Could "knows that" be inconsistent?*

Gillian Russell

Draft at 15th September 2009

In his recent *Philosophers' Imprint* paper "The (mostly harmless) inconsistency of knowledge attributions" [Weiner, 2009], Matt Weiner argues that the semantics of the expression "knows that", as it is used in attributions of knowledge like "Hannah knows that the bank will be open," are inconsistent, but that this inconsistency is "mostly harmless." He presents his view as an alternative to the invariantist, contextualist and relativist approaches currently prevalent in the literature, (e.g. [DeRose, 1995], [Lewis, 1996], [Stanley, 2005], [MacFarlane, 2005], [Hawthorne, 2006]) and argues that it avoids important disadvantages of each.

Yet in calling the supposed inconsistency of knowledge attributions "mostly harmless" Weiner implies that his view does not have new disadvantages of its own. My purpose in the present paper is to argue that this is not so, and that if we accept that the semantics of 'know'—or indeed any word—are inconsistent, we face a dilemma: either we accept dialetheism, the view that there there are true contradictions, or we must accept that semantic competence in English requires belief in, or similar commitment to, claims that are not true. I will show that neither of these options is well described as "mostly harmless."

The paper is structured as follows: in the first part I present Weiner's view and his arguments for it. Then in section 2 I compare the question of whether the semantics of 'knows that' are inconsistent with the much older controversy over whether the semantics of the expression 'is true' are inconsistent. In section 3 I will present Herzberger's arguments from the 1960s for thinking that no expression in a natural language can have inconsistent semantics. Finally, in section 4 I argue that although Herzberger's argument seems anachronistic today, both the contemporary ways of avoiding his conclusion have significant disadvantages.

1 Weiner on Knowledge Attributions

Weiner holds that in order to be competent with the expression 'know', speakers must be disposed to accept three schemata:

*With thanks to Stephen Finlay, Andrei Marmor, Lewis Powell, Jacob Ross, Mark Schroeder, Scott Soames and the other participants at the USC philosophy department colloquium for helpful discussion.

Disquotation Principle An utterance of "S knows that p" at time t is true iff at time t, S knows$_{tenseless}$ that p.

Practical Environment Principle S's evidence for p is good enough for knowledge iff S's evidence for p is good enough to make it epistemically rational for her to act on the assumption that p.

Parity of Evidence Principle If the evidence concerning p for S and T is the same, then S's evidence is good enough for knowledge if T's is good enough for knowledge.[1]

Each of these principles seems reasonable enough at first blush (though the issue of what is really required for semantic competence is a fraught one.) Yet if we allow a few background assumptions, the most important of which are that that skepticism is false, and that people's practical interests differ, then these principles entail a contradiction.[2]

We can use a well-known example to bring this out. Suppose that Hannah and Sarah are driving home from work on Friday night with their paycheques. They plan to deposit them at the bank on the way home but as they drive past they notice that the lines for the tellers' windows are very long. Hannah was at the bank on a Saturday a few weeks ago and noticed that it was open. She suggests that they wait and deposit their cheques in the morning, saying "I know that the bank will be open on Saturday."

Suppose that Hannah is right: she does know that the bank will be open on Saturday. Then by the Practical Environment Principle she must have evidence sufficient to make it practically rational for her to act on the assumption that the bank is open on Saturday. Yet her evidence might be sufficient in part because it is not very important that she deposit her cheque. Sarah—to whom it is vitally important that her cheque is deposited before Monday, since otherwise her rent cheque will bounce—may have the same evidence even though in her case it is insufficient to make it epistemically rational for her to act on the assumption that the bank will be open on Saturday, because for her the stakes are higher. So by the Practical Environment Principle Sarah's evidence is insufficient for knowledge, and hence Sarah does not know that the bank will be open on Saturday. But by the parity of evidence principle, Sarah's evidence *is* sufficient for knowledge, since it is the same as Hannah's, and hence (assuming everything else is in place) Sarah knows that the bank is open on Saturday. This is a contradiction already, but using the disquotation principle we can also produce a metalinguistic version of this contradiction: "Sarah knows that the bank will be open" is true and "Sarah knows that the bank will be open" is not true.

Each of the three mainstream non-sceptical views of the semantics of 'knows that'—contextualist, relativist and sensitive invariantist—avoids this contradiction by denying one of the three Principles. Contextualists like DeRose and

[1] Being disposed to accept these schematic principles will at least entail being disposed to accept each of their instances.

[2] Thus they are not actually inconsistent all on their own, but rather inconsistent with some theses that we might assume to be true for the purposes of the debate.

relativists like MacFarlane, deny the Disquotation Principle, on the grounds that whether or not 'S knows that p' is true at time t can depend on the epistemic standards in place at the context in which the sentence is uttered, (or in the case of relativists, the context with respect to which the utterance is assessed) and not just the time and quality of S's evidence at the time of utterance. Sensitive invariantists like Stanley and Hawthorne will deny the Parity of Evidence Principle, arguing, in Stanley's case, that the differing practical interests of Hannah and Sarah may lead to 'Sarah knows that the bank will be open' being false, while 'Hannah knows that the bank will be open' is true, even where Hannah and Sarah have the same evidence.

But there is a cost to the rejection of these principles, since ordinary speakers, not in the grip of a sophisticated philosophical theory, will often speak as if they take the principles to be true, so that theorists are forced to explain the intuitions of ordinary speakers away as a kind of "semantic blindness"—an inability on the part of competent speakers to see what the correct semantics of 'knows that' requires them to judge in these cases. For example, speakers may be apt to judge that since Sarah does not have sufficient evidence to count as knowing that the bank will be open on Saturday and Hannah has the same evidence for that proposition, 'Hannah knows that the bank will be open on Saturday' cannot be true. Yet from the point of view of sensitive invariantism, this judgement would be a mistake, caused perhaps by missing the semantic fact that the standards required for knowledge can vary with the *subject's* practical interests.

On Weiner's view however, the naive speakers have got the semantic facts exactly right. The three knowledge principles are not only correct, but essential to semantic competence with 'knows that'; that is, in order to be semantically competent with 'knows that', one must be disposed to accept these principles. This need not mean that competent speakers will never deny them; Weiner follows [Eklund, 2002] in allowing that speakers may reject, or suspend belief in constitutive semantic principles, or the consequences of such principles, when they conflict with theoretical convictions or appear to lead to contradiction.

However, to say that being disposed to accept the three principles is a condition of semantic competence *is* to be committed to the following: that sentences which conflict the these principles will sound semantically inconsistent—on a par with "some bachelors are unmarried" or "there have been black dogs and there have been white cats, but there have not been white cats." [Weiner, 2009, 3] The Knowledge Principles are, in a sense, analytic, and claims that conflict with them, like those below, should sound, not just false, but actually incoherent or "analytically false":

(1) The evidence concerning whether the bank will be open on Saturday is the same for Hannah and Sarah, but Hannah's evidence is sufficient for knowledge and Sarah's is not.

(2) "Hannah knows that the bank will be open on Saturday" at time t is true but at time t Hannah knows$_{tenseless}$ not that the bank will be open on Saturday.

(3) Hannah's evidence concerning whether the bank will be open on Saturday is good enough for knowledge but Hannah's evidence concerning whether the bank will be open on Saturday is not good enough to make it epistemically rational for her to act on the assumption that the bank will be open on Saturday.

This then, is Weiner's argument for favouring the inconsistency account: unlike its rivals it is able to respect each of the Knowledge Principles, and in doing so it explains the pull of the paradoxes they seem to lead us into. Yet Weiner seems to be aware that, as is often the case with inconsistency accounts, the main objection to the view is unlikely to be that there is no argument for it, but rather that it is incoherent, or somehow incomprehensible. He goes out of his way to combat this anticipated objection by providing "a model for inconsistent knowledge-talk" [Weiner, 2009, 11], a fictional world in which talk of *time* is intended to have many of the properties which Weiner wishes to attribute to knowledge talk. He also stresses the harmlessness of his account, suggesting both that it does not have any strong or unpalatable consequences and that we are not compelled to change our semantics:

> However, the thesis is radical only at the theoretical level; it does not call for a radical revision of our practice of knowledge ascription. In particular, even if knowledge-talk is inconsistent, we need not and should not abandon it. The cases in which the inconsistency might lead to actual confusion are rare enough that knowledge-talk is an efficient way of communicating. [...] The inconsistency of 'know' is mostly harmless. [Weiner, 2009, 1]

2 Tarski and Inconsistent Languages

From these claims of harmlessness, it seems unlikely that Weiner intends to commit *himself* to a contradiction, as he happily might if he were endorsing dialetheism. It seems, rather, as if he means to claim that there is an inconsistency *in the language*, rather as a non-dialetheist might claim that there is an inconsistency in a certain text, or in the rules of baseball, or in the law. One can claim that there is an inconsistency in the rules of baseball without thereby committing oneself to a contradiction. In this Weiner has an illustrious predecessor in Alfred Tarski, who held that the rules governing the truth-predicate in semantically-closed languages led quickly to contradiction (given some obvious facts):

> A characteristic feature of colloquial language...is its universality... But it is presumably just this universality of everyday language which is the primary source of all the semantical antinomies, like the antinomies of the liar or of heterological words. These antinomies seem to provide a proof that every language which is universal in the above sense, and for which the normal laws of logic hold, must be inconsistent. [Tarski, 1944, 347–8]

Tarski was persuaded of this view by another argument to a contradiction: the Liar paradox. He discovered that if we have a sufficiently rich language—one in which it is possible to form names for the sentences of the language (perhaps by putting them in quotation marks)—and we introduce a truth predicate whose definition entails all instances of the schema below[3]

[T] 'S' is true if, and only if, S.

then we can derive a contradiction. For consider the sentence:

[L] *The only sentence printed in italics in this paper is not true.*

We can substitute [L] into schema [T], to get this:

(4) 'The only sentence printed in italics in this paper is not true' is true if, and only if, the only sentence printed in italics in this paper is not true.

Then we make the empirical observation that the only sentence printed in italics in this paper is sentence [L] above. This could of course have been otherwise—perhaps the printer could have made a mistake. But given that [L] *is* the only sentence in italics, the following identity sentence is true:

[ID] The only sentence printed in italics in this paper = 'The only sentence printed in italics in this paper is not true.'

As a result we preserve truth when we substitute the description on the left hand side for the name on the right, and when we do this in [5] we get:

(5) The only sentence printed in italics in this paper is true if, and only if, the only sentence printed in italics in this paper is not true.

But (5) is of the form $S \leftrightarrow \neg S$, and so leads quickly to a contradiction using classical truth-functional logic. Thus what Tarski seems to have meant when he said that natural languages were inconsistent was that once one formulated the correct definitions governing the expressions in the language, one was able to use classical logic to derive a contradiction from those definitions and (in this version) some obvious non-linguistic truths, such as [ID].

Like Wiener, Tarski does not seem to take *himself* to be committed to contradictions. That position was not really on the philosophical radar in a serious way at the time (Heraclitus and Protagoras having died some years earlier, Routley and Priest not yet having written their contributions to the subject) and Tarski seems to consider the necessary revision in logic beyond the pale:

> It would be superfluous to stress here the consequences of rejecting the assumption (II), that is, of changing our logic (supposing this were possible) even in its more elementary and fundamental parts. [Tarski, 1944, 349]

[3]Instances of this schema are formed by uniformly replacing instances of the schematic letter 'S' with a sentence.

But Tarski's view also differs from Weiner's in two significant respects: Tarski did not take the claim that a language was inconsistent to be a claim about what could be derived from what was required for *semantic competence*, but rather about what could be derived from what was required for a *correct definition* of the truth-predicate. Also unlike Weiner, Tarski did not think the presence of an expression that led to inconsistency should be regarded as harmless:

> It is a fact that we are here in the presence of an absurdity, that we have been compelled to assert a false sentence [...] If we take our work seriously, we cannot be reconciled with this fact. We must discover its cause, that is to say, we must analyze premises upon which the antimony is based; we must then reject at least one of these premises, and we must investigate the consequences which this has for the whole domain of our research. [Tarski, 1944, 348]

3 Herzberger's Arguments

Tarski's writings brought the idea that languages could be inconsistent to the notice of the philosophical community, where they were accepted by Carnap, who wrote that "customary usage of the terms 'true' and 'false' leads...to a contradiction." [Carnap, 1937] It became common to think that natural languages were vague and inconsistent, and that this was something that proper scientific languages should seek to avoid. But not everyone accepted that languages *could* be inconsistent in the first place. One objector was Hans Herzberger, who wrote:

> [N]o such antimony can show our language to be logically inconsistent. Nor can any of the antimonies be properly said to to be 'present' or 'derivable' within natural languages; for if they were, that would render the language inconsistent, which is impossible [Herzberger, 1965, 478]

In [Herzberger, 1965] and [Herzberger, 1967], he argued that any claim that a language is inconsistent is *itself* an inconsistent claim. Herzberger took these arguments to be reductios ad absurdum on the idea of an inconsistent language. His original argument went like this:

A1 To say that a language is inconsistent is to say that the set of analytic sentences in that language is inconsistent.

A2 If a set of sentences is inconsistent, then at least one element of the set is not true.

A3 If a sentence is analytic, then it is true.

AC To say that a language is inconsistent is to be committed to the claim that there is a set of sentences, every element of which is true, and at least one element of which is not true.

Thus to claim that a language is inconsistent is commit oneself to a first order inconsistent claim.

The argument is refreshingly simple, but there are some prima facie questions as to how it connects up with Tarski and Weiner's views for, first, it is not at all clear that what each meant by "inconsistent language" was a language in which the analytic sentences are inconsistent. Although Weiner relies on a notion a lot like analyticity, Tarski appears to have shared Quine's suspicions about the notion, at least in (erstwhile) private correspondence with Morton White. [White and Tarski, 1987] And second, both Tarski's version of the Liar and Weiner's paradox concerning knowledge show their constitutive meaning principles conflicting, not with another analytic sentence, but with obvious or accepted synthetic claims; in Tarski's case the fact that the sentence on page 347, line 44 of that paper is the sentence "the sentence on page 347, line 44 of this paper is false,"[4] in Weiner's the falsity of skepticism and fact that people's practical interests may vary.

One way around the problem is to adapt Herzberger's definition of inconsistency for a language, as we do here in B1:

B1 To say that a language is inconsistent is to say that the union, $A \cup B$, of the set of analytic sentences of the language A, and some set B containing containing only true sentences of the language is inconsistent.

B2 If $A \cup B$ is inconsistent, then at least one element of that set is not true.

B3 By hypothesis, every element of B is true.

B4 Every element of A is analytic, and if a sentence is analytic, then it is true.

B5 Every element of $A \cup B$ is either a member of A or a member of B.

B6 Every element of $A \cup B$ is true. [B3, B4 and B5]

BC To say that a language is inconsistent is to be committed to the claim that there is a set of sentences, every element of which is true, and at least one element of which is not true.

This fix disposes of the problem caused by Tarski and Weiner's relying on empirical sentences to achieve their contradictions.

But what if one has doubts about the very notion of an analytic sentence? Herzberger addresses this problem in [Herzberger, 1967]. He suggests a way to frame the argument in terms of truth-conditions, instead of in terms of analytic sentences. We let T be the relation that holds between a sentence s and a condition w if the condition is sufficient for the truth of the sentence, a situation that we express as follows: $T(w, s)$. Then a Type B set of sentences is a set of sentences, Q, meeting the following conditions: i) Q is logically inconsistent, ii)

[4]This claim plays the same role in Tarski's argument as the claim made by [ID] plays in my recreation of it.

there is a condition w such that $T(w, s)$ for every $s \in Q$ and iii) it is logically possible that w obtain.[5] We will say that a language L is *inconsistent in its truth conditions* just in case L gives rise to at least one Type-B set of sentences with respect to those truth conditions. Here is the argument that no such language is possible:

C1 Suppose that L is inconsistent in its truth-conditions.

C2 Then it gives rise to at least one Type-B set, that is, there is a set of sentences Q of L such that

  i) Q is logically inconsistent,

  ii) there is a condition w such that $T(w, s)$ for every $s \in Q$ and

  iii) it is logically possible that w obtain.

C3 Given what $T(w, s)$ means, we know that if w obtained, every member of Q would be true.

C4 But Q is inconsistent, so at least one member of Q must be false.

C5 Hence it is not logically possible for w to obtain.

CC So Q is not a Type-B set of sentences.

Again we derive a contradiction from the assumption that a language meets the criteria for being inconsistent.

One final hurdle remains for the project of adapting Herzberger's reductio ad absurdum of the claim that a language is inconsistent to Weiner's version of the inconsistency view. The problem is that Weiner is very ambivalent about the notion of truth-conditions, feeling that his view sits more easily with inferentialist approaches to semantics:

> My account of inconsistency does fit well with inferentialist semantics, and the idea that the inconsistent predicate does not correspond to anything in the world, but it may also be possible to reconcile it with a truth-conditional semantics. [Weiner, 2009, 4]

Fortunately Weiner's account of what it is for a language to be inconsistent does not require Herzberger's twist involving truth-conditions, since Weiner relies on a notion from the analytic family, that of being "semantically inconsistent" or analytically false.

[5] The last condition is added to rule out the chance that we think that if w is the condition that $2 + 2 = 4$ and $2 + 2 \neq 4$, then it stands in the T relation to the sentence $2 + 2 = 4$ and $2 + 2 \neq 4$.

> And because the disposition comes from semantic competence rather than general knowledge, the resulting utterances will be not just strange-sounding but akin to contradictions. They will sound not just obviously false but as though they do not make sense." [Weiner, 2009, 4]

This makes Weiner's view susceptible to the B-version of Herzberger's argument.

4 Dialetheism and Analyticity Without Truth

Herzberger thought that his arguments were straight-forward reductios ad absurdum of the view that languages can be inconsistent, but no contemporary writer could leave the argument here. There are now two rather surprising options open to someone who likes inconsistent languages, and both these options are endorsed by philosophers for reasons that are independent of the Herzberger arguments.

The first option is a move to outright dialetheism, holding that a set containing only true sentences *can* be inconsistent—that is, entail a contradiction—and that the reason this is possible is that contradictions so derived are *true*. For Herzberger and Tarski this seems to have been an almost unimaginable position, but the move has allowed dialetheists like Priest to maintain that English is inconsistent. [Priest, 2006, 5] This is a fascinating position and there is a lot to be said in favour of it, but I think it is fairly obvious that one thing that could not be said about it is that it is *harmless*. If an argument is wanted, there is this: in classical logic everything follows from a contradiction, so if we retain classical logic, dialetheism collapses into *trivialism*, the view that everything is true. Dialetheists do not want to be trivialists, and so must give up classical logic for a weaker one—one which does not have $\alpha, \neg\alpha \vdash \beta$ as a theorem. Hence dialethesim is not harmless; it forces wholesale weakening of our logic, and that can only be regarded as a significant cost.

The other option—which would also have seemed alien to Tarski and Herzberger, but which has been endorsed by both Tappenden [Tappenden, 1993] and Eklund [Eklund, 2002]—is that there are some analytic sentences *that are not true*. As a response to Herzberger's argument, this option works by denying the premise (A3 in the simplest version) which states that analytic sentences are true. And if a set of analytic sentences is not a set of true sentences, then of course there is no problem with it being inconsistent.

The view might seem to be easily dismissed. After all, isn't "analytic" more or less short for "analytically *true*" and glossed by the phrase "*true* in virtue of meaning"? I rather think that's so [XXXX, 2008], but in order to give the opposing idea some traction, we can make use of Boghossian's distinction between epistemic and metaphysical analyticity. [Boghossian, 1996] A sentence is *metaphysically analytic* just in case it is true solely in virtue of its meaning, and *epistemically analytic* just in case someone who is semantically competent with it knows, or is justified in believing what it says. Eklund holds that it is

only epistemically analytic sentences that may be inconsistent [Eklund, 2002, 253] so let us try taking 'analytic' to mean something more like 'epistemically analytic'.

To this end, we will assume for the sake of argument that to say that a sentence is analytic is to say that someone who understands it is *justified in believing what it says*. Since justification does not entail truth, it no longer follows immediately that an analytic sentence must be true. But what is it that semantically competent speakers have that justifies them in their belief? There are a few possible answers to this question, corresponding to different things we can say about the actual status of the analytic-but-not-true sentences. Semantic competence may be taken to require special kinds of knowledge (e.g. in order to be competent with the word 'bachelor' one has to know that it applies only to men) but weaker views may require only special kinds of belief (in order to be competent with the word 'true', one has to believe the T-schema) or weaker still, semantic competence may not require the belief itself, but only a disposition to believe.[6] In the remainder of the paper, I will argue that none of these three versions of the view is "harmless."

The Knowledge Version

The simplest and strongest version of the 'analytic and false' view, will say that a sentence can be analytic and false, where to say that a sentence is analytic is to say that anyone who has the relevant semantic competence will *know* things that entail the proposition expressed by the sentence. This strong version of the view is of dubious coherence. Since 'know' is factive, one cannot *know* that something false is true (at least, not if dialetheism is off the table for this part of the discussion.)

The Belief Version

A more feasible version of the view would employ the second view of semantic competence, according to which a sentence can be analytic and false, where to say that a sentence is analytic is to say that anyone who has the relevant semantic competence believes things that entail the proposition expressed by the sentence. This view has two non-harmless consequences. First, suppose that the sentence 'L is true and L is not true' (where 'L' is a name for the Liar sentence) is analytic but false in this way. Then, ex hypothesi, it is entailed by sentences that one must accept if one is to be semantically competent. But the Liar sentence is contradictory, so those things are inconsistent. So at least one of the things one must believe in order to be semantically competent must be false. If this view is right, *the ability to speak English requires belief in falsehoods.*

The view is both implausible and such that if it did turn out to be true, it would be far from harmless. To see that it is implausible, consider this more

[6] This trichotomy holds for inferences required for semantic competence too. One might require anything from knowledge that the inference is valid to a mere disposition to make the inference.

morally loaded example. It has sometimes been suggested that the reason that certain racial epithets like "nigger" (in the derogatory sense) are so objectionable is that semantic competence with them requires racist beliefs, i.e. one can only really understand the word if one believes the falsehood that the colour of one's skin is causally related to with one's value as a human being. This view of the semantics has the value of providing an explanation of the interesting fact that many people who do not hold this view will not use the word at all (not even to say approbative things) whilst many people who do hold the false view will. However, it also entails that people who do not believe this cannot really understand sentences containing the word; they are not semantically competent with it. This is surely wrong; those who do not hold the false belief refrain from using the word, and object when others use it, because they understand all too well what it means.[7]

Similarly, if one took this approach to the Liar paradox, one might hold that semantic competence with 'true' requires belief in the T-schema (that is, belief in all instances of it) even though there are instances of the T-schema (those involving the Liar sentence, for example) which are not true. The advantage of this view is that it explains our feeling that the T-schema is not just plausible, but analytic; counter-instances seem semantically wrong in at least as strong a way as sentences that violate Weiner's knowledge principles. Because of this the view can be credited with explaining the *pull* of the paradox, the reason why we ended up in trouble in the first place

But this approach to the Liar also entails that people who don't believe the T-schema are not semantically competent with the English word 'true' and this is surely false. One of the two most common responses to the Liar is to restrict the T-schema in some way (e.g. [Kripke, 1984], [Glanzberg, 2004]), and it just seems false to say that in giving up a belief in the T-schema, such philosophers have lost their semantic competence with the English word 'true'. That would suggest that they don't know what they are talking about anymore, and should just shut up and let those of us who have retained the appropriate *false* belief tell them what's what!

Moreover, even if the view were true, it would not be harmless. Suppose there really were a word semantic competence with which *required* belief in a

[7] I have heard it suggested that a possible response here is to distinguish two different kinds of semantic competence: what is required for competent *understanding*, and what is required for competent *use*. One might then say that one can be competent to understand the racial epithet without having the racist beliefs, but that such beliefs are necessary for competent use. However I think this distinction is both mistaken and insufficient to solve the problem. It is mistaken because the semantic knowledge required for understanding *just is* the semantic knowledge required for use. To be able to understand an expression is to be able to use it in your thoughts. To be able to use an expression competently requires being able to understand it. But even if what was required for competent use was different from what was required for competent understanding, the problem would remain: non-racists will not only refrain from using 'nigger', but they will appropriately criticise others' uses of it. Hence they *do* know what is required for competent use of words like 'nigger'; if they didn't, they would have no grounds to criticise the uses of others. The best explanation of their refraining from using the word is not that they lack the semantic *competence* to use it, but rather that they are not racists.

falsehood. Surely this *would itself* be a reason to prune the word from our language—unless we do so we must have false beliefs! Someone might try to argue that there can sometimes be good reasons to believe false things, and that the simplicity or familiarity of an expression could be reason enough to retain it, despite the fact that competence with it requires belief in a falsehood. But this presupposes two unlikely things. First, it presupposes a very strong brand of voluntarism about belief, i.e. the view that we are able to choose believe something even when we know it is false. But how plausible is that? Is it really that likely that someone might read a lot about the Liar paradox, come to the conclusion that the T-schema was false, and successfully elect to go on believing it anyway—on the grounds of simplicity and entrenchment—in order to remain competent with the word 'true'?

And second, I think it underestimates the infectiousness and dangers of having false beliefs. To be sure, there are examples in the literature which make it seem plausible that there are some occasions on which it would be appropriate to attempt to believe things that are probably false: that I will make the leap, that I will get the job, that I will win the race. But as W. C. Clifford argued, anything we believe can have consequences for what other things we believe, for what other things we will relate to others, and thence for our actions and the actions of others; both testimony and action can have serious, unforseen, moral consequences. Even were it possible, deciding to believe the false is epistemically reckless, and, on the whole, epistemic recklessness is morally reckless.

The Dispositional Version

The last and weakest version of the analytic-and-false response that I will consider says that a sentence is false and analytic if i) it is false and ii) anyone who is semantically competent will be *disposed* to accept things which entail the sentence. This is the view endorsed by both Weiner and Eklund, the most contemporary of my targets:

> Let us call the untrue premises and invalid steps in an unsound argument the *culprits* of the argument. Then what I shall argue is that being competent with the expressions employed in the paradoxes involves being disposed to accept the culprit as true or valid: so long, anyway, as you do not have what you take to be evidence against its truth or validity. [Eklund, 2002, 252]

> The Knowledge Principles [...] are deep, if not constitutive, principles about how we use 'know'. A speaker who is competent to use the word 'know' will be disposed to accept inferences in accordance with the principles. These dispositions can be overridden, but insofar as you completely lack them, you are not using 'know' the way it is used in English. [Weiner, 2009, 2]

This version of the view does have certain advantages over the previous, stronger version of the view, but it seems to me that the advantages are the advantages

that a weak version of a wrong view has over a stronger version of a wrong view. In particular, this view is also implausible, and it would still have non-harmless consequences similar to the ones that the belief version of the view had.

To really bring out that the view is implausible, consider again the attempt to make the same view work in a more morally loaded case, and suppose, not that one needs to have a racist *belief* to be competent with the word 'nigger', but that one needs to have a disposition towards holding the racist belief. On this view, in the absence of undermining evidence, no-one could be competent with the word unless they had the racist belief. But again, this is clearly wrong. Never having being disposed towards racists beliefs, and not holding racist beliefs even in the absence of undermining evidence, does not make one unable to understand the word 'nigger'. Why think things are any different with 'true' or 'knows that'?

The belief-version of the analytic-but-false view also had the non-harmless consequence that in order to be competent with certain expressions, one would have to believe a falsehood. The disposition-version has the consequence that in order to be competent with certain expressions, one has to be disposed to believe a falsehood—sufficiently strongly disposed that in the absence of undermining evidence, one will believe the falsehood. I suppose we might think that this is a slightly less bad consequence, but again, we can ask the question, why isn't this reason enough to give up the expression? It puts us in an epistemically prejudiced position.[8]

5 Conclusion

It has recently been suggested that the heroic attempts of epistemologists and philosophers of language to find a consistent semantics for 'knows that' have been misplaced: the semantic rules governing 'knows that' are inconsistent, and we should accept that, and then go on using the expression as before. I have argued here that closer attention to the history of debate over the Liar paradox shows this complacency to be misplaced: holding part of our language to be inconsistent forces us to choose between dialetheism and a view on which our very competence in English is epistemically compromising. A commitment to the disjunction of these views adds substantially to the inconsistency view's cost. If I am right, then the honest toil required to uncover a consistent semantics looks worthwhile after all.

[8]For much of this section I have written as though if an analytic sentence is not true, it must be false. However, as Mark Schroeder has pointed out to me, one might instead hold that they are *truth-valueless*. I think he is entirely right that this is another option, and in fact the view fits better with Tappenden's position on analytic sentences. [Tappenden, 1993] Moreover, once one has agreed to move to a 3-valued approach like this (true/false/neither), one can take the usual steps to make it more palatable (e.g. as in [Soames, 1999].) However I think it is fairly clear that this option has consequences from *both* the dialetheist and the analytic-but-false camps. If we take it, we must allow *both* that competence in English requires commitment (in a stronger or weaker way) to untruths *and* that we need a new logic.

References

[Boghossian, 1996] Boghossian, P. A. (1996). Analyticity reconsidered. *Nous*, 30(3):360–391.

[Carnap, 1937] Carnap, R. (1937). *The Logical Syntax of Language*. Keagan Paul, London.

[DeRose, 1995] DeRose, K. (1995). Solving the skeptical problem. *Philosophical Review*, 104(1).

[Eklund, 2002] Eklund, M. (2002). Inconsistent languages. *Philosophy and Phenomenological Research*, 64:241–275.

[Glanzberg, 2004] Glanzberg, M. (2004). A contextual-hierarchical approach to truth and the liar paradox. *Journal of Philosophical Logic*, 33:27–88.

[Hawthorne, 2006] Hawthorne, J. (2006). *Knowledge and Lotteries*. Oxford, New York.

[Herzberger, 1965] Herzberger, H. G. (1965). The logical consistency of language. *Harvard Educational Review*, 35:469–480.

[Herzberger, 1967] Herzberger, H. G. (1967). The truth-conditional consistency of natural languages. *Journal of Philosophy*, LXIV(2):29–35.

[Kripke, 1984] Kripke, S. A. (1984). Outline of a theory of truth. In Martin, R. L., editor, *Recent Essays on Truth and The Liar Paradox*, pages 53–81. Oxford University Press, New York.

[Lewis, 1996] Lewis, D. (1996). Elusive Knowledge. In Dretske, F. and Bernecker, S., editors, *Knowledge: Readings in Contemporary Epistemology*, pages 367–384. Oxford University Press.

[MacFarlane, 2005] MacFarlane, J. (2005). The assessment sensitivity of knowledge attributions. In *Oxford Studies in Epistemology 1*, pages 197–233. Oxford University Press, Oxford.

[Priest, 2006] Priest, G. (1987/2006). *In Contradiction*. Oxford University Press, Oxford, 2nd edition.

[XXXXXX, 2008] XXXXXX (2008). *XXXXXXXXXXXX*. Oxford University Press, Oxford.

[Soames, 1999] Soames, S. (1999). *Understanding Truth*. Oxford University Press, New York.

[Stanley, 2005] Stanley, J. (2005). *Knowledge and Practical Interests*. Oxford University Press, Oxford.

[Tappenden, 1993] Tappenden, J. (1993). Analytic truth—it's worse (or perhaps better) than you thought. *Philosophical Topics*, 21:233–261.

[Tarski, 1944] Tarski, A. (1944). The semantic conception of truth: and the foundations of semantics. *Philosophy and Phenomenological Research*, 4(3):341–376.

[Weiner, 2009] Weiner, M. (2009). The (mostly harmless) inconsistency of knowledge attributions. *Philosophers' Imprint*, 9(1).

[White and Tarski, 1987] White, M. and Tarski, A. (1987). A philosophical letter of Alfred Tarski. *The Journal of Philosophy*, 84(1):28–32.