

## TRUTH'S SAVIOUR?

BY KEVIN SCHARP

*Saving Truth from Paradox*. BY HARTRY FIELD. (Oxford UP, 2008. Pp. xv + 406. Price £19.99.)

Hartry Field's book *Saving Truth from Paradox* is without question among the best works on truth and the liar paradox in the analytic tradition; it should become the standard reference on the liar paradox for years to come. Field offers lucid, technically accurate but accessible discussions of most of the approaches to the liar paradox that are currently being debated in the literature. He also defends his favoured approach, which requires a change from classical to paracomplete logic. After a brief flirtation with dialetheism around the turn of the century, he now offers a novel, powerful and technically dazzling way of dealing with the liar paradox to accompany his influential version of disquotationalism.<sup>1</sup> Together they provide a unified view of the nature and logic of truth.<sup>2</sup> Field's solution to the liar, together with his fair and charitable discussion of the alternatives, make this book required reading by anyone remotely interested in issues associated with truth, philosophical logic and philosophy of language.

The book covers much the same ground as several of Field's recent papers on the liar paradox,<sup>3</sup> but this is not a collection; instead, Field has written the book from

<sup>1</sup> See his *Truth and the Absence of Fact* (Oxford UP, 2001); the discussion of dialetheism is at pp. 145–6.

<sup>2</sup> These works also dovetail neatly with Field's 2008 John Locke Lectures, 'Logic, Normativity, and Rational Revisability', which focus on the normative and rational aspects of logic and the kind of rational revision in our norms of reasoning that Field thinks is required to deal with the liar paradox; however, since these lectures are currently unpublished, I shall not discuss them further.

<sup>3</sup> Field, 'A Revenge-Immune Solution to the Semantic Paradoxes', *Journal of Philosophical Logic*, 32 (2003), pp. 139–77; 'The Semantic Paradoxes and the Paradoxes of Vagueness', in J. Beall (ed.), *Liars and Heaps* (Oxford UP, 2003), pp. 262–311; 'No Fact of the Matter', *Australasian Journal of Philosophy*, 81 (2003), pp. 457–80; 'The Consistency of the Naïve Theory of Properties', *The Philosophical Quarterly*, 54 (2004), pp. 78–104; 'Variations on a Theme by Yablo', in J. Beall and B. Armour-Garb (eds), *Deflationism and Paradox* (Oxford UP, 2005), pp. 53–74; 'Truth and the Unprovability of Consistency', *Mind*, 115 (2006), pp. 567–605; 'Solving the Paradoxes, Escaping Revenge', in J. Beall (ed.), *The Revenge of the Liar* (Oxford UP, 2008), pp. 78–144.

scratch in a way that informs the reader of the many formal results contained in the papers without actually presenting the technical details. The result is very readable and pitched at a level that allows even readers with little expertise in mathematical logic to grasp the complexities in Field's assessments of the myriad solutions to the liar paradox.

Field classifies the solutions to the liar by their background logic – roughly, solutions that permit classical logic, solutions that require a weakening to paracomplete logic (which denies the law of excluded middle), and solutions that require a weakening to paraconsistent logic (which accepts some contradictions). This way of classifying solutions is illuminating and a welcome break from the received way of categorizing them according to gappy theories, revision theories, context-dependence theories, and dialethic theories. For example, Field's classification highlights the differences between his solution and superficially similar solutions like the one offered by Tim Maudlin.<sup>4</sup> In addition, the first part (chs 1–3) of the book, designed to provide the reader with the relevant background, contains excellent discussions of Tarski's theorem, Gödel's diagonalization lemma and Kripke's fixed-point construction. A real gem in this part is the chapter on the unprovability of soundness (ch. 2), which brings out a potentially surprising consequence of Kreisel's squeezing argument. Throughout the book Field's primary objection to views he rejects (see e.g., p. 210) is that they are not consistent with the intersubstitutability principle

If C and D are alike except that (in some transparent context) one has a sentence 'A' where the other has '<A> is true', then one can legitimately infer D from C and C from D.

He argues convincingly that if truth does not obey the intersubstitutability principle, then truth predicates could not play their stereotypical role as a device of generalization. The point is familiar from the literature on deflationism, and is usually accepted by deflationists and their detractors alike.

One can think of Field's favoured theory as a solution to an optimization problem: given that the intersubstitutability principle is required for truth to play its role in our linguistic practice, and that this principle is inconsistent with classical logic, what is the best non-classical logic that is consistent with it? As conditions on the 'best' non-classical logic, Field wants an intuitive conditional that is both powerful enough to reason with and renders valid all the T-sentences (even the ones for paradoxical sentences), and he wants a way of classifying liar sentences in the object-language.

His solution begins with the internal strong-Kleene theory familiar from Kripke's construction (Field calls this theory KFS); he then uses a revision construction, popularized by Gupta and Belnap,<sup>5</sup> to define an adequate conditional. The resulting logic is paracomplete in that it does not validate the law of excluded middle (LEM), but it is not a version of intuitionism (e.g., it validates double negation elimination). Field proves a conservativeness theorem (i.e., if we add 'true' to a ground language that

<sup>4</sup> T. Maudlin, *Truth and Paradox* (Oxford UP, 2004).

<sup>5</sup> A. Gupta and N.D. Belnap, *The Revision Theory of Truth* (MIT Press, 1993).

has a classical model, then there is a non-classical model of the resulting language which is identical with the ground model of the true-free fragment), and shows both that the intersubstitutability principle holds and that all the T-sentences (in the object-language) are valid (chs 15–17). Field's conditional acts just like a material conditional when LEM is assumed. Moreover, it allows him to define a determinateness operator  $D$  ('DA' is defined as ' $A \wedge \neg(A \rightarrow \neg A)$ ', p. 236), which can be used to classify the liar sentences in the object-language as not determinately true and not determinately false. Furthermore, the determinateness operator iterates non-trivially, so it can even be used to classify liar-type sentences that contain occurrences of the determinateness operator. For example, a sentence  $Q$  that is provably equivalent to ' $Q$  is not determinately true' is not determinately determinately true. Indeed, by iterating the determinateness operator, one can generate a transfinite hierarchy of determinateness operators. It is a delicate issue just how far this hierarchy extends, since the language in question also contains a truth predicate, which can be used to generalize over the determinateness operators. Field argues (chs 15, 16, 22) that the hierarchy eventually breaks down, but in the interesting cases, the point of breakdown is indeterminate. Thus according to Field, the determinateness operators serve the purpose of classifying liar-type sentences without giving rise to pesky 'revenge' paradoxes that plague other solutions. I regret that a critical study of this length cannot do justice to the technical details of Field's elegant paracomplete logic, his conditional, or the hierarchy of determinateness operators.

On the philosophical side, we can think of Field's solution to the liar as part of a broader project on indeterminateness that goes back to the early 1970s (see chs. 6–10 of *Truth and the Absence of Fact*). According to Field, the mistake in the reasoning that leads to the liar paradox is assuming that LEM holds of truth claims in general. Moreover, he thinks that indeterminateness is also the culprit in the paradoxes of definability, the sorites paradox, and the property version of Russell's paradox, and he proposes essentially the same solution to them as well. These aspects of his view are touched on briefly in the book (chs 5, 9, 20), but they are not given the attention they deserve. In fact, I found the scattered discussions of vagueness distracting, since it is such a complex problem with a huge literature, and Field's approach to it does not do anything to support his solution to the liar. However, the discussion of the naïve theory of properties, relations and propositions is relevant. One especially important result of the theory of propositions (as 0-place properties) is the existence of paradoxical propositions (p. 296). This is a significant accomplishment, given that it is so easy to find oneself inadvertently believing, asserting or reasoning from something paradoxical. The received view has been that there are no such propositions, but Field points out that this view is highly counter-intuitive, and offers a solution to the problem.

In what follows, I shall express some doubts about whether Field's solution really is the Ascalon he makes it out to be. One issue is whether Field's solution is supposed to be descriptive or prescriptive. That is, is it supposed to describe a natural language and the way people actually use it, or is it supposed to be a suggestion for how to change our language and practice? Field does address this issue, but not in detail. He seems (p. 15) to think that adopting a paracomplete logic would

be a change in our logic, and although he is not interested in the question of whether changing our logic or the principles that govern our truth predicate constitutes a change in meaning, when pressed, he writes ‘perhaps the best thing to say is that all the principles leading to paradox are “meaning constituting” ... In that case, the ordinary meanings are jointly incoherent, and a “change of meaning” is required to restore coherence’ (p. 17). These claims support the prescriptive reading of Field’s solution, which leaves us with a question: what is the best way to explain the semantic features of our language as we currently use it? Perhaps Field thinks our language is trivial (i.e., every sentence is true and false), or maybe he thinks our language only seems classical. It is hard to tell, and his views on this topic affect how we assess his prescriptive solution. At the very least, I would have liked to see Field engage with these issues instead of dismissing them so quickly.

A related issue is Field’s discussion of whether his solution requires expressive limitations. It is no secret that seemingly legitimate linguistic expressions like Boolean negation, the intuitionist conditional and idempotent determinateness operators (and many others) are not expressible in a paracomplete language (with a truth predicate) upon pain of triviality. Field discusses each of these examples,<sup>6</sup> and concludes in each case that although the expressions in question are meaningful, they are incoherent. To his credit, he does not adopt the wildly implausible view that these expressions are simply meaningless.

I was disappointed with these discussions, mainly because the rules of the game are far from clear. Field seems to think that the onus is on his opponent to show that these expressions are coherent. However, that is hardly fair, since Field permits himself to assume that truth is coherent, and he uses that assumption to argue that the other expressions are incoherent. In addition, he claims that any argument for the coherence of Boolean negation begs the question by assuming that Boolean negation is coherent. Of course, we know that we cannot argue for the coherence of our basic logical concepts without begging the question, so Field’s reasoning here just seems confused. The important thing to recognize is that if speakers of natural language use these expressions and have good reason to use them (and I think it very plausible that they do), then Field’s solution will have a bunch of hidden costs. Further, there is good reason to think that any solution to the liar requires treating certain meaningful expressions as incoherent; so we need to have a debate about how to understand such expressions, and we need to discuss whether it makes the most sense to treat truth as incoherent, as theorists like Matti Eklund and myself think,<sup>7</sup> or follow Field in treating these other expressions as incoherent. Again Field takes a dismissive attitude (pp. 8, 309 fn. 1), instead of recognizing how crucial this issue is to disputes about the liar.

Another problem for Field’s solution comes when we try to apply his theory to a natural language like English. Field readily admits that we use ‘true’ to express disagreements (by saying that a sentence or theory is not true), and some of his

<sup>6</sup> See pp. 309–12, 312–24, 343–6, respectively; presumably an incoherent expression has incompatible meaning-constitutive principles.

<sup>7</sup> M. Eklund, ‘Inconsistent Languages’, *Philosophy and Phenomenological Research*, 64 (2002), pp. 251–75; K. Scharp, ‘Replacing Truth’, *Inquiry*, 50 (2007), pp. 606–21.

criticisms of alternative solutions (e.g., pp. 205–8) depend on the fact that they cannot accommodate this usage. However, Field's own solution falls prey to a version of this problem. For example, he disagrees with paraconsistent dialetheism (i.e., the view that some sentences are true and false), but he cannot use a truth predicate to express his disagreement since ‘“some sentences are true and false” is not true’ comes out as indeterminate on Field's view. Instead, he has to say that the sentence in question is not *determinately* true. In other words, he takes some of the work traditionally given to the truth predicate and outsources it to his determinateness operators. The problem for Field's view comes when we think about how to interpret an ordinary speaker of English who has never heard of Field's solution or even the liar paradox itself; we are forced to make a hard choice as to how we should interpret some of that person's uses of ‘true’ to express disagreement. We have to conclude either that ‘true’ is not univocal and invariant or that some sentences containing the truth predicate which are legitimately used to express disagreement are indeterminate. It seems to me that either Field's theory is seriously revisionist (not just about our logic, but about our use of ‘true’), which even in the best-case scenario would leave us without a theory of truth *as we use it*, or he is forced to treat truth predicates as ambiguous or context-dependent. Both of these alternatives have major costs associated with them. The first makes it seem that Field's points about how we use truth predicates are just cherry-picked to support his theory. The second is even worse, since in everyday conversational contexts, speakers have no idea whether the sentences they are discussing might be paradoxical. If the semantic features of a sentence (e.g., whether an occurrence of ‘true’ expresses the concept of truth or a concept of determinate truth) are not in general available to competent speakers, then it seems impossible to make sense of how they can communicate successfully.<sup>8</sup> I think this is a major problem for Field, but he does not address it, and I can only gesture at it here.

A final worry concerns Field's discussion of validity as truth preservation, which is a theme that runs throughout the book (chs 2, 12, 19, 26). He argues convincingly that no theory of truth will be able to accept that all and only valid arguments preserve truth, since any theory that accepts this claim will fall foul of Gödel's second incompleteness theorem (i.e., it will be able to prove its own consistency). Field argues that this fact is not embarrassing for a theory of truth, and that instead one can say that valid arguments preserve truth *when it matters*, i.e., when it comes to non-paradoxical sentences: ‘the paracomplete theorist can and should take the same line as the weakly classical theorist as regards restricted truth-preservation: there is no reason to doubt that the rules preserve truth *when applied to sentences we're committed to*’ (pp. 287–8). I can imagine two ways of interpreting ‘sentences we're committed to’: (i) theorems of the theory of truth, or (ii) sentences one believes or entertains. If (i) is correct, then being truth preserving when it matters is not very important, since one reasons from many things that are not theorems of the theory, and presumably if truth preservation is important, then one would like those

<sup>8</sup> I take it that this worry is similar to one Kripke expresses about the orthodox Tarskian approach to the liar: see S. Kripke, ‘Outline of a Theory of Truth’, *Journal of Philosophy*, 72 (1975), pp. 690–716, at pp. 695–6.

inferences to be truth preserving. If (ii) is right, then the theory is not truth preserving when it matters, since one might inadvertently believe something contingently paradoxical. So being truth preserving when it matters is either not important or not a feature of these theories. Either way, it does not make much sense to offer restricted truth preservation as a consolation prize.

Despite my misgivings about Field's solution, his book is a breathtaking achievement. His work is challenging in all the right ways, and raises the bar on discussions of truth, the liar, and the role of mathematical logic in illuminating natural languages. It is, quite simply, the state of the art.

*The Ohio State University*