long as application of the truth predicate is governed by enough discipline to make overt contradictions repugnant, we have to deal with paradoxes that may arise for it. This is the case whether or not the assertions in question wind up being truth apt in the sense relevant to, say, non-cognitivism. Thus, I believe that the issues primarily under consideration by Jackson et al., as well as those debated by Boghossian (1990), Wright (1992), and Soames (1999), are somewhat different than those being investigated here. (Surely no one ever really thought non-cognitivism could solve the Liar!)

[17] As was pointed out in Parsons (1974a), *any* view that contains a theory of propositions as truth conditions leads to a strong version (T). On such a view, the truth predicate is just the operation of evaluating an intension on the actual world. In the notation of intensional logic, we have something like $\ulcorner Tr(p) \leftrightarrow {}^{\vee}p \urcorner$. Coupled with the principle of intensional logic $\ulcorner {}^{\vee\wedge}\phi \leftrightarrow \phi \urcorner$, we get the T-schema in the form $\ulcorner Tr({}^{\vee\wedge}\phi) \leftrightarrow \phi \urcorner$. This has the status of a truth of mathematics, if not logic. The truth predicate is basically the operation of function evaluation. Even so, it is not clear that we should conclude that it is entirely trivial. Having self-applicative truth puts us in an untyped world, and experience with, say, the untyped λ-calculus shows that function application in such a setting is far from trivial.

[18] I suspect that Jackson et al. (1994) miss this point. They consider a view of truth aptness based in part on whether a sentence expresses the content of a belief, while belief states are said to be "designed to fit the way things are" (p. 297). This

Such a view might be divisive about propositions. If it holds that whether or not there is a proposition depends on more than the minimal criteria we have looked at, it could be.[19] But then such a view faces a challenge. Propositions are supposed to play the role of truth bearers. So, how do we account for the link between propositions, divisively and non-truth-conditionally described, and truth, described as deflationary? Can a theory be divisive about propositions, link truth with propositions by insisting that propositions are truth bearers, and yet be deflationist? I am inclined to doubt it can.[20]

References

Austin, J. L.: 1950, 'Truth'. *Aristotelian Society Supplementary Volume* **24**, 111–129. Reprinted in (Austin, 1961).

Austin, J. L.: 1961, *Philosophical Papers*. Oxford: Oxford University Press. Edited by J. O. Urmson and G. J. Warnock.

Ayer, A. J.: 1946, *Language, Truth and Logic*. London: Victor Gollancz, second edition.

Barwise, J. and J. Etchemendy: 1987, *The Liar*. Oxford: Oxford University Press.

Boghossian, P. A.: 1990, 'The Status of Content'. *Philosophical Review* **99**, 157–184.

appears already to brings with it a correspondence-like notion of truth for what sentences express, and so is not anything a deflationist can accept.

[19] Though this does seem to be a line worth considering, I am again unsure if anyone has actually held it or not. Horwich (1990) may have. He has a use-based theory of proposition individuation, and casts propositions as truth bearers. But, he also seems to have very minimal standards for proposition existence and expression. So, it is unclear to me if his account can be divisive about propositions. (See Field, 1992 for some discussion.)

Field (1986; 1994) certainly allows for substantial non-truth-conditional accounts of content, coupled with a deflationary account of truth specifically applied to utterances that the speaker understands. I am unsure if the notion of understanding he relies upon could be divisive enough, though. Which part of $\ulcorner \neg Tr(\ulcorner \lambda \urcorner) \urcorner$ don't you understand?

[20] One view that seems not to be vulnerable to the objections I have raised is that sketched in Soames (1984). The idea there is to construe truth as a property of *abstract* interpreted languages. Though he does not pursue the matter, much of what I have said about how a theory could be made sufficiently divisive could easily be carried over to this setting. As I understand it, the view is offered as deflationist in that truth is construed as not a metaphysically important notion, but more a piece of mathematics. (Soames (1999) takes a similar stance towards deflationism, and discusses the Liar explicitly. In response to the Strengthened Liar, he there appeals to a "Tarski-like hierarchy" (p. 181).) As Soames is well aware, this is very far indeed from minimalism, and even from most other positions that offer themselves as deflationist. I see no reason to argue over who gets the term 'deflationist', but I believe Soames' position is significantly different from the kind under discussion here.

Boolos, G.: 1971, 'The Iterative Conception of Set'. *Journal of Philosophy* **68**, 215–232. Reprinted in (Boolos, 1998).

Boolos, G.: 1989, 'Iteration Again'. *Philosophical Topics* **42**, 5–21. Reprinted in (Boolos, 1998).

Boolos, G.: 1993, *The Logic of Provability*. Cambridge: Cambridge University Press.

Boolos, G.: 1998, *Logic, Logic, and Logic*. Cambridge: Harvard University Press.

Burge, T.: 1979, 'Semantical Paradox'. *Journal of Philosophy* **76**, 169–198. Reprinted in (Martin, 1984).

Feferman, S.: 1984, 'Toward Useful Type-Free Theories, I'. *Journal of Symbolic Logic* **49**, 75–111. Reprinted in (Martin, 1984).

Feferman, S.: 1991, 'Reflecting on Incompleteness'. *Journal of Symbolic Logic* **56**, 1–49.

Field, H.: 1986, 'The Deflationary Conception of Truth'. In: G. MacDonald and C. Wright (eds.): *Fact, Science and Morality*. Oxford: Blackwell, pp. 55–117.

Field, H.: 1992, 'Critical Notice: Paul Horwich's *Truth*'. *Philosophy of Science* **59**, 321–330.

Field, H.: 1994, 'Deflationist Views of Meaning and Content'. *Mind* **103**, 249–284.

Friedman, H. and M. Sheard: 1987, 'An Axiomatic Approach to Self-Referential Truth'. *Annals of Pure and Applied Logic* **33**, 1–21.

Glanzberg, M.: 2001, 'The Liar in Context'. *Philosophical Studies* **103**, 217–251.

Glanzberg, M.: MS, 'Quantification and Realism'. Manuscript.

Grim, P.: 1991, *The Incomplete Universe*. Cambridge: MIT Press.

Herzberger, H. G.: 1970, 'Paradoxes of Grounding in Semantics'. *Journal of Philosophy* **67**, 145–167.

Horwich, P.: 1990, *Truth*. Oxford: Basil Blackwell.

Horwich, P.: 1994, 'The Essence of Expressivism'. *Analysis* **54**, 19–20.

Jackson, F., G. Oppy, and M. Smith: 1994, 'Minimalism and Truth Aptness'. *Mind* **103**, 287–302.

Jech, T.: 1978, *Set Theory*. New York: Academic Press.

Kripke, S.: 1975, 'Outline of a Theory of Truth'. *Journal of Philosophy* **72**, 690–716. Reprinted in (Martin, 1984).

Martin, R. L. (ed.): 1984, *Recent Essays on Truth and the Liar Paradox*. Oxford: Oxford University Press.

McGee, V.: 1991, *Truth, Vagueness, and Paradox*. Indianapolis: Hackett.

McGee, V.: 1992, 'Maximal Consistent Sets of Instances of Tarski's Schema (T)'. *Journal of Philosophical Logic* **21**, 235–241.

Parsons, C.: 1974a, 'The Liar Paradox'. *Journal of Philosophical Logic* **3**, 381–412. Reprinted in (Parsons, 1983).

Parsons, C.: 1974b, 'Sets and Classes'. *Noûs* **8**, 1–12. Reprinted in (Parsons, 1983).

Parsons, C.: 1983, *Mathematics in Philosophy*. Ithaca: Cornell University Press.

Quine, W. V.: 1986, *Philosophy of Logic*. Cambridge: Harvard University Press, second edition.

Ramsey, F. P.: 1926, 'The Foundations of Mathematics'. *Proceedings of the London Mathematical Society, Second Series* **25**, 338–384. Reprinted in (Ramsey, 1931).

Ramsey, F. P.: 1927, 'Facts and Propositions'. *Aristotelian Society Supplementary Volume* **7**, 153–170. Reprinted in (Ramsey, 1931).

Ramsey, F. P.: 1931, *The Foundations of Mathematics and Other Logical Essays*. London: Routledge and Kegan Paul.

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

Soames, S.: 1984, 'What Is a Theory of Truth?'. *Journal of Philosophy* **81**, 411–429.

Soames, S.: 1999, *Understanding Truth*. Oxford: Oxford University Press.

Tarski, A.: 1935, 'Der Wahrheitsbegriff in den formalizierten Sprachen'. *Studia Philosophica* **1**, 261–405. Translation by J. H. Woodger as "The Concept of Truth in Formalized Languages" in (Tarski, 1983).

Tarski, A.: 1983, *Logic, Semantics, Metamathematics*. Indianapolis: Hackett, second edition. Edited by J. Corcoran with translations by J. H. Woodger.

Wright, C.: 1992, *Truth and Objectivity*. Cambridge: Harvard University press.