Erkenntnis (2005) 62: 47–69

IGOR DOUVEN

A PRINCIPLED SOLUTION TO FITCH'S PARADOX

ABSTRACT. To save antirealism from Fitch's Paradox, Tennant has proposed to restrict the scope of the antirealist principle that all truths are knowable to truths that can be consistently assumed to be known. Although the proposal solves the paradox, it has been accused of doing so in an *ad hoc* manner. This paper argues that, first, for all Tennant has shown, the accusation is just; second, a restriction of the antirealist principle apparently weaker than Tennant's yields a non-*ad hoc* solution to Fitch's Paradox; and third, the alternative is only apparently weaker than, and even provably equivalent to, Tennant's. It is thereby shown that the latter is not *ad hoc* after all.

1. FITCH'S PARADOX AND TENNANT'S SOLUTION

Fitch's Paradox (Fitch 1963) shows that from the antirealist principle that all truths are *knowable*, i.e.,

(1) $\forall\varphi(\varphi \rightarrow \Diamond K(\varphi))$,

it follows that all truths are *known*, i.e.,

(2) $\forall\varphi(\varphi \rightarrow K(\varphi))$.

(Here "$K(\varphi)$" is short for "$\exists S \exists t K(\varphi, S, t)$", which in turn stands for "someone at some time knows φ".) The paradox assumes classical logic plus the following rules of inference:

$$\frac{\Diamond\varphi \quad \begin{matrix}[\varphi]^i \\ \vdots \\ \bot\end{matrix}}{\bot}{\scriptstyle i}\Diamond\bot \qquad \frac{K(\varphi)}{\varphi}F \qquad \frac{K(\varphi \wedge \psi)}{K(\varphi)}D$$

The rule $\Diamond\bot$ corresponds to the modal principle that absurd propositions are impossible (cf. Tennant 1997, p. 257); F corresponds to the assumption that knowledge is factive (knowledge requires truth); and D corresponds to the assumption that knowledge distributes across conjunctions.[1]

Given these rules, the argument runs as follows:[2]

$$\dfrac{\dfrac{[\exists\varphi(\varphi \wedge \neg K(\varphi))]^1 \quad \dfrac{\dfrac{[\varphi \wedge \neg K(\varphi)]^2 \quad \dfrac{\forall\varphi(\varphi \to \Diamond K(\varphi))}{(\varphi \wedge \neg K(\varphi)) \to \Diamond K(\varphi \wedge \neg K(\varphi))}\,{\scriptstyle \forall E}}{\Diamond K(\varphi \wedge \neg K(\varphi))}\,{\scriptstyle \to E} \quad \dfrac{\dfrac{\dfrac{[K(\varphi \wedge \neg K(\varphi))]^3}{K(\neg K(\varphi))}\,{\scriptstyle D}}{\neg K(\varphi)}\,{\scriptstyle F} \quad \dfrac{[K(\varphi \wedge \neg K(\varphi))]^3}{K(\varphi)}\,{\scriptstyle D}}{\bot}\,{\scriptstyle \to E}}{\bot}\,{\scriptstyle 3\Diamond\bot}}{\bot}\,{\scriptstyle 2\exists E}}{\neg\exists\varphi(\varphi \wedge \neg K(\varphi))}\,{\scriptstyle 1 \to I}$$

By elementary logic, the conclusion is equivalent to (2). Anyone committed to (1) will thus have to buy (2) as well. While this is not a paradox in the strict sense – it does not show (1) to be inconsistent – it does show (1) to have a consequence that is hard to believe. Naturally that is worrying enough for defenders of the principle.

To thwart the threat to antirealism posed by Fitch's result, Tennant (1997, Chapter 8) proposes to replace (1) by the weaker claim that all true *Cartesian* propositions are knowable, where a proposition is Cartesian precisely if the assumption that it is (or was, or will be) known is consistent (a proposition that is not Cartesian he calls *anti-Cartesian*). That is to say, Tennant's proposal is to replace (1) by

(3) $\varphi \to \Diamond K(\varphi)$, for all φ such that $K(\varphi)$ is consistent.

With (3) in place of (1) the above proof no longer goes through. After all, as the subproof dependent on $K(\varphi \wedge \neg K(\varphi))$ shows, the assumption that $\varphi \wedge \neg K(\varphi)$ is known is inconsistent. Thus the instantiation of $\forall\varphi(\varphi \to \Diamond K(\varphi))$ by $\varphi \wedge \neg K(\varphi)$, which occurs in the proof, is no longer permissible as the latter proposition is anti-Cartesian and therefore does not satisfy the proviso that now goes with the former.

But whereas it is clear that Tennant's proposal blocks the paradox, it is not equally clear that it does so in a principled or non-*ad hoc* way. In fact, Hand and Kvanvig (1999) dispute that it does. In the critical part of this paper I argue, first, that Hand and Kvanvig's critique is not very telling because it is based on an overly strict criterion for non-*ad hoc*ness (Section 2), and second, that, for all Tennant has shown, (3) comes out *ad hoc* also on a more reasonable conception of non-*ad hoc*ness (Section 3). In the constructive part I then present a restriction on (1) that, given two assumptions to be specified, can be shown to block Fitch's Paradox in a non-*ad hoc* manner (given the more reasonable understanding of non-*ad hoc*ness; Section 4); in motivating the assumptions, I draw upon the writings of, among others, DeRose, Unger, and Williamson concerning the relations between knowledge, belief, and assertion. In that part I further show that, given the same assumptions, the restriction I propose is in effect equivalent to Tennant's (Section 5). I close by briefly evaluating the proposed solution from a realist perspective (Section 6).

2. WHAT MAKES A RESTRICTION PRINCIPLED?

Hand and Kvanvig are quite explicit about what, according to them, must be done to solve Fitch's Paradox in a non-*ad hoc* way. They say:

> To address the paradox in a philosophically substantive way, one must go beyond such arbitrary approaches [as simply excluding from the range of (1) the truths that lead to problems]. Realists do this by observing that truth is "radically nonepistemic", thereby giving themselves a reason based on their conception of truth for denying (1). Tennant must do something comparable. We should expect him to find some feature of truth, antirealistically conceived, that disarms the paradox by allowing some truths to be unknowable. (Hand and Kvanvig 1999, p. 423)

In other words, if a restriction strategy such as Tennant's is to qualify as non-*ad hoc*, it ought to be motivated by a reason that is independent of Fitch's Paradox and that emanates from the antirealist conception of truth. And Tennant has failed to provide such a motivation, Hand and Kvanvig think.[3]

Suppose Tennant has indeed failed to do so. Then how bad is that for his proposal? In order to answer this question, we must look a bit closer at the criterion for non-*ad hoc*ness Hand and Kvanvig suppose. I agree with their requirement that there be an independent reason for restricting (1) in whatever way one proposes to restrict it. I also agree that an independent reason alone is not enough. Rather than reconfirmation *that* there is something wrong with (1), we seek some understanding of *what* is wrong with it. So, the independent reason should have explanatory cash value, it should be "one that allows an explanation of [(l)]'s failure" (Hand and Kvanvig 1999, p. 426). But this is where my agreement with them ends. In particular, I fail to see why the requisite reason has to be a feature of truth, why one must "cite something about one's conception of truth that calls for [the given restriction]", as Hand and Kvanvig say (*ibid.*). Why could it not be something about one's conception of, for instance, knowledge that explains what is wrong with (1)? Or something about the concept of belief, or about that of justification (or about both), which, at least on most current analyses of knowledge, are involved in (1) via the concept of knowledge? Suppose that we are given an excellent, explanatory reason for restricting (1) in a particular way; that, thus restricted, the principle blocks the derivation of Fitch's Paradox; but that the reason given has to do with the concept of (say) justification. Would we in that case still lack a principled solution to the paradox? Should the antirealist still provide a further reason for the restriction, this one based on her conception of truth?

It might be suggested that the foregoing *would* count as a reason for restricting (1) on Hand and Kvanvig's criterion, given that antirealist truth crucially involves the concept of knowledge, which in turn involves that of justification; the same would hold, for instance, for a reason that had to do with the concept of belief. I suspect that this is not how Hand and Kvanvig want their demand for providing a reason related to antirealist truth to be understood, but should they agree with the suggestion, then I may be in perfect agreement with them after all. For I submit the following as a reasonable criterion for the non-*ad hoc*ness of proposals such as Tennant's:

> In order to qualify as principled or non-*ad hoc*, it is necessary and sufficient that a proposal for restricting (1) in a particular way be accompanied by a reason for adopting it other than its capability to solve the paradox, and that reason must be related, in an informative or explanatory way, to one or more of the concepts that are either implicitly or explicitly involved in (1).[4]

I should also note that, if this is accepted, then there might be nothing specifically antirealist about the reason we have for restricting (1). The realist does not, or at least need not, have a conception of, for instance, belief different from that of the antirealist. And it is possible that a conception of belief shared by both parties impels us to revise (1) in a way which also solves Fitch's Paradox.

While Tennant is less explicit than Hand and Kvanvig are, or than I have just tried to be, about what it takes for a restriction strategy to be principled, from his (2001) reply to Hand and Kvanvig's paper we can safely infer that he would reject the criterion just proposed. For in that paper he argues that (3) is "substantive, informative and important", and that this is enough to render "[t]he objection that the restriction invoked is *ad hoc* ... groundless" (Tennant 2001, p. 111).[5] Let us grant the first part of this claim and focus on the sufficient condition for non-*ad hoc*ness supposed by the second – a restriction on a thesis is non-*ad hoc* if the thus restricted thesis is substantive, etc. To my eye, this condition is far too permissive. For take any substantive, informative, important, but paradoxical principle P. Then presumably a restricted version P' of it that applies whenever P applies except in those cases in which the latter leads to paradox will be no less substantive, informative, or important than P itself. But, intuitively, we can still distinguish between *ad hoc* and non-*ad hoc* ways of obtaining P' from P. For instance, if the restricted principle were obtained simply by declaring that principle P applies except in cases in which its application leads to paradox, then it would certainly strike us as being *ad hoc*, its supposed substantiveness, etc., notwithstanding.

If needed, many examples from the history of mathematics or from that of analytic philosophy could be adduced to buttress the claim that it is one thing for a restricted theory or principle to be substantive, etc., and that it is quite another for the restriction imposed upon the theory or principle to be non-*ad hoc*. I mention but three.

One would be hard put to find among working mathematicians or philosophers of mathematics someone willing seriously to deny that ZF set theory is a substantive, informative, and important theory. Yet not few in those quarters appear to share Putnam's (2000, p. 24) opinion that the Axiom Scheme of Replacement – which is more or less ZF's substitute for the Unrestricted Comprehension Scheme of naive set theory, and which does the real work in blocking the paradoxes that plagued the naive theory – is a mere "formal maneuver" for which there is "no intuitive basis at all".

Similarly for Tarski's theory of truth. Tennant (2001, p. 111) is surely right to suggest that the theory is substantive, informative, and important.[6] However, his claim that by restricting truth predicates to language levels "Tarski can hardly be accused of making an *ad hoc* restriction" (*ibid.*) seems wrong. As a matter of fact, many philosophers have actually made the accusation. See for instance Quine (1961, p. 9), who calls the language-level restriction a "desperate resort" and an "artificial departure from natural and established usage." Putnam likewise thinks it is *ad hoc*, calling the restriction a "desperate device" (1990, p. 16) and "just a technical solution" (2000, p. 5). For similar criticisms see, among many others, Haack (1978, p. 144), Fox (1989), Kirkham (1992, p. 281), and Glanzberg (2004).

As a final example, consider Horwich's minimalist theory of truth, which avoids the Liar Paradoxes by being expressly limited to the "*uncontroversial* instances of the equivalence schema [It is true that p if and only if p]" (Horwich 1990, p. 7; italics mine, original italics omitted). To my knowledge, no one has objected that the theory fails to be substantive or informative or important because of how it dodges paradox. But does that mean Horwich could legitimately claim to have given a principled solution to the Liar Paradoxes? Horwich (1990, p. 41 ff) himself, at any rate, makes it quite clear that he is not under the impression to have done so, or even to have given anything at all worthy of the name "solution to the Liar Paradoxes". And, significantly, none of the many commentators on his work has complained that Horwich is too modest on this count.[7]

Even though in my view these examples give the lie to Tennant's conception of non-*ad hoc*ness, I must at once admit that the lack of unanimity about the status, *qua* being *ad hoc*, of many proposed solutions to paradoxes, or of the theories in which these solutions are embedded, indicates that intuitions about this matter tend to vary. In order to avoid unproductive quibbling, then, let me note that even those whose intuitions about *ad hoc*ness differ from mine, and who are inclined to agree with Tennant that (3)'s being substantive, informative, and important suffices to make the restriction it imposes on (1) non-*ad hoc*, should find it of some interest to learn that this restriction has a motivation that is independent of Fitch's Paradox and that has explanatory power, in brief, that it also qualifies as non-*ad hoc* on the non-*ad hoc*ness criterion proposed three paragraphs back, as I endeavor to show in Sections 4 and 5. First, however, I want to consider a further reason Tennant has for thinking his solution is not *ad hoc*, one that is also discussed by Hand and Kvanvig (1999, p. 424 f), and which prima facie holds some promise of satisfying our criterion.[8]

3. UNTHINKABLE TRUTHS

The further reason involves the "necessitarian" version of (1), that is,

$$(4) \qquad \Box\forall\varphi(\varphi \rightarrow \Diamond K(\varphi)).$$

Arguably, the antirealist is committed to (4), and not merely to (1).[9] After all, according to the antirealist it is a purely conceptual matter that all truths are knowable. But now consider this. Since knowing requires thinking, there can be no dispute about the validity of the following rule of inference ("$T(\varphi)$" is short for "$\exists S \exists t T(\varphi, S, t)$", which means that φ is thought by someone at some time):

$$\frac{K(\varphi)}{T(\varphi)}\,T$$

It further would seem possible that there exist no thinkers and thus that no proposition is thought, i.e., the following seems plausible:

(5) $\Diamond\neg\exists\varphi T(\varphi)$.

Nevertheless, the antirealist cannot admit as much, as can be seen thus:[10]

$$\dfrac{\dfrac{[\Diamond\neg\exists\varphi\, T(\varphi)]^1 \quad \dfrac{\dfrac{[\neg\exists\varphi\, T(\varphi)]^2 \quad \dfrac{\forall\varphi(\varphi\to\Diamond K(\varphi))}{\neg\exists\varphi\, T(\varphi)\to\Diamond K(\neg\exists\varphi\, T(\varphi))}\,\forall E}{\Diamond K(\neg\exists\varphi\, T(\varphi))}\,{\to}E \qquad \dfrac{\dfrac{\dfrac{[K(\neg\exists\varphi\, T(\varphi))]^3}{\exists\psi\, K(\psi)}\,\exists I \quad \dfrac{\dfrac{[K(\psi)]^4}{T(\psi)}\,T}{\exists\varphi\, T(\varphi)}\,\exists I}{\exists\varphi\, T(\varphi)}\,{}_4\exists E \quad \dfrac{[K(\neg\exists\varphi\, T(\varphi))]^3}{\neg\exists\varphi\, T(\varphi)}\,F}{\bot}\,{\to}E}{\bot}\,{}_3\Diamond\bot}{\bot}\,{}_2\Diamond\bot}{\neg\Diamond\neg\exists\varphi\, T(\varphi)}\,{}_1{\to}I$$

Thus it is a consequence of (4) that in *all* possible worlds – that is, even in those uninhabited by any thinkers – there are some propositions which are thought. That's a patently absurd result by any standards, and hence a further problem for the antirealist who endorses (4).[11]

It is easy to see, however, that the proposition that no proposition is thought is anti-Cartesian, for if it is known then it is thought, but if it is thought then it is false and thus cannot be known. By consequence, restricting (4) to Cartesian truths rids us not only of Fitch's Paradox but also of the further problem just presented (for that restriction makes the application of the rule $\forall$-elimination in the above proof incorrect). Tennant (1997, p. 274) seems to believe that this shows his solution to the paradox to be non-*ad hoc*. Unsurprisingly, Hand and Kvanvig (1999, p. 425) are adamant that he is wrong about this, for although the problem may give us an independent reason for restricting (4), that reason does not emanate from the antirealist conception of truth. However, we rejected their criterion as being unduly demanding. And if, as on our criterion, the requisite independent reason need not emanate from the concept of truth, but may also have to do with some other concept presupposed by (4) – such as, patently, that of thinking – then doesn't the above problem show that Tennant's solution is a principled one?

It is certainly true that the problem points the way to a perfectly good explanation of why some truths are unknowable, and hence of why (4) is wrong, namely:

> Something can be known only if it can be thought, and, as was shown above, there may exist truths which are unthinkable. Thus, either (4) wrongly assumes that, necessarily, all truths are thinkable, or it overlooks an important conceptual connection between knowing and thinking.

Even so, however, the above problem is of no help to Tennant's proposal. As will be recalled, our criterion requires any proposal for restricting (1) or (4) in a certain way to be accompanied by an independent and explanatory reason for adopting *it*, and not just by an independent and explanatory reason for restricting (1) or (4) (in some way or other). Differently put, what the independent reason must explain is not just why (1) and/or (4) ought to be restricted, but why they ought to be restricted in the specific way the proposal proposes. Clearly, while the above problem does the former, it does *not* explain why (4) ought to be restricted to *Cartesian* truths. After all, solving that problem, and doing justice to the insight it provides us with, requires no more than that we restrict (4) to all *thinkable* truths, i.e., more precisely, that we supplant it by the following principle:

(6) $\Box(\varphi \rightarrow \Diamond K(\varphi))$, for all φ such that $T(\varphi)$ is consistent.

And *that* principle is just not restrictive enough to prevent Fitch's Paradox from arising: no contradiction ensues from the assumption that someone at some time entertains the thought that, for example, snow is white and no one knows or knew or will ever know that snow is white. In short, though the problem involving (5) gives a reason for restricting (4) that is independent of Fitch's Paradox, and that is even related in the right way to a concept presupposed by (4), it does *not* give a reason for adopting the particular restriction Tennant proposes. It thus fails to show that his proposal is not *ad hoc*, not only on Hand and Kvanvig's non-ad *hoc*ness criterion, but also on ours.

4. A PRINCIPLED SOLUTION TO FITCH'S PARADOX

If all truths are knowable, then, given that knowledge requires justified belief, all truths must be justifiedly believable. But there are truths, among them ones of the form $\varphi \wedge \neg K(\varphi)$, that cannot be so believed. This claim plays a pivotal part in the solution to Fitch's Paradox to be offered below. Since my argument for it builds on Tennant's analysis of Moore's Paradox, I start by describing that in some detail.

Consider the following instance of Moore's Paradox:

(7) It's raining, but I don't believe that it's raining.

As has been amply noted in the literature, (7) is not inconsistent. If it were, then we could infer "It's not raining" from "I don't believe it's raining", which obviously we can *not*. Still, it seems clear that there is *something* amiss with (7). On Tennant's (1997, p. 251 f) diagnosis, while (7) is not itself inconsistent, by *asserting* it the speaker does imply a contradiction. Of course this is not only so on his diagnosis; the foregoing represents more or less the customary way of thinking about Moore's Paradox nowadays. Nevertheless, Tennant's analysis of the paradox is specially noteworthy because, firstly, it aims to make plausible not only that (7) and other instances of the paradox cannot be consistently asserted but also that they cannot be consistently believed, and secondly and at least as importantly, it aims to do so in a neatly formal fashion.

To achieve the latter goal, Tennant needs three extra rules. The first he calls the *rule of rational commitment* (Tennant 1997, p. 247):

$$\cfrac{B(\varphi_1, S, t), \ldots, B(\varphi_n, S, t) \quad \begin{matrix}\underbrace{[\varphi_1]^i, \ldots, [\varphi_n]^i} \\ \vdots \\ \psi\end{matrix}}{B(\psi, S, t)}\,{}_iRC$$

("$B(\varphi, S, t)$" stands for "S at t believes φ"; I will sometimes use "$B(\varphi)$" as shorthand for "$\exists S \exists t B(\varphi, S, t)$".) Note that it follows from *RC* that belief distributes across conjunctions. Thus the following rule is weaker than *RC*:[12]

$$\frac{B(\varphi \wedge \psi, S, t)}{B(\varphi, S, t)}\,BD$$

In the proofs below we will not need *RC* and can make do instead with *BD*.

The second rule is what Tennant (1997, p. 248) calls the *rule of credibility*:

$$\cfrac{\begin{matrix}\underbrace{B(\varphi, S, t), [\varphi]^i} \\ \vdots \\ \bot\end{matrix}}{\bot}\,{}_iC$$

It is the natural-deduction analogue of the thought that a belief claim is consistent only if it is consistent with the truth of the belief's content.

Let "$\alpha(\varphi, S, t)$" stand for "S at t asserts φ". Then the final additional rule, called αB, is this (Tennant 1997, p. 250):

$$\frac{\alpha(\varphi, S, t) \quad \begin{matrix}[B(\varphi, S, t)]^i \\ \vdots \\ \bot\end{matrix}}{\bot}\,{}_i\alpha B$$

Its rationale, as Tennant explains (*ibid.*), is that assertion betokens belief, and that hence someone who asserts φ "is rationally committed to (the consequences of) our taking him to believe that φ". So, in particular, if it is inconsistent for a person to believe φ, then it is inconsistent for her to assert φ, too. Accordingly, αB allows us to hold accountable an asserter for any inconsistency that follows from the supposition that she believes what she asserts.

It should be noted that RC, C, and αB all involve some idealization. It is not true for most of us, or even any of us, that we are able to recognize all logical consequences of our beliefs, let alone that our beliefs are closed under logical consequence; according to RC they are, however. And we normally would not want to say that it is *inconsistent* to believe a necessary falsehood that none of us (nonideal reasoners) is capable of recognizing as such; yet according to C it is. As to the third rule, while S's asserting φ at t may commit her to an inconsistency if the supposition that S at t believes φ is inconsistent, it does not seem to follow what according to αB we may conclude in that case, namely, that the supposition that S at t asserts φ is inconsistent – *unless*, of course, S can be assumed to be a rational individual who would never do anything that could commit her to an inconsistency and who, in addition, will notice whenever the supposition that she believes a given proposition at a given time is inconsistent.

Tennant (1997, p. 248 ff) argues at some length, and convincingly to my mind, that in the present context it is legitimate to assume that we are rational thinkers with cognitive powers transcending those we actually possess. Nevertheless, the assumption, specifically the part concerning our cognitive powers (and not the rationality part), has recently been criticized by DeVidi and Kenyon (2003). One way to respond to their critique would be to restrict application of each of the above rules to propositions φ such that the logical relation the given rule requires us to perceive is *obvious*; so, for instance, one could restrict C to propositions φ such that $\varphi \wedge B(\varphi)$ is *obviously* inconsistent.[13] It appears that the rules thus restricted suffice for both Tennant's and our concerns on any intuitively plausible understanding of obviousness, considering that the applications of the rules in proofs to be given below only involve propositions to which they very obviously apply. Of course to maintain their status as rules, we would have to make the notion of obviousness precise. I will not bother to do so here and instead will go along with Tennant's idealization, as have done almost all commentators on his work.

With *BD*, *C*, and αB it is easy to demonstrate that instances of Moore's Paradox are not consistently assertible, as follows (Tennant, 1997, p. 252):

$$\dfrac{\alpha(\varphi\wedge\neg B(\varphi,S,t),S,t) \qquad \dfrac{\dfrac{\dfrac{[B(\varphi\wedge\neg B(\varphi,S,t),S,t)]^1}{B(\varphi,S,t)}{\scriptstyle BD} \quad \dfrac{[\varphi\wedge\neg B(\varphi,S,t)]^2}{\neg B(\varphi,S,t)}{\scriptstyle \wedge E}}{\bot}{\scriptstyle \to E}}{\bot}{\scriptstyle 2C}}{\bot}{\scriptstyle 1\alpha B}$$

And, as the application of the rule of credibility exhibits, instances of Moore's Paradox cannot be consistently believed either.

If we agree on this, then, given that undoubtedly there exist true instances of Moore's Paradox, we can conclude that there are truths that I cannot consistently believe. Arguably, what I cannot consistently believe I cannot be justified in believing. Since I cannot know what I cannot justifiedly believe, it follows that there exist true instances of Moore's Paradox that I cannot know. Clearly enough, it does *not* follow that there are truths that cannot be known by anyone, at any time. Many have noted that "It's raining, but John does not believe that it is", is not an instance of Moore's Paradox and not philosophically problematic at all; similarly for "It's raining, but I did not believe it".[14] Consequently, the proposition expressed by (7) *can* be consistently believed – by someone else, or by me at some other time. This may make Moore's Paradox seem largely irrelevant to Fitch's Paradox, which says that some truths cannot be known *by anybody*, *ever*. But it is not, for consider the following generalization of it:

(8) φ, but no one believes (or believed, or will ever believe) φ.

We need only two of Tennant's three extra-logical rules to show that instances of (8) can never be consistently believed, by anyone:

$$\dfrac{\dfrac{\dfrac{B(\varphi\wedge\forall S'\,\forall t'\neg B(\varphi,S',t'),S,t)}{B(\varphi,S,t)}{\scriptstyle BD} \quad \dfrac{\dfrac{\dfrac{[\varphi\wedge\forall S'\,\forall t'\neg B(\varphi,S',t')]^1}{\forall S'\,\forall t'\neg B(\varphi,S',t')}{\scriptstyle \wedge E}}{\forall t'\neg B(\varphi,S,t')}{\scriptstyle \forall E}}{\neg B(\varphi,S,t)}{\scriptstyle \forall E}}{\bot}{\scriptstyle \to E}}{\bot}{\scriptstyle 1C}$$

Thus it is provably the case that $\neg B(\varphi \wedge \neg B(\varphi))$, for all φ. Since no doubt there are many humdrum truths that will never even be entertained by anyone, we may assume that (8) has true instances. As we just saw, however, these cannot be consistently believed. But that means they cannot be known either. Hence there are unknowable truths.

So, already in the face of our generalized version of Moore's Paradox, (4) (and (1), for surely there are *actually* true instances of (8)) cannot be maintained, at least not unadorned. That gives us another problem for (4) independent of Fitch's Paradox, and again it is one that clearly suggests an explanation of what is wrong with the principle, to wit:

> Something can be known only if it can be justifiedly, and thus consistently, believed. And, as the above generalized version of Moore's Paradox shows, there exist truths which cannot be consistently believed. Hence, either (4) wrongly assumes that all truths can be consistently believed, or it overlooks an important conceptual connection between knowing and consistently believing.

This immediately suggests an emendation of (4), namely, to restrict it to truths that are consistently believable, as follows:

(9) $\Box(\varphi \rightarrow \Diamond K(\varphi))$, for all φ such that $B(\varphi)$ is consistent.

What I want to show now is that (9) is impervious not only to the generalization of Moore's Paradox but also to Fitch's Paradox. In doing so, I make use of two extra premises. Both will be assumed without argument, which seems excusable given that they have received extensive and – not only in my view – forceful defenses elsewhere in the philosophical literature.

The first premise is the so-called *knowledge account of assertion*. It is hinted at by Moore (1962, p. 277) when he says that "by asserting *p* positively you *imply*, though you don't assert, that you know that *p*", but it is only fully developed in Unger (1975). There the gist of the account is presented as the claim that "asserting that something is so entails not just representing the thing as being so, but representing oneself as *knowing* that it is" (Unger 1975, p. 256). Unger is able to provide an impressive amount of linguistic evidence supporting this claim. For instance, as he points out, it accounts for the fact that the question "How do you know?" is a normal and socially perfectly acceptable response to an assertion and the further fact that "the asserter [cannot] get off the hook by saying 'I never said I knew it'" (Unger 1975, p. 263 f). This and other evidence has convinced many analytic philosophers that the knowledge account of assertion is correct.[15]

By an argument parallel to the one that led Tennant to his rule αB, the knowledge account of assertion can be seen to warrant a crucial strengthening of that rule. Arguably, by representing things as being a certain way, a person commits herself to the supposition that things are that way. In particular, by representing herself as knowing a certain thing, a person commits herself to the supposition that she knows the thing. Thus, if Unger is right that in asserting something a person represents herself as knowing it, then if it is inconsistent to suppose that S at t knows φ, S commits herself to an inconsistency by asserting φ at t, or, in natural-deduction format:

$$\dfrac{\alpha(\varphi, S, t) \quad \begin{matrix}[K(\varphi, S, t)]^i \\ \vdots \\ \bot\end{matrix}}{\bot}\ {}_i\alpha K$$

In other terms, closer to Moore's: since by asserting φ at t the person S implies that she knows φ at t, by asserting φ at t she implies an inconsistency if the supposition that she knows φ at t is inconsistent.

The second extra premise I call, following Adler (2002), the *belief–assertion parallel*. It is the claim that belief is a species of assertion, namely, subvocalized assertion or, as Adler (2002, p. 74) puts it, "assertion to oneself". Williamson (2000) seems to be making basically the same assumption when he says that "assertion is the exterior analogue of judgement, which stands to belief as act to state" (p. 238), and that "occurrently believing p stands to asserting p as the inner stands to the outer" (p. 255).[16] However, the most elaborate defense of the assumption is to be found in Adler (2002, Chapters 5 and 10). For our purposes here, the important consequence of this parallel is that it is not only the case that, as Tennant already assumed, if it is inconsistent to believe φ, then it is inconsistent to assert φ, but that the converse holds as well:

(10) If it is inconsistent to assert φ, then it is inconsistent to believe φ.

After all, what one cannot consistently assert, one cannot consistently assert to oneself, and hence, given the belief–assertion parallel, one cannot consistently believe either.[17]

Now first note that with the rule αK added to those we already had we can readily derive that propositions of the form $\varphi \wedge \neg K(\varphi)$ are not consistently assertible by anyone at any time:

$$\dfrac{\alpha(\varphi \wedge \neg\exists S'\,\exists t'K(\varphi, S', t'), S, t) \qquad \dfrac{\dfrac{\dfrac{\dfrac{[K(\varphi \wedge \neg\exists S'\,\exists t'K(\varphi, S', t'), S, t)]^1}{K(\varphi, S, t)}\,{\scriptstyle D}}{\exists t'K(\varphi, S, t')}\,{\scriptstyle \exists I}}{\exists S'\,\exists t'K(\varphi, S', t')}\,{\scriptstyle \exists I} \qquad \dfrac{\dfrac{[K(\varphi \wedge \neg\exists S'\,\exists t'K(\varphi, S', t'), S, t)]^1}{\varphi \wedge \neg\exists S'\,\exists t'K(\varphi, S', t')}\,{\scriptstyle F}}{\neg\exists S'\,\exists t'K(\varphi, S', t')}\,{\scriptstyle \wedge E}}{\bot}\,{\scriptstyle \to E}}{\bot}\,{\scriptstyle 1\alpha K}$$

And from this, by invoking (10), we infer that propositions of the form $\varphi \wedge \neg K(\varphi)$ cannot be consistently believed by anyone at any time. So it is provably the case that $\neg B(\varphi \wedge \neg K(\varphi))$, for all φ. As a result, propositions of the form $\varphi \wedge \neg K(\varphi)$ do not satisfy (9)'s proviso, which in turn means that we cannot instantiate $\forall\varphi(\varphi \rightarrow \Diamond K(\varphi))$ by $\varphi \wedge \neg K(\varphi)$ in Fitch's proof, whereby the paradox is blocked.

It may be of some interest to see how we could have reached the same conclusion about propositions of the form $\varphi \wedge \neg K(\varphi)$, though not in an equally rigorous fashion, by mounting an argument parallel to Unger's analysis of Moore's Paradox.[18] As Unger (1975, p. 257) remarks, since in asserting something one represents oneself as knowing the thing, one thereby *ipso facto* represents oneself as *at least* believing what one asserts. With this said, Unger is able to offer the following explanation of what is wrong with (7):

> We first think of an utterer [of (7)] as asserting that it's raining and, so, as representing himself as at least believing that it is. We then think of this utterer as going on to assert that he doesn't even believe that it is and, so, as going on to represent himself as not even believing the thing. We thus think of the utterer as representing all of the following as being the case: that he at least believes that it's raining and he doesn't even believe that it is. But, then, as this last is quite clearly inconsistent, we think of the utterer as *representing something inconsistent* as being the case. (Unger 1975, p. 258)

But if, as per Unger's assumption, in asserting a proposition one represents oneself as knowing it, and not merely as believing it, then an argument showing that propositions of the form $\varphi \wedge \neg K(\varphi)$ are not consistently assertible suggests itself. For in asserting the first part of such a proposition, one would be representing oneself as knowing a particular thing. In asserting the second part, one would be representing it as being the case that no one knows (or knew or will know) that very same thing (by the part of Unger's assumption that says that in asserting something one represents it as being the case). Hence, in asserting the proposition as a whole, one would be representing oneself both as knowing something and (by instantiation) as not knowing that thing. One thereby would be representing something inconsistent as being the case. Accordingly, it is not possible to consistently assert $\varphi \wedge \neg K(\varphi)$ for any φ and so, again by (10), not possible to consistently believe it either.[19]

To sum up thus far, we have a solution to Fitch's Paradox that has a motivation that both is independent from the paradox and gives an explanation of what is wrong with (4): the solution is to restrict (4) to truths that are consistently believable; the independent motivation for it is that it is able to solve the problem the generalized version of Moore's Paradox creates for (4); and the explanation it provides us with is the one displayed just above (9) which points to the conceptual ties between knowledge and consistent belief. In other words, we have a solution to Fitch's Paradox that satisfies our criterion of non-*ad hoc*ness.[20]

5. TENNANT'S PROPOSAL RESCUED

We must still make good on our claim, made at the outset, that, given the assumptions called upon by our solution, that solution can be proved to be equivalent to Tennant's, or as we can now say more precisely, (9) is equivalent to the principle we obtain by restricting (4) to Cartesian truths, as Tennant urges we do. At first blush, this may seem implausible. For let us call "warrant" whatever it is that turns true belief into knowledge,[21] and let "$W(\varphi, S, t)$" mean that the person S at time t has warrant for believing φ. Then it holds, formally put, that $\exists S \exists t K(\varphi, S, t) \leftrightarrow [\varphi \wedge \exists S \exists t(B(\varphi, S, t) \wedge W(\varphi, S, t))]$ So obviously, if $B(\varphi)$ is inconsistent, then $K(\varphi)$ must be inconsistent as well, i.e., contrapositively,

(11) If $K(\varphi)$ is consistent, then so is $B(\varphi)$.

But why should the converse hold? Couldn't we have that, for instance, $\exists S \exists t W(\varphi, S, t)$ is inconsistent, and thus $\exists S \exists t K(\varphi, S, t)$ is inconsistent, while $\exists S \exists t B(\varphi)$ is consistent? It is easy to show that, given our assumptions, the answer is negative. If we use "$\alpha(\varphi)$" for "$\exists S \exists t \alpha(\varphi, S, t)$", then (10) says that

(12) If $B(\varphi)$ is consistent, then so is $\alpha(\varphi)$.

Further, by the rule a αK, we have that

(13) If $\alpha(\varphi)$ is consistent, then so is $K(\varphi)$.

And from (12) and (13) it follows that

(14) If $B(\varphi)$ is consistent, then so is $K(\varphi)$.

Finally, combining (14) with (11) yields that

(15) $K(\varphi)$ is consistent if and only if $B(\varphi)$ is consistent.

Hence, there is no difference between restricting (4) to Cartesian truths and restricting it to truths that are consistently believable. As a result, the conclusion about (9) reached in the previous section – that supplanting (4) by (9) solves Fitch's Paradox in a principled manner – transfers to Tennant's proposal.

Lest it be said that this does little to render plausible the claim that our solution and Tennant's are equivalent, because (15) seems implausible itself, let me try to explain the result, more verbally, in terms of Adler's (2002) account of knowledge, which was one of the sources of inspiration for two of the principles involved in the above derivation. What Adler argues in his book is that, as a matter of conceptual necessity, one must regard one's beliefs as being warrantedly held. So, for instance, on his view it is not possible to believe that electrons have negative charge but that one's evidence for believing so is deficient. In particular, it follows that if it is inconsistent to suppose that a person believes to have warrant for a proposition, then, by conceptual necessity, it is inconsistent to suppose that she believes the proposition. Thus, since by the rule of credibility it follows that it is inconsistent to suppose that a person believes to have warrant for a given proposition if the supposition that she has warrant for that proposition is inconsistent, Adler's view implies that it is inconsistent to suppose that a person believes a proposition if it is inconsistent to suppose that she has warrant for the proposition. And if so, then the putative possibility considered above – that it might be consistent to suppose a person believes a given proposition even if it is inconsistent to suppose she has warrant for it – is not really a possibility.

There remains a worry to be addressed that could easily be raised by (15), namely, the worry that if the consistently believable propositions are precisely the consistently knowable ones, then our purported explanation of what is wrong with (4) may not be genuinely explanatory. For that explanation basically tells us that not all truths are knowable, and thus (4) is false, because there exist truths that cannot be consistently believed – which seems explanatory indeed. On the other hand, to be told that not all truths are knowable because there exist truths that cannot be consistently known seems much less explanatory, or even not explanatory at all. But given (15), how could the one be more explanatory than the other? How, given that equivalence, could it be more informative to be told that there exist truths that cannot be consistently believed than it is to be told that there exist truths that cannot be consistently known? And if it cannot, then we may still lack an explanation of why (4) goes wrong and, accordingly, still lack a principled solution to Fitch's Paradox.

What this worry overlooks, however, is that explanation is a "hyperintensional" notion, meaning that how we *put* a proposition may matter to its capacity to explain. Compare, for instance,

(16) Lois Lane had nothing to fear because Clark Kent is Superman,

with

(17) Lois Lane had nothing to fear because Clark Kent is Clark Kent.

Whereas the former may be truly explanatory, the latter evidently cannot be, even though "Clark Kent is Superman" and "Clark Kent is Clark Kent" have the same truth conditions (supposing "Superman" to refer rigidly) and thus express the same proposition. Or consider explanatory proofs in mathematics.[22] It is obvious that replacing any such proof by the theorem it is meant to explain does not yield another explanation of the theorem, even though the theorem and the proof express the same proposition (namely, the necessary proposition). The same is true in our case. If the assumptions we made in Section 4 hold, then "Not all truths are consistently believable" is true if and only if "Not all truths are consistently knowable" is true. But that does not mean we can always substitute one for the other in explanatory contexts without affecting the status of the explanation *qua* explanation. In particular it does not mean that if "Not all truths are consistently knowable" cannot explain why not all truths are knowable, then "Not all truths are consistently believable" cannot explain that fact either. It thus would be a mistake to think that (15) impugns our explanation of what is wrong with (4).

6. GOOD NEWS FOR THE REALIST, TOO

According to the citation from Hand and Kvanvig (1999) at the beginning of Section 2, realists explain what is wrong with (1) by arguing that truth is radically non-epistemic. Thereby they have given a principled solution to Fitch's Paradox, or so Hand and Kvanvig suggest. I doubt that that is right, however. What realists typically mean by saying that truth is radically non-epistemic is that truth is not in any way constrained by our cognitive capacities. Note, however, that in itself this does not imply that there are any impediments to our coming to know each and every truth. Of course, realists typically have a bagful of reasons for believing that some truths are, or at least may be, beyond our ken. But these will not do to explain why (1) should be abandoned.

To see why not, consider that the reasons hinted at here are invariably of a methodological nature and are often lumped together under the heading of underdetermination of theory by the data. For instance, it is often said that for any theory that postulates unobservables there exist empirically equivalent rivals (roughly, rivals that share the same empirical consequences but tell incompatible stories about the unobservable part of reality), and that, unless they are refuted by the data, there is nothing in our methodology which could guarantee that we will be able to determine the truth values of such rivals.[23]

Next note that Fitch's Paradox can make do with an utterly minimal assumption about the logic governing the possibility operator (namely, that it comprises the rule $\Diamond\bot$). So, if we wish, we can assume that operator to indicate logical possibility in the broadest sense (as believed to be captured by the logic S5), and then we still have our paradox. Then (1) merely makes the claim that if a proposition is true, it is logically possible to know it.

Now, how could reasons relating to, for example, the problem of underdetermination explain that (1) is wrong, given that they seem perfectly compatible with it? In fact, these reasons are perfectly compatible even with the stronger (4). After all, that there are methodological reasons for believing that not all truths are epistemically accessible does not imply that there also are purely structural or logical reasons for believing that not all truths are so accessible: the methodological impossibility, or at least improbability, of coming to know all truths does not preclude that for each truth there is the logical possibility that it is known. In fact, it would be unsurprising if, at least prior to seeing Fitch's proof, a realist would subscribe to (1) or even to (4) – given a reading of the modal operator as indicating (broad) logical possibility, that is.

In short, Fitch's Paradox shows that already for purely logical reasons we cannot know all truths (unless we are willing to accept that all truths are known). And, contrary to what Hand and Kvanvig seem to suppose, the existence of such logical reasons is not explained by pointing to methodological constraints on inquiry. So the result of Section 4 should be welcomed by the realist camp, too.

ACKNOWLEDGEMENTS

I am greatly indebted to Michael Hand, Frank Hindriks, Bjørn Jespersen, Gert-Jan Lokhorst, Fred Muller, Johan van Benthem, Christopher von Bülow, and an anonymous referee for *Erkenntnis* for valuable comments on a previous version of this paper.

NOTES

1 The assumption underlying F is entirely uncontroversial. The one underlying D is not quite so uncontroversial (cf., e.g., Nozick 1981), but, as Williamson (2000, p. 283 ff) points out, it is not quite so essential to the paradox either: without it, (1) can still be shown to have a consequence that is hardly more palatable than (2).

2 Tennant (1990) and van Dalen (1994) contain careful introductions to the natural deduction system I will make use of at various points in this paper.

3 Cf. also Hand (2003, p. 221).

4 I take the criterion to apply, *mutatis mutandis*, to proposals to restrict the stronger principle (4) below, too. It in effect seems generalizable to other types of approach to Fitch's Paradox as well (and even to solutions to paradoxes *tout court*; see in this vein Haack (1978, p. 139)). A good example of such an approach that would satisfy the generalized criterion is Beall's (2000) proposal to abandon the classical logic underlying the paradox in favor of a paraconsistent one. The independent reason he provides is the Knower Paradox. Presuming there is no viable solution to this paradox (which is debatable, as he admits), it also explains *why* we should change to a paraconsistent logic: human knowledge in that case just happens to be inconsistent. For a similar proposal, see Wansing (2002), who is more specific than Beall about the type of paraconsistent logic that (in his view) is needed to deal with Fitch's Paradox, but who is less concerned with providing an independent motivation for his solution.

5 Literally, the passage reads: "The Restricted Thesis about knowledge [i.e., (3)] is substantive, informative and important. The objection that the restriction invoked is *ad hoc* is groundless". So, strictly, he does not say that the (putative) non-*ad hoc*ness is a consequence of the thesis' being substantive, etc. However, that this is what he means seems to be the only way to make sense of the succession of the cited sentences.

6 Or at any rate the theory seems to be as substantive and informative as a theory of truth can get; see Davidson (1996).

7 Rather some have complained that he has not said enough about the paradoxes; see, e.g., Field (1992).

8 Tennant (1997, p. 246 f, 258 f) seems to have a third reason for believing his solution is not *ad hoc*, namely, that unless we place similar restrictions as he proposes we place on (1) on what we can believe and, respectively, on what we can wonder about, the seemingly sensible claims that every truth is believable and that every truth can be wondered about come out paradoxical, too. This seems a non-starter, however, for the fact that in all these cases something has to be done to avoid paradox evidently does not mean that whatever is done to that effect in each case automatically qualifies as non-*ad hoc*. On the contrary, for instance the

"principle P' approach" mentioned in the text, although (of course) effective in all three cases, would still strike one as being utterly *ad hoc*.

9 Cf. Tennant (1997, p. 273 f). Tennant already earlier, in his own representation of Fitch's Paradox, implicitly assumes the necessitarian reading, for in that representation he makes use of the rule "$\varphi \backslash \Diamond K(\varphi)$" (p. 260). According to him, this rule "expresses the anti-realist principle that all truths are knowable" (*ibid.*). But given that it is a *rule*, and thus that the inference of $\Diamond K(\varphi)$ from φ (for any φ) is supposed to be universally valid (and not just when φ is a truth about the actual world), it really expresses the principle that, *necessarily*, all truths are knowable.

10 It is worth mentioning that the use of $\forall\varphi(\varphi \rightarrow \Diamond K(\varphi))$ in the necessitated subproof depending on $\neg\exists\varphi T(\varphi)$ requires (4) and would be inadmissible with only (1) as a premise; in such a subproof we may only assume, apart from the subproof's hypothesis, what holds necessarily according to the argument's premises.

11 The further problem, having an absurd conclusion, might seem to be even more serious than Fitch's Paradox, which has a conclusion that is hard to believe (as we said) but not (I take it) downright absurd. But this is not so. First observe that, given (4), we can strengthen the conclusion of Fitch's Paradox to "*Necessarily*, all truths are known", as follows: By Fitch's result, we have

$$\forall\varphi(\varphi \rightarrow \Diamond K(\varphi)) \vdash \forall\varphi(\varphi \rightarrow K(\varphi)),$$

or, by the deduction theorem,

$$\vdash \forall\varphi(\varphi \rightarrow \Diamond K(\varphi)) \rightarrow \forall\varphi(\varphi \rightarrow K(\varphi)).$$

So, by an application of the necessitation rule that is comprised by any normal modal logic,

$$\vdash \Box[\forall\varphi(\varphi \rightarrow \Diamond K(\varphi)) \rightarrow \forall\varphi(\varphi \rightarrow K(\varphi))].$$

From this and (4) it follows by the characteristic rule of K that $\Box\forall\varphi(\varphi \rightarrow K(\varphi))$. Now, given this result, and given that there are some truths about every possible world, it follows that there are known truths in all possible worlds, even in those devoid of knowers, a conclusion no less absurd than that, necessarily, some propositions are thought.

12 "Weaker" in the sense that whatever can be derived by means of *BD* can be derived by means of *RC* but not vice versa.

13 According to DeVidi and Kenyon (2003, p. 490), restricting the rules in the way proposed in the text is no option because "considerations of obviousness are of entirely the wrong sort for answering questions of logical possibility". The remark is puzzling, however, for it would seem that in determining what logical principles govern our reasoning about belief and knowledge, considerations of the sort of obviousness at issue are amongst the most relevant ones.

14 See, e.g., Moran (1997) and Adler (2002) on the specifically first-person and present-tense character of Moore's Paradox.

15 See, e.g., DeRose (1991, 2002), Brandom (1994), Alston (2000), Williamson (2000), Adler (2002), and Sundholm (in press). For some rare dissenting opinions see Bach and Harnish (1979) and Williams (2002).

16 See also Dummett (1981, p. 362): "judgment… is the interiorization of the external act of assertion"; and Sundholm (1999).

17 For those doubtful of the belief–assertion parallel, I should mention that, since the consequence just stated is the only consequence of the parallel that matters to the

solution to Fitch's Paradox to be given, it would not be necessarily troubling for us if certain, perhaps even quite substantial, disanalogies between the notions of belief and assertion must be admitted to exist.

18 In effect, Unger (1975, p. 258 f) already draws the parallel, albeit with "It's raining, but I don't know it is" (cf. also Moore 1962, p. 277; Williamson 2000, p. 253; Adler 2002, p. 194 f). What follows in the text is an obvious extension of that parallel to the type of proposition that is of more direct concern in the present context.

19 A third way of arguing for our conclusion would be by means of the inference rule "$B(\varphi, S, t) \backslash B(K(\varphi, S, t), S, t)$", which has been defended by some working in epistemic logic (see van der Hoek 1993, where what is here presented as a rule is presented as an axiom; see also Kraus and Lehmann 1988). Given that rule, the proof that $\varphi \wedge \neg K(\varphi)$ cannot be consistently believed (for any φ) is straightforward and left to the reader. As van der Hoek (*op. cit*., Section 6) notes, the rule is incompatible with the assumption of negative introspection, i.e., the assumption that $\neg K(\varphi) \rightarrow K\neg(K(\varphi))$ for all φ; it is again easy to show that the same holds for (9). But Meyer (2001, p. 188) is surely right that the assumption of negative introspection is at most tenable for artificial agents.

20 The solution still seems vulnerable to another objection against Tennant's solution raised in Hand and Kvanvig (1999). The objection is that if there can be unknowable truths (as is the case both on Tennant's and on my account), then why cannot *all* truths be unknowable? The objection would seem problematic for anyone led to antirealism by Dummettian concerns about manifestability of grasp of meaning; for such a person, the notion of an unknowable truth must be incoherent. However, not all antirealists are motivated by meaning-theoretic concerns. For instance, some seem to be driven by the idea that it is simply absurd to suppose the truth about the world might completely evade us (Peirce and Putnam may be cases in point). That is compatible with the claim that there are single truths which are unknowable. Of course it remains open to Hand and Kvanvig, or anyone else, to challenge whatever non-Dummettian motivations for denying realist truth these other philosophers may have.

21 I take the notion of warrant from Plantinga (1993). Note that warrant must be stronger than justification, given that the latter fails to turn true belief into knowledge in so-called Gettier cases (cf. Gettier 1963).

22 See, e.g., Hersh (1997) and Mancosu (2001) for a defense of the claim that explanatory mathematical proofs exist.

23 See on this argument, among others, van Fraassen (1980), Earman (1993), Douven and Horsten (1998), Douven (2000), and Devitt (2002).

REFERENCES

Adler, J.: 2002, *Belief's Own Ethics*, MIT Press, Cambridge, MA.

Alston, W.: 2000, *Illocutionary Acts and Sentence Meaning*, Cornell University Press, Ithaca.

Bach, K. and R. Harnish: 1979, *Linguistic Communication and Speech Acts*, MIT Press, Cambridge, MA.

Beall, J. C.: 2000, 'Fitch's Proof, Verificationism and the Knower Paradox', *Australasian Journal of Philosophy* **78**, 241–247.

Brandom, R.: 1994, *Making it Explicit*, Harvard University Press, Cambridge MA.
Davidson, D.: 1996, 'The Folly of Trying to Define Truth', *Journal of Philosophy* **93**, 263–278.
DeRose, K.: 1991 'Epistemic Possibilities', *Philosophical Review* **100**, 581–605.
DeRose, K.: 2002, 'Assertion, Knowledge, and Context', *Philosophical Review* **111**, 167–203.
DeVidi, D. and T. Kenyon: 2003, 'Analogues of Knowability', *Australasian Journal of Philosophy* **81**, 481–495.
Devitt, M.: 2002, 'Underdetermination and Realism', in E. Villanueva and E. Sosa (eds.), *Philosophical Issues*, Vol. 12, Blackwell, Oxford, pp. 26–50.
Douven, I. 2000 'The Antirealist Argument for Underdetermination', *Philosophical Quarterly* **50**, 371–375.
Douven, I. and L. Horsten: 1998, 'Earman on Underdetermination and Empirical Indistinguishability', *Erkenntnis* **49**, 303–320.
Dummett, M.: 1981, *Frege: Philosophy of Language,* (2nd edn) Duckworth, London.
Earman, J.: 1993, 'Underdetermination, Realism, and Reason'', in P. French, T. Uehling, Jr., and H. Wettstein (eds.), *Midwest Studies in Philosophy*, Vol. XVIII, University of Notre Dame Press, Notre Dame, pp. 19–38.
Field, H.: 1992 'Critical Notice: Paul Horwich's *Truth*', *Philosophy of Science* **59**, 321–330.
Fitch, F.: 1963, 'A Logical Analysis of Some Value Concepts', *Journal of Symbolic Logic* **28**, 135–142.
Fox, J.: 1989, 'What Were Tarski's Truth-Definitions For?', *History and Philosophy of Logic* **10**, 165–179.
Gettier, E.: 1963, 'Is Justified True Belief Knowledge?', *Analysis* **23**, 121–123.
Glanzberg, M.: 2004, 'A Contextual–Hierarchical Approach to Truth and the Liar Paradox', *Journal of Philosophical Logic* **33**, 27–88.
Haack, S.: 1978, *Philosophy of Logics*, Cambridge University Press, Cambridge.
Hand, M.: 2003, 'Knowability and Epistemic Truth', *Australasian Journal of Philosophy* **81**, 216–228.
Hand, M. and J. Kvanvig: 1999, 'Tennant on Knowability', *Australasian Journal of Philosophy* **77**, 422–428.
Hersh, R.: 1997, 'Prove – Once More and Again', *Philosophia Mathematica* **5**, 153–165.
Horwich, P.: 1990, *Truth*, Clarendon Press, Oxford.
Kirkham, R.: 1992, *Theories of Truth*, MIT Press, Cambridge, MA.
Kraus, S. and D. Lehmann: 1988, 'Knowledge, Belief and Time', *Theoretical Computer Science* **58**, 155–174.
Mancosu, P.: 2001, 'Mathematical Explanation: Problems and Prospects', *Topoi* **20**, 97–117.
Meyer, J.-J.: 2001, 'Epistemic Logic', in L. Goble (ed.), *The Blackwell Guide to Philosophical Logic*, Blackwell, Oxford, pp. 183–202.
Moore, G. E.: 1962, *Commonplace Book: 1919–1953*, Allen and Unwin, London.
Moran, R.: 1997, 'Self-Knowledge: Discovery, Resolution, and Undoing', *European Journal of Philosophy* **5**, 141–161.
Nozick, R.: 1981, *Philosophical Explanations*, Harvard University Press, Cambridge, MA.
Plantinga, A.: 1993, *Warrant: The Current Debate*, Oxford University Press, Oxford.
Putnam, H.: 1990, 'Realism with a Human Face', in *Realism with a Human Face*, Harvard University Press, Cambridge, MA, pp. 3–29.

Putnam, H.: 2000, 'Paradox Revisited', in G. Sher and R. Tieszen (eds.), *Between Logic and Intuition*, Cambridge University Press, Cambridge, pp. 3–26.

Quine, W. V.: 1961, 'Paradox', *Scientific American* **206**, 84–96. (Reprinted in his *The Ways of Paradox*, Harvard University Press, Cambridge, MA, 1976, pp. 1–18; the page reference is to the reprint.)

Sundholm, G.: 1999, 'MacColl on Judgement and Inference', *Nordic Journal of Philosophical Logic* **3**, 119–132.

Sundholm, G.: in press, 'Antirealism and the Roles of Truth', in I. Niiniluoto, M. Sintonen, and J. Woleński (eds.), *Handbook of Epistemology*, Kluwer, Dordrecht.

Tennant, N.: 1990, *Natural Logic*, Edinburgh University Press, Edinburgh.

Tennant, N.: 1997, *The Taming of the True*, Oxford University Press, Oxford.

Tennant, N.: 2001, 'Is Every Truth Knowable? Reply to Hand and Kvanvig', *Australasian Journal of Philosophy* **79**, 107–113.

Unger, P.: 1975, *Ignorance: A Case for Scepticism*, Clarendon Press, Oxford.

van Dalen, D.: 1994, *Logic and Structure* (3rd edn), Springer, New York.

van der Hoek, W.: 1993, 'Systems for Knowledge and Belief', *Journal of Logic and Computation* **3**, 173–195.

van Fraassen, B. C. 1980, *The Scientific Image*, Clarendon Press, Oxford.

Wansing, H.: 2002, 'Diamonds Are a Philosopher's Best Friends', *Journal of Philosophical Logic* **31**, 591–612.

Williams, B.: 2002, *Truth and Truthfulness: An Essay in Genealogy*, Princeton University Press, Princeton.

Williamson, T.: 2000, *Knowledge and Its Limits*, Oxford University Press, Oxford.

Department of Philosophy
Erasmus University Rotterdam
Postbus 1738
Rotterdam 3000 DR
The Netherlands
E-mail: douven@fwb.eur.nl

Manuscript submitted 7 December 2003
Final version received 27 April 2004