Published in *Erkenntnis* **77** (1), pp.133-148, available at
www.springerlink.com, DOI 10.1007/s10670-011-9310-2.

ON USING INCONSISTENT EXPRESSIONS

Arvid Båve, Stockholm University

Abstract: The paper discusses the Inconsistency Theory of Truth (IT), the view that "true" is inconsistent in the sense that its meaning-constitutive principles include all instances of the truth-schema (T). It argues that (IT) entails that anyone using "true" in its ordinary sense is committed to all the (T)-instances and that any theory in which "true" is used in that sense entails the (T)-instances (which, given classical logic, entail contradictions). More specifically, I argue that theorists are committed to the meaning-constitutive principles of logical constants, relative to the interpretation they intend thereof (e.g., classical), and that theories containing logical constants entail those principles. Further, I argue, since there is no relevant difference from the case of "true", inconsistency theorists' uses of "true" commit them to the (T)-instances. Adherents of (IT) are recommended, as a consequence, to eschew the truth-predicate. I also criticise Matti Eklund's account of how the semantic value of "true" is determined, which can be taken as an attempt to show how "true" can be consistently used, despite being inconsistent.

Since Tarski (1944), the idea that the Liar paradox shows that ordinary language is defective has long remained in the shadow of the numerous attempts, including Tarski's own, to work out a consistent notion of truth. Recently, however, a view similar to Tarski's, the Inconsistency Theory of Truth (IT), has gained popularity among truth-theorists.[1] Though there are many different formulations of (IT), I will adopt Matti Eklund's terminology, on which (IT) is the view that the truth equivalence-schema

(T) That *p* is true iff *p*

(or " '*p*' is true iff *p*"—the difference will not concern us here) is a *meaning-constitutive principle* of "true". By this, I will mean that *competent speakers are disposed in virtue of their semantic competence with "true" to accept every instance of (T)*. In other words, the

competence conditions for "true" entail that a competent speaker will be disposed to accept these instances. I will use "(IT)" and related terms ("inconsistency theorists", etc.) in this narrow sense. Perhaps not all self-labelled "inconsistency theorists" would accept this claim, but that is not essential to this paper. Still, Patterson (2009) is committed to (IT) in my sense, since he takes our competence to consist in our "cognising" an inconsistent T-theory, including all the (T)-instances, and he takes "cognising" to entail a (defeasible) disposition to accept (2009: 419). Many authors take the "inconsistency theory" to be a truth-theoretic claim rather than one about competence and dispositions (cf. Azzouni (2003, 2007) and Ludwig (2002)). However, given how speakers' cognising or believing truth-theoretic semantic axioms is (typically) thought to play a part in explaining linguistic behaviour, it may well be that they, too, are ultimately committed to (IT) in my sense). Still, I will not press this point further, but rather focus on those theorists who explicitly accept (IT) in my narrow sense, in particular, Eklund.

The reason that (IT), as defined here, is taken to be an "inconsistency" view is of course that, together with various further allegedly meaning-constitutive principles (sentences or rules of inference), some of the instances of (T) entail a contradiction (which amounts to the Liar paradox). Assuming, furthermore, that our competence with negation and conjunction disposes us to *reject* contradictions, (IT) has the consequence that our language is inconsistent in the sense that competence with its expressions disposes us to accept principles that allow the inference of sentences that our semantic competence disposes us to reject. It is plausible that the notion of an inconsistent *expression* must ultimately be defined in terms of an inconsistent language, in the sense just explained. On that take, "true" is inconsistent in that adding it (by way of (T)) to a language with certain expressive devices already in place yields an inconsistent language in this sense.

The support for (IT) comes, firstly, from its attractive simplicity—we do not have to assume that our competence with "true" consists in some relation to the kind of complex and unnatural principles that would have to define "true" if it were consistent. Further, and even more importantly, (IT) is arguably the only consistent theory that can properly deal with the totality of linguistic intuition pertaining to "true", especially those relating to the Liar paradox. In particular, it explains why we find the steps in the derivation of the contradiction so intuitive, even to the extent that someone who does *not* find them intuitive thereby displays semantic ignorance (cf. Eklund (2002a: 260)).

(IT) should be distinguished from *dialethism*, which says that there are true contradictions. Dialethism fares as well as (IT) in explaining the intuitions involving "true", but arguably fares worse with accounting for our intuitions about contradictions. It is also committed to a non-classical logic, again in conflict with widespread intuitions. So although (IT) and dialethism agree about the meaning-constitutive principles of "true", inconsistency theorists typically reject dialethism, and thus intend, like most philosophers, to accept only consistent theories. Thus, (IT) is meant to be a consistent description of an inconsistent language.

The present paper raises the question whether this ambition to give a consistent theory can be met if the theory also *uses* "true", given that it is inconsistent in the sense just explained. My main thesis is that it cannot, i.e., that given (IT), *any theory in which "true" is used in its ordinary sense* (and which contains the further expressive devices required to derive a contradiction, like classical negation, etc.) *is inconsistent*. This matters, firstly, because almost all inconsistency theorists make essential use of "true" in their theories about "true".[2] Thus, if they also grant (IT) in my sense, their theories are all inconsistent. Secondly, it matters because if I am right, then (IT) has the rather radical-seeming commitment of *eschewing the term completely*, i.e., of disallowing any use of "true".

I will proceed by first stating an argument against Eklund's view about the semantic value of "true". This serves the general purpose of the paper because that view can be taken as a demonstration of how and why "true" can be consistently used like any ordinary, consistent expression (contrary to the main thesis of this paper). Section II then presents a more general case for the claim that using inconsistent expressions in a theory makes the theory inconsistent. In section III, finally, I discuss the costs of eschewing "true", and conclude that this commitment of (IT) is not as radical and costly as it may seem at first.

1. Eklund on the determination of the semantic value of "true"

Inconsistency theorists typically claim that the derivation involving a liar sentence to a contradiction is *unsound*; specifically, that the relevant instances of the truth-schema are *untrue*, despite being meaning-constitutive.[3] This may seem like a commitment for anyone who rejects dialethism. But a different response is possible, namely, to simply *eschew* "true", i.e., to refuse to use it, on the grounds that it is inconsistent.[4] Such a view of course does not instead take the Liar derivation to be *sound*, since this is inconsistent. Rather, one simply refuses to use "true" and all expressions defined in terms of it, such as "sound". (Of course, when I speak of using "true", I throughout mean using it *in its ordinary sense*. It is trivial that the word "true" can be consistently used if it is used in a sense other than the ordinary. And there is no indication that my opponent inconsistency theorists intend to use "true" in any such alternative sense.)

Now, Eklund does not merely claim the Liar derivation to be unsound; he develops an elaborate view within the truth-theoretic framework about "true" in order to show how the Liar derivation *can* be unsound, although its premises and inferences are all meaning-constitutive (of some expression). On his view, "true" should be treated the way David Lewis treats theoretical terms. To wit, Eklund thinks that the semantic value of "true" is that which

makes true *most* (but not necessarily all) of the meaning-constitutive principles for "true" (2002a: 264f.). Thus, "true" and related notions figure rather centrally in his account. It thus clashes in an obvious way with the view to be defended here. But there is a more important reason to focus here on this part of Eklund's account, namely, that it can plausibly be taken as an attempt to show how and why "true" can be consistently used, despite its inconsistency. I will therefore devote this section to trying to undermine this view, by arguing that there are differences between theoretical terms and "true", in view of which the view fails.[5]

According to Lewis (1970), the semantic value of a theoretical term, like "electron" is determined as that which *best satisfies* a certain set of principles involving the term. These principles are thus those that count towards determining the semantic value of the term. I shall call them the *determinant principles*. So, if these principles, in the case of "electron", are, "F_1(electrons)", "F_2(electrons)", ..., "F_n(electrons)", then the semantic value of "electron" is what best satisfies "$F_1(x)$", "$F_2(x)$", ..., "$F_n(x)$". Two obvious questions are, first, what the determinant principles for a given expression are, and, secondly, what is to count as "best" satisfying them. In general, a determinant principle should be *accepted* by (at least some distinguished subset of) competent speakers (perhaps the experts). A common idea is to take these principles to be all and only the meaning-constitutive principles for the expression. That an entity satisfies a set of principles better than another may be understood in terms of some kind of counting and weighting, but we need not here elaborate on this. Lewis later proposed (1984) that not only the satisfaction of principles, but the intrinsic nature of the entity, may count toward determining whether it is the semantic value of a term. To wit, he thought that if one entity is *more natural* than another, then that weighs in favour of the former being the semantic value of a term. This further condition is also important for Eklund's theory, but will not play any role in my argument.

Of course, a term may have a certain semantic value even if it does not satisfy *all* the determinant principles. If we didn't allow for that, we would face a difficult choice between, on the one hand, a view on which most of our theoretical terms may turn out to be empty, and, on the other hand, taking the sentences that count towards determining the semantic value to be so few and weak as to fail to determine a semantic value. In any case, since Lewis grants that a determinant principle can fail to be satisfied by a term's semantic value, his account can solve the puzzle of how a theoretical term can have a constant referent, although the determinant principles may be revised and taken back, and so forth, and, indeed, although many of them may be false. The solution accords well with our preconceptions about theory revision and the use and reference of theoretical terms.

Now, Eklund thinks that "true" and "electron" are alike in that their determinant principles may not all be satisfied by their respective semantic values. For "true", he thinks, the determinant principles are all and only the meaning-constitutive principles for "true", i.e., the instances of (T). However, since not all of these can be satisfied, on pain of contradiction, some instances must be untrue. Thus, truth itself—the semantic value of "true"—must be something that does not in fact satisfy every meaning-giving principle of "true".[6]

One might say that on Eklund's view, although "true" is inconsistent, truth is consistent. That "truth is consistent" can be understood as meaning simply that the most exhaustive true description of truth is consistent. This, furthermore, is a simple consequence of Eklund's rejection of dialethism. On this view, then, the meaning-constitutive principles of "true" do not perfectly capture the nature of truth. However, if dialethism is false, the semantic value of "true" must still be something that can be described (exhaustively, I presume) in a consistent way. Thus, Eklund thinks, there is a set of principles that are determinant for some other, *consistent* expression, which has the same semantic value as "true" (2002a: 268f.). Perhaps one of the many proposed truth-definitions in the literature that

are designed to "solve the Liar paradox" is such an expression. It will not, then, have the same meaning as "true", but it will better capture truth than do the meaning-constitutive principles for "true".

This account seems somewhat counter-intuitive. Granted, it is an unavoidable commitment of (IT) that there are meaning-constitutive principles for "true" that are not true, and this seems like an independently plausible claim. But it is quite another thing to say that such a principle can be untrue because the semantic value of "true" does not satisfy it. If that were the case, we should say that the principles in question—to wit, the pathological instances of (T)—are not merely untrue, but *false*. On this view, then, although there is such a thing as truth, it is not the way it has to be to make the meaning-constitutive principles of "true" come out true. This seems implausible: if "true" is inconsistent, it seems, we should rather say that there is no such thing as truth.

My main objection here, however, is not simply that Eklund's account is counterintuitive in this way. Rather, I will argue that although it is plausible that the semantic value of an expression can fail to satisfy a determinant principle, it is not plausible to claim that it can fail to satisfy the *meaning-constitutive* principles for the expression. And, since the totality of meaning-constitutive determinant principles for "true" is unsatsifiable (assuming (IT)), "true" cannot have a semantic value. Therefore, the view that a term's semantic value can fail to satisfy its determinant principles cannot be appealed to in order to show how "true" can be used like any ordinary, consistent expression. That is, the use of "true" cannot be justified by the claim that "true", like "electron", has a semantic value and can be useful although we may be unsure as to its exact nature, and although some of our beliefs involving the notion are false. For the two expressions are unlike in that the principles, if any, that fail to be satisfied by electrons are not meaning-constitutive for "electron".

I would like to make two further, minor points before proceeding with the argument. First, one might think that, if I am right and Eklund is wrong, then the pathological instances of (T) are merely *not true*, due to the lack of a semantic value, rather than false. However, this claim involves a use of "true", wherefore I will not make it. Rather, again, the alternative I propose is that "true" and its cognates be eschewed. Secondly, I want to note that my objection is neutral on what is taken as the determinant principles for "true", as long as they include the instances of (T). But it is uncontroversial that the meaning-constitutive principles for an expression will all be among the determinant principles, and adherents of (IT) are committed to the claim that the instances of (T) are among the meaning-constitutive principles for "true". Thus, the view that "true" has a semantic value cannot be saved by "adding" further determinant principles.

Now, in order to determine whether the semantic value of an expression must satisfy its meaning-constitutive principles, we should look at some unrelated case, in order to see what it seems plausible to say, quite generally, about determinant principles and meaning-constitutivity. Let us therefore look at an example in which there is no question as to what are the meaning-constitutive principles, namely, a stipulatively defined term. Suppose that a mathematician believes that there is a number satisfying conditions C_1, ..., C_n, and wants to speculate about the further properties of this number, and stipulates, for reasons of brevity, that τ is the number satisfying conditions C_1, ..., C_n. In this case, each sentence, "τ meets condition C_i" ($1 \leq i \leq n$) is meaning-constitutive of "τ", i.e., to be competent with "τ", one must be disposed to accept it (at least assuming that one is competent with the other expressions therein). Suppose, further, that it turns out that no number can satisfy these conditions, on pain of inconsistency. In such a case, we would *not* take the term to refer to what satisfies them "best", but rather hold that it failed to refer. For suppose someone discovered that a certain number satisfies the conditions "well", but not perfectly, and claimed

that τ is in fact *this* number. Surely, such a claim would be, as it were, semantically reproved. His statement would be taken to display negligence of the relevant linguistic convention. One would say, "Look, this is how τ is defined. You may if you wish redefine it, but you cannot say that, in fact, τ, as defined, does not satisfy C_i". Thus, the case of terms *defined* as having referents satisfying certain predicates is different from the case of theoretical terms, some of the determinant principles of which are merely are *accepted* or *believed*, but not taken as definitional.

It seems clear that "τ" lacks semantic value because its meaning-constitutive principles are unsatisfiable. The question is then whether "true" goes with "τ" or with "electron", some of whose determinant principles may fail to be satisfied by its semantic value. But it is important to note exactly what the difference is between "τ" and "electron". The difference is merely that "electron" has determinant principles that its semantic value may fail to satisfy, namely, the ones that are not meaning-constitutive. They are the *same* in the respect that their meaning-constitutive principles must be satisfied by their semantic values. So, with regard to the meaning-constitutivity of determinant principles, "true" is like "τ", rather than "electron", but the commonality consists merely in that all of their determinant principles are meaning-constitutive. Also, of course, the determinant principles for both are unsatisfiable. Since in the case of "τ", we do not hesitate to say that it, for this reason, has no semantic value, we should say the same of "true".

It may be thought that I have here mistakenly assumed that the determinant principles for "electron" are not meaning-constitutive. For, surely, someone who denied all of the supposedly non-meaning-constitutive principles for "electron" must be semantically incompetent. Here, it is important to mind a distinction between absolutely and relatively meaning-constitutive principles. It is reasonable that the "substantial" determinant principles for "electron" are such that accepting them *counts toward* semantic competence with

"electron". However, it should be clear that any one of them may still be denied consistently with competence (although perhaps some can be denied consistently with competence only by coming to accept or ceasing to accept certain other sentences). Competence with such expressions as "electron" is thus a matter of surpassing a threshold, i.e., a matter of degree, while, on (IT), this is not so for "true". This view of expressions like "electron" fits nicely with the intuition that some sentences containing them seem like border-line cases of meaning-constitutivity, like perhaps, "Electrons have negative charge". It is also rather generally plausible that expressions with an empirical content have mainly "relatively" meaning-constitutive principles. Possible exceptions will be rather trivial, like "Electrons are particles". If we grant Putnam's (1962) argument, even "Cats are animals" is not absolutely meaning-constitutive. In any case, then, "meaning-constitutive" should here be read as meaning *absolutely meaning-constitutive*, i.e., the notion defined in the beginning of this paper. Once this is clarified, this objection can be held to rest on an equivocation on "meaning-constitutive".

To make essentially the same point in a different way, consider a different way we might have introduced a term for a particle. Suppose that we *defined* a term, say "electrine", as meaning "things which satisfy conditions $F_1(x)$, $F_2(x)$, ..., $F_n(x)$", in which case all sentences "F_1(electrines)", "F_2(electrines)", ..., "F_n(electrines)" would be (absolutely) meaning-constitutive. In that case, if we found out that there is nothing that satisfies all these conditions, although there is something that satisfies a "good deal" of them, we would still say that "electrine" failed to refer. It is just that "electron" was not introduced, and is not used, in this way.

Before considering some objections to the above line of reasoning, I would like to present an objection very similar to that above, which targets Eklund's idea that "true" may express the same property as a consistently defined predicate. Two terms that are both

introduced by meaning-constitutive principles that make them inequivalent, cannot, it seems, express the same property, no matter how "similar". Define a predicate "$F_1(x)$" as equivalent with "$G_1(x)$ & ... & $G_n(x)$", for a very large n, and then define "$F_2(x)$" as equivalent with "$G_1(x)$ & ... & $G_{n+1}(x)$" so that "$F_1(x)$" and "$F_2(x)$" become inequivalent, if ever so "similar". Surely, they still express different properties, if any (though their extensions may of course be the same). This is the very common intuition expressed by sentences like, "The property of being *F* is distinct from the property of being *G*, although all and only *F*s are *G*s". The only motivation for denying this in this connection is that one of the terms would otherwise not express a property at all, but, of course, it would beg the question against the present argument to rule this out.

The first objection against my main argument I want to consider is that there might be other expressions whose determinant principles are both inconsistent and meaning-constitutive, but which we would still take to have a semantic value. I cannot think of any example I find ultimately persuasive, but I have noticed that a common reaction to the argument above is to mention Frege's inconsistent set theory laid out in his *Grundlagen* (1884/1974). Surely, it is said, Frege was still talking about *sets*, although his axioms governing the notion were meaning-constitutive. My reply is that it is not clear whether Frege's axioms were intended as absolutely meaning-constitutive. If they were, then the notion he defined (which may or may not coincide with the pretheoretical notion) is inconsistent, in which case it must be treated like "true"—as lacking semantic value. If, on the other hand, his axioms are *not* absolutely meaning-constitutive, then his term "set" is *not* an example of the kind of expression the objection speaks of, wherefore the objection fails.

Secondly, it may be thought that I have unduly neglected the fact that "τ" has been introduced by means of an explicit definition, whereas "true" has been introduced into our language by way of an implicit linguistic convention. The corresponding principles have thus

become meaning-constitutive by different routes, as it were. However, I do not see why this difference should entail a difference concerning the conditions upon having a semantic value (this difference will be further discussed below). What seems to be relevant is meaning-constitutivity, not whether the meaning-constitutive principles have been made explicit.

A third and important objection is that without an account like Eklund's, we would flout the Principle of Charity (cf. Eklund (2002: 263f.)). For present purposes, the principle is that the semantics for an expression must make *true* as many as possible of the sentences involving the expression of some particular kind (say, the accepted ones). I will grant, for the sake of argument, that the Principle of Charity requires the kind of account Eklund proposes. Still, there are good reasons to reject such a principle, if taken as holding for all expressions (cf. McGinn (1977) and Daly and Liggins (2010)). A less contentious variant of charity takes it to require rather that the semantics make speakers come out as accepting sentences in maximal accordance with their meanings. Since the semantics will assign to expressions various meaning-constitutive principles (absolute and relative) and perhaps defeasible verifying perceptions for expressions with empirical contents, it is a form of conceptual role semantics. The satisfaction of the weakened principle of charity for such a semantics is a relatively straightforward matter. Basically, a semantic hypothesis is regarded as justified to the extent that actual speakers fit the description of what the semantics takes as constitutive, or co-varying, with semantic competence. If a semantic hypothesis scores well in this regard, it will *ipso facto* be successful in predicting speakers' linguistic behaviours and intuitions. It is unclear why we should need anything over and above this. Of course, adherents of truth-theoretic semantics often say that the phenomenon of semantic compositionality can only be done justice by their type of semantics. We cannot enter into this debate here, but let us simply note that this is a contested claim, and one that its defenders have tended to assume or

conjecture, rather than argue (Davidson's original statement of the claim in "Truth and Meaning" (1967) being a case in point).

There are even more reasons for adherents of (IT) specifically to adopt a broadly conceptual role semantics, rather than a truth-theoretic semantics. For the most important advantage of (IT) is its promise to account in a simple way for all linguistic behaviour and all intuitions concerning truth and the related paradoxes. A standard truth-theoretic semantic theory must, on pain of contradiction, distinguish the pathological instances of (T) from the well-behaved ones. But that means precisely that this theory could *not* have this predictive property that (IT) has in virtue of *not* distinguishing them. So what, one may wonder, is the truth-theoretic semantics good for, if it does not connect to semantic competence and linguistic behaviour? Thus, it is a commitment of (IT) to deny that standard truth-theoretic semantics plays an explanatory role with respect to semantic competence and linguistic behaviour. (Ludwig, of course, does take his Davidsonian truth-theoretic semantics for "true" to play such a role. But since he accepts (IT), this semantics is precisely *in*consistent. His resolution of this conflict is to refrain from accepting its axioms, and merely say that speakers acceptance of them constitute their semantic competence. So his theory is really more similar to conceptual role semantics than standard truth-theoretic semantics. For it ultimately trades in the notion of speakers' accepting/believing/cognising things involving truth-theoretic properties, rather than in expressions just *having* them (cf. Horwich (2008)).

This argument relates to the original objection from the truth-maximising Principle of Charity as follows. It is commonly held that the way to empirically verify a semantic hypothesis within the truth-theoretic semantic framework goes by showing that speakers come out maximally truth-speaking, given the semantic theory. But not only are there reasons to be sceptical about truth-theoretic semantics in general; an inconsistency theorist is committed to denying that such a semantic hypothesis (since it would have to be consistent)

could play any explanatory role in the explanation of our linguistic intuitions and behaviour concerning "true". Further, since the arguments in favour of truth-maximising that are independent of truth-theoretic semantics can well be questioned as well, the appeal to charity can well be resisted by an inconsistency theorist.

Finally, I want to note an important difference between Lewis's account of theoretical terms and Eklund's account of "true". Lewis's account is both intuitive and solves a universally recognised problem. The account Eklund needs is not intuitive and also only solves the problem of holding on to (IT) while giving a truth-theoretic semantics and granting a property of truth. Although (IT) has good support (in my opinion), this support counts just as much in favour of a theory like my own, which eschews "true". Thus, the plausibility of (IT) does not extend to Eklund's view of the semantic value of "true", and neither does the plausibility of Lewis's account of theoretical terms.

2. Why using inconsistent expressions is inconsistent

In this section, I will present a direct case for the claim that an inconsistent expression cannot be consistently used. More precisely, I will argue that a theorist who uses a given expression and intends it to be understood in a certain sense is committed to accepting the meaning-constitutive principles for the expression relative to that sense. Of course, one can deny this claim simply by rejecting the idea that the meaning of the expression in question is determined by meaning-constitutive principles, but this is clearly not open to inconsistency theorists. A consequence of theorists being committed in this way, further, is that any *theory* in which an expression is used and intended to be understood in a certain sense will entail the meaning-constitutive principles relative to that sense. Though correct, however, this claim might be misleading: what I mean is that the theory will *entail* these principles because it will *contain* them, if only implicitly (and every claim entails itself). This connection between a

theorist being committed to a meaning-constitutive principle and her theory entailing it might not be transparent, but I hope to justify it in due course.

One can of course use "true", and claim to be using it in its ordinary sense if one holds that its meaning-constitutive principles are consistent with the rest of one's theory. But this possibility is irrelevant, since it is not open to inconsistency theorists about truth. Perhaps there is a sense of "committed" in which any theorist using "true" in its ordinary sense will be committed to the totality of (T)-instances, assuming they are meaning-constitutive, whether she accepts this claim or not. However, this is also irrelevant to my main purpose here, which is to show that inconsistency theorists are committed to accepting these principles, if they use "true" in its ordinary sense.

I said I would provide a "direct" case for this claim. This is opposed to the case made in the previous section against Eklund's view. That criticism was "indirect" in the sense that it aimed to undercut a view which aims to show how it can be possible to consistently use "true", despite its inconsistency. Thus, it was merely an *argument against an argument for the claim that "true" can be consistently used*. Here, instead, I will provide an argument against the claim that "true" can be consistently used, and this argument is wholly independent of the foregoing criticism.

Of course, my own view is *not* that "true" cannot be consistently used because the relevant instances of (T) are true (or that inferences between their halves are truth-preserving), since that is in itself an inconsistent claim. Thus, I must argue that one can be committed to a claim in the relevant sense without this entailing that the claim is true. But this is quite innocuous, for there is clearly a sense in which speakers are committed to that which they assert, whether true or not. Of course, inconsistency theorists do not assert the problematic instances of (T). Still, I will argue that they are committed to them, if they use "true", in the same sense as that in which they would be committed to them if they asserted them.

What, then, is the difference between them and myself? Why am I not committed to these principles, and everyone else? The difference is precisely that they use "true", intend it to be understood in the ordinary sense, and hold that the meaning-constitutive principles for the word relative to that sense include all the (T)-instances. This entails that they are committed to the instances of (T), and that their theories in which "true" is used entails them. My main argument for this thesis begins with a simple comparison with a different case. Consider any theory in which logical constants occur. For any such theory, someone endorsing it must be able to specify what logic is to hold in it. Suppose our theorist decides that the logic is classical, i.e., that the constants are to be interpreted classically. Then, surely, she is committed to the classical theorems. Likewise, they are all entailed by the theory. These two claims seem inseparable, given the way this example is set up. Nevertheless, it seems to follow from the simple fact that the theorist, relative to whose intended interpretation we are taking the theory, has claimed that she is using the logical constants in their classical senses. Hence, if the meanings of these constants are determined by meaning-constitutive principles (which, as such, would have to allow the inference of all the classical theorems), then anyone who uses them and intends them to be understood classically is committed to their meaning-constitutive principles.

Note that we do not even have to assume that the meanings of logical constants *are* determined by meaning-constitutive principles (although it would be odd for an inconsistency theorist to think otherwise, since the meaning of "true" is so determined, on their view, and it is hard to see why there should be a difference between them). All we need to ensure is that if the meaning of an expression is determined by meaning-constitutive principles, then a theorist who uses it and intends it to be understood in the sense determined by those principles is committed to those same principles. And since, on (IT), the meaning of "true" is determined precisely by the totality of (T)-instances, an inconsistency theorist who uses "true" in its

ordinary sense is committed to these instances. Again, if she is committed in this sense, then it seems obvious that the theory in which the expression is used entails (indeed, contains) the principles.

We might say that for the kind of case described, the schema (T) figures as an implicit axiom schema in the theory using "true". Thus, its instances need not be explicitly endorsed for it to function as an axiom schema—commitment is sufficient. This claim can equally be attested by comparing with the case of logical constants. If a theorist proposes a theory and claims that the logic is meant to be classical, then the theory will contain principles allowing inference to the classical theorems as implicit axioms or inference rules. In any case, the theory will clearly *entail* these theorems, whether the principles codifying the inferences have been made explicit or not. And this is really all I mean by saying that they figure as implicit axiom schemata. The claim in virtue of which this holds, however, is that something *has* been made explicit, namely, the claim that the expression is to be understood in a certain sense. So, by parity of reasoning, if a theorist uses "true" and holds that it is used in its ordinary sense, then the theorist is committed to *its* meaning-constitutive principles, which, on (IT), include every instance of (T). Since these instances are for that reason contained in the theory, just like classical theorems, etc., are contained in any theory which is meant to be logically classical, the theory will entail a contradiction, assuming it also contains the further necessary expressive devices.

A natural reaction against this on the part of adherents of (IT) is simply to refuse the above reasoning, and insist that they are only committed to what they have explicitly stated, and that they have not so stated the inconsistent (T)-instances, but rather claimed them to be false. To assess this move, let us again compare with the case of logical constants. Suppose someone claims that the constants of her theory are to be understood classically, yet claims that a certain instance of a classical logical truth (i.e., a schema) is not true. Surely, this is not

a viable combination of claims. Rather, this theorist must either retract from the latter claim or give up the claim that the logical constants are to be understood as classical.

Similarly, suppose we have an objection against this theorist to the effect that her theory classically entails an absurdity. It would be very odd for her to defend the theory by saying that since she has not explicitly stated the instance of the relevant axiom schema or inference rule at some time or place, she is not committed to the absurd claim. Even if she were to explicitly reject the relevant sentence or inference, what is classically entailed by her theory is simply entailed, given her claim that the constants are classical. The rejection of a sentence that is a classical logical truth would just be inconsistent, and not something that somehow annulled the commitment to it.

I take these claims about logical constants to be obvious. But it seems equally obvious that there is no reason to regard "true" as any different from logical constants in the relevant respects. But then we must conclude that, given (IT), "true" cannot be consistently used in its ordinary sense. That is, if someone claims that "true" is to be understood in the ordinary sense, and that the meaning-constitutive principles for the word relative to that sense include all of the (T)-instances, then they are theorems of her theory, just as with classical theorems. (Though I will not expand on this, one may also consider the analogous situation where someone claims that the word "bachelor", as she uses it, is to be understood as having its ordinary meaning. Surely, any claim of hers of the form "*x* is a bachelor" will then entail "*x* is unmarried", whether she rejects this entailment or not.)

However, couldn't a defender of (IT) object that for all I have said, it might still be consistent to both hold that the logical constants of their theory are classical in the sense that the expressions' respective meaning-constitutive principles are classical, and yet deny some of these principles? Yes, and this would be true, but in this sense of "taking the constants of one's theory to be classical", the problem is not that these claims would be logically

inconsistent, but rather that it is irrational both to take the constants of one's theory to be stipulated (in the relevant sense) by classical principles and denying some of them, rather than taking the constants to be stipulated by whatever weaker principles one takes to be true. This is not to say that the claim that a given principle is meaning-constitutive of an expression entails that the principle is true, but rather that one cannot rationally intend one's expressions to be understood in accordance with principles one takes to be false. In this sense, the theorist is rationally committed to taking the meaning-constitutive principles of the terms she uses to be true. All that is assumed here is that the rational thing to do for, e.g., an intuitionist is to stipulate the meanings of their theory's constants by merely intuitionistically valid principles, rather than classical ones. This does not preclude one from also saying that the logical constants in natural languages have classical meaning-constitutive principles. Thus, an intuitionist with such a view about ordinary language must, in order to be fully rational, accept only theories whose constants she takes to have different meanings from those of ordinary language.

Here is another possible move that an objector might try at this point. One simply says that "true" is to be understood in its ordinary sense, but this just means that the property it expresses is, à la Eklund, that which best satisfies the meaning-constitutive principles. The reply—and note that this is not the criticism of section I—is that this is simply an inconsistent instruction for how to interpret "true" as used by the theorist. For the second claim, that "true" is to be understood as expressing whatever property best satisfies the instances of (T) simply means that it is not to be understood in accordance with the meaning it has according to (IT). Rather, on this stipulation, the meaning-constitutive principle for "true" is *consistent*. To wit, it is to the effect that something is true iff it has the property that best satisfies the (T)-instances.

That stipulation clearly gives a different meaning to "true". For instance, with such a stipulation, there would be nothing semantically incompetent about denying a pathological instance of (T). In fact, doing so with the motivation that it leads to contradiction would be to display one's *competence* with "true", relative to the sense given by the stipulation under consideration. Thus, on this stipulation of "true", the promise to use it in its ordinary sense is frustrated. Once again, we could verify this claim by comparing with the case of logical constants. Surely, saying that by "and", one will express a concept which is such that most, but not all, instances the introduction and elimination rules of conjunction come out as truth-preserving is different from saying that one will express the concept of classical conjunction. Indeed, even those who reject these rules—and so reject classical logic—agree that most of their instances are truth-preserving. This concludes my argument.

The argument of this section essentially envisages a case analogous to that of inconsistency theorists using "true", and claims that whatever holds for the former should hold for the latter. The reason for this somewhat roundabout procedure is that the correct analysis of the crucial notions of commitment and entailment are highly contested. I thus wanted to avoid presupposing any specific view on such matters. However, although their correct analysis is uncertain, these concepts have a fairly clear role in our philosophical reasoning about theories, inferences, epistemic justification, etc. I submit that nothing beyond this general, widely agreed upon role has been appealed to in the above argument. Perhaps an inconsistency theorist using "true" could come up with an account of why "true" should be treated differently from logical constants and "bachelor" (and the examples could probably be multiplied). I cannot see what possible relevant difference could be appealed to for such purposes, but we should at least grant that there might be some relevant difference between these expressions that has gone unnoticed. Still, until any such case has been made, I think, we must conclude that inconsistencists' theories that use "true" in its ordinary sense are

inconsistent. At the very least, we should agree that, in view of this argument, inconsistency theorists who want to use "true" must explain how they do not thereby become committed to every instance of (T), and thus to a contradiction.

3. Living without the truth-predicate (and property designators)

Committing "true", with Neurath, to the list of forbidden words, is not a very conservative move. Still, I think, it is not quite as radical as it may first seem. If, for instance, we accept the deflationist claim that the only theoretically important function of "true" is its ability to increase the expressive power of a language in which it is introduced, then, since this logical role is essentially the same as that of a propositional quantifier, we could simply replace "true" with the latter, assuming it can be consistently defined. Or, we could perhaps define a predicate, which is similar to the truth-predicate except that it is consistent, or use one of the many predicates of this kind proposed in the literature on the truth-paradoxes.

If we accept the deflationist stance, then we are committed to showing that all phenomena we admit as real—in particular, semantic and mental ones—be explicable without the use of a truth-predicate that is not simply defined by (T). But the conclusion of this paper is that no consistent theory can use such a predicate at all, why deflationism and (IT), with its commitment to eschew "true" form a coherent (if controversial) whole. Semantic and mental content must then be understood in other terms, presumably use-theoretic or functionalistic ones, but this is not the place to discuss such precarious matters. Furthermore, logic must be reconceived as not dealing with validity in the ordinary sense of truth-preservation in a case. But there are many alternatives available, for instance, purely normative theories dealing in notions such as correct acceptance and rejection. There are also arguments (e.g., Field (2008, 2009) and Hofweber (2007, forthcoming)) aiming to show that validity should not be

understood in terms of truth at all. So, even if eschewing "true" in this domain is radical, it is not so in any obviously objectionable way.

It may also be asked whether we should propose a general linguistic reform, now that we have found "true" to be inconsistent. Aside from being very unrealistic, I believe this proposal is also undermotivated, since, assuming that "true" *is* inconsistent, this feature is demonstrably harmless for practical purposes: we have not collapsed intellectually from learning of the Liar. However, from a philosophical viewpoint, and, in particular, as concerns the question of how to formulate a consistent and complete theory of the world, we must conclude that "true" will not be in it.

I will close by considering a puzzle for any inconsistency theorist who agrees with me about using "true": should we say that there is no such thing as the property of truth (or the property of being true)? Naturally, we cannot say that there *is* such a property, but saying that there isn't comes with troubles of its own. Firstly, we would then be *using* "true". Secondly, we seem thereby to run into problems emerging from the semantics of "property". A plausible view of this expression is that it is (implicitly) defined by the schema

(P) x has $P(F)$ iff $F(x)$,

where "$P(F)$" stands proxy for the property designator formed from the predicate "F". But if so, then the claim that there is no property of truth would seem to entail, *via* (P), that it is not the case that a given sentence has the property of being true. But this entails that the sentence is not true, and thus, since this conclusion does not depend on any specific sentence, we could derive that all sentences are untrue.

Now, (P) is like (T) in generating contradictions. Yet, (P) seems just as meaning-constitutive for "property" as (T) is for "true", and for the same reasons. I think the obvious

way for the inconsistency theorist about truth of dealing with the puzzle is therefore to adopt an inconsistency theory of the notion of property. If the arguments of this paper are correct, we must then eschew "property" as well, and then the problems involving the alleged property of truth are dealt with by claiming that they have no consistent solution.

Some philosophers (e.g., Schnieder (2010)) think that the self-referential paradoxes involving the notion of property differs from the Liar paradox in that the former has a solution that is not available in the case of truth. The paradox involving "property" that is analogous to the Liar involves, instead of the Liar sentence, the term, "the property of being a property that does not have itself". With a derivation very similar to the Liar derivation, we can infer a contradiction involving this expression, given the unrestricted schema (P). It is thought, however, that we could treat this alleged paradox the same way we treat the "paradox" involving the barber who shaves every man that does not shave himself, i.e., by simply denying the existence of this property. However, this objection misses the point, for the present idea is that "property" is an inconsistent expression the same way that "true" is, i.e., that every instance of (P) is a meaning-constitutive principle. The reason why the barber solution is irrelevant to this theory is that, independently of whether there is such a property or not, the meaning-constitutive principles for "property" entails a disposition to accept the relevant instance of (P), and (P) entails that there is such a property. So the adherent of (IT) should accept a fully-fledged inconsistency theory about the notion of property, and thus eschew "property" and its cognates. This conclusion provides an independent reason to avoid the dilemma presented above, the choice between agreeing and denying that there is a property of truth.

[1] Such views have been defended by Wittgenstein (1956: app. I: 12), Chihara (1979, 1984), Yablo (1993), Barker (1998), Burgess (2002), Eklund (2002a, 2002b), Ludwig (2002),

Azzouni (2003, 2007), Patterson (2007a, 2007b, 2009), and Scharp (2007a, 2007b). Some of these papers are published in *Inquiry* **50** (6), a special issue on (IT).

[2] Eklund explicitly states that the semantics of "true" should be truth-theoretic, and so holds that "true" may—indeed, should—be used in the metalanguage (2002a: 264ff.). Charles Chihara does not quite engage in these metalinguistic questions, but does use "true" in his reasoning about various paradoxes, and, as we will note below, formulates his canonical assessments of the paradoxes using "true" (1979, 1984). Kirk Ludwig (2002) takes pains to ensure that his Davidsonian semantic theory does not *state* T-biconditionals. However, it does state "M-sentences" of the form "*S* means that *p*", and one of these is "'*Liar* is not true' means that *Liar* is not true" (where "*Liar*" is the name of a strengthened liar sentence). This, however, is a use, not a mention, of "true", and its occurrence in the scope of "means that" does not change this fact. Patterson's views have changed from his (2007a, b) to his (2009), but it is clear from his wording that he does not object to using "true". Scharp, however, sides with my own stance (2007a: 302), and will be discussed further below.

[3] See Eklund (2002a: 252, 266). Though Chihara never states that the argument is unsound in one sentence, it is clear from his "diagnoses" of other paradoxes (1979: 593, 597), and his claim that the Liar should be similarly treated (1979: 606, 611), that he means that the principle that gives "true" its meaning (e.g., his "[Tr]") is not true. His principal idea is that some sentences that appear to be made "true by *fiat*" (e.g., by definition or convention) are not true (see esp. (1979: 593-7)).

[4] Scharp (2007a: 302) holds a view similar to mine on this issue, but formulates it somewhat differently: "It is my view that inconsistent concepts should be replaced with consistent ones; they aren't fit for employment. The rationale for this view is simple: other things considered, one should avoid undertaking incompatible commitments. Thus, if one discovers that a concept is inconsistent, one should stop employing it if one can.". However, he does not

elaborate these points further.

[5] Scharp (2007b) and Patterson (2009) also criticise Eklund's views about the determination of the semantic value of "true", but their arguments are quite different in that they do not appeal to intuitions about distinct cases of meaning-constitutive principles determining semantic values.

[6] Eklund's wording (2002a: 272) suggests that it is truth, i.e., a property, rather than an extension, that is the semantic value of "true", but I think this is immaterial to the present discussion. I will also not enter into more technical (but fairly unproblematic) questions of what it is for a property or extension to satisfy a principle.

References

Azzouni, J. (2003), "The Strengthened Liar, the Expressive Strength of Natural Languages, and Regimentation", *The Philosophical Forum* **34**, pp. 329-50.

Barker, J. (1998), *The Inconsistency Theory of Truth*, unpublished doctoral dissertation, Princeton University.

Beall, J. C. (2007), *The Revenge of the Liar: New Essays on the Paradox*, Oxford, Oxford University Press.

Burgess, J. P. (2002), "Is There a Problem about the Deflationary Theory of Truth?", Halbach, V. and L. Horsten (2002), pp. 37-55.

Chihara, C. (1979), "The Semantic Paradoxes: A Diagnostic Investigation", *The Philosophical Review* **88**, pp. 590-618.

—— (1984), "The Semantic Paradoxes: Some Second Thoughts", *Philosophical Studies* **45**, pp. 223-9.

Daly, C. & D. Liggins (2010), "In Defence of Error Theory", *Philosophical Studies* **149**, pp. 209-230.

Davidson, D. (1967), "Truth and Meaning", *Synthèse* **17**, pp. 304-323.

Eklund, M. (2002a), "Inconsistent Languages", *Philosophy and Phenomenological Research* **64**, pp. 251-75.

—— (2002b), "Deep Inconsistency", *Australasian Journal of Philosophy* **80**, pp. 321-31.

Field, H. (2008), *Saving Truth from Paradox*, Oxford, Oxford University Press.

—— (2009), "The Normative Role of Logic", *Proceedings of the Aristotelian Society Supplementary Volume* **83**, pp. 251-268.

Frege, G. (1884), *Die Grundlagen der Arithmetik: Eine Logisch-Mathematische Untersuchung über den Begriff der Zahl*, Breslau, W. Koebner. Translated by J. L.

Austin (1974), *The Foundations of Arithmetic: A Logico-Mathematical Enquiry into the Concept of Number*, Oxford, Blackwell.

Halbach, V. and L. Horsten (eds.) (2002), *Principles of Truth*, Frankfurt a. M., Hänsel-Hohenhausen.

Hofweber, T (2007), "Validity, Paradox, and the Ideal of Deductive Logic", in Beall (2007), pp. 145-158.

Hofweber, T. (forthcoming), "Inferential Role and the Ideal of Deductive logic", forthcoming in *Meaning, Understanding, and Knowledge*, D. Patterson (ed.), Riga, University of Latvia Press.

Horwich, P. (2008), "What's Truth Got To Do with It", *Linguistics and Philosophy* **31**, pp. 309-322.

Lewis, D. (1970), "How To Define Theoretical Terms", *The Journal of Philosophy* **67**, pp. 427-446.

——— (1984), "Putnam's Paradox", *Australasian Journal of Philosophy* **62**, pp. 221-36.

Ludwig, K. (2002), "What is the Role of a Truth Theory in a Meaning Theory?", in Campbell, O'Rourke and Shier (eds.), *Meaning and Truth: Investigations in Philosophical Semantics*, New York, Seven Bridges Press, pp. 142-63.

McGinn, C. (1977), "Charity, Interpretation, and Belief", *Journal of Philosophy* **74**, pp. 521-535.

Patterson, D. (2007a), "Understanding the Liar", in Beall (2007), pp. 197-224.

——— (2007b), "Inconsistency Theories: The Significance of Semantic Ascent", *Inquiry* **50**, pp. 575-589.

——— (2009), "Inconsistency Theories of Semantic Paradox", *Philosophy and Phenomenological Research* **79**, pp. 387-422.

Putnam, H. (1962), "It Ain't Necessarily So", *Journal of Philosophy* **59**, pp. 658-671.

Scharp, K. (2007a), "Aletheic Vengeance", in Beall (2007), pp. 272-319.

——— (2007b), "Replacing Truth", *Inquiry* **50**, pp. 606-621.

Schnieder, B. (2010), "Inexpressible Properties and Grelling's antinomy", *Philosophical Studies* **148**, pp. 369-385.

Tarski, A. (1944), "The Semantic Conception of Truth", *Philosophy and Phenomenological Research* **4**, pp. 341-76.

Wittgenstein, L. (1956), *Remarks on the Foundations of Mathematics*, Oxford, Blackwell.

Yablo, S. (1993), "Definitions, Consistent and Inconsistent", *Philosophical Studies* **72**, pp. 147-75.