CUTTING THE GÖDELIAN KNOT

Frederick W. KROON

1. Introduction

Logicism has often been declared dead and ready for burial. In a recent paper ([6]), Geoffrey Hellman updates the charge that Gödel killed logicism by arguing that logicists better not replace the doctrine that mathematical truth coincides with provability from logicist logic (a doctrine falsified by Gödel's first incompleteness theorem if the theorems of such a logic can be effectively generated) with the epistemological doctrine that at least all *knowable* mathematical truths coincide with theorems of logicist logic. The latter doctrine too, claims Hellman, stands refuted by Gödel: this time by his second incompleteness theorem. The present paper is an attempt to rescue epistemological logicism from this new attack.

Let me begin my account of this attack with a more precise formulation of epistemological logicism. It is a doctrine that is taken to imply, at the very least, that:

EL: There is a formal system K such that

(1) for any knowable mathematical claim, P, there is a sentence S in the language of K such that S represents P and S is a theorem of K; and

(2) any theorem S of K represents some knowable mathematical claim.

The versions of epistemological logicism that Hellman thinks he is able to rule out most directly are the versions that imply:

EL_1: There is a *finitely axiomatizable* formal system K meeting conditions (1) and (2) of EL.

But Hellman is also able to argue against certain versions of EL that do not impose the requirement of finite axiomatizability, specifically ones that imply (pp. 457-495):

EL_2: There is a recursively axiomatizable formal system K meeting conditions (1) and (2) of EL and satisfying the condition:

there is an enumeration of the axioms of K such that, if B_i is the conjunction of the first i members of K in this enumeration, then

(i) ⌜B_i is true⌝ is knowable for all i

materially implies

(ii) ⌜$\forall i B_i$ is true⌝ is knowable.

Hellman argues in addition that even if the logicist advocates a version of EL that does not entail EL_2 (and hence does not entail EL_1) it cannot be known of the logicist system K he advocates that it satisfies the conditions of EL (p. 459). Since the logicist presumably thinks that he will know that his ideal logicist system is a logicist system once it is finally constructed (call this epistemic version of EL "EL_3"), this result also seems to conflict sharply with the hopes that logicists have for their program.

Note, by the way, that Hellman's argument doesn't even exploit features of logicism that are peculiar to logicism. His argument, if valid, is a quite general argument against certain forms of the thesis that mathematics constitutes a unified science, whether or not that science is taken to be logic, set theory, category theory, or some science of structure even more general than category theory. Of course if the unified science in question is to be *logic* then there might well be good reasons for supposing that the system better be finitely axiomatizable (for how could a non-finitely axiomatizable science contain the principles that the mind relies on in thinking logically? How could such a science be encoded in a finite brain?), and in that sense, perhaps, Hellman's argument against EL_1 might well count as a disproof of logicism in particular.

I shall argue in this paper, that Hellman's attacks on EL_1, EL_2 and EL_3 fail. In the next section, I state Hellman's argument and assumptions, and then discuss what I take to be the erroneous move in Section 3. I then argue (in Section 4) that logicists have always had reason to reject this move.

2. *The argument against epistemological logicism*

Hellman's attacks on EL_1, EL_2 and EL_3 all have the same form. Each argument proceeds by using certain lemmas connecting knowability, truth and consistency to establish that any logicist system of the specified kind will be able to prove its own consistency, *contra* Gödel's second incompleteness theorem. I shall state only the first proof in detail.

Metatheorem I There is no formal system that is finitely axiomatizable and that meets the conditions of EL. (That is, EL_1 is false).

Hellman begins by briefly defending the following lemmas ([1]):

L_1: If P is a mathematical statement that is knowable (or: "$\Diamond Kn(P)$"), then P expresses a mathematical truth.

L_2: For any system K meeting EL and any sentence X in the language of K, L(K): $\Diamond Kn(X) \rightarrow \Diamond Kn(True_{L(K)}(X))$

L_3: For any system K meeting EL and any sentence X in the language of K, L(K): $\Diamond Kn(True_{L(K)}(X)) \rightarrow \Diamond Kn(X \nvdash_{UL(K)} A \& \sim A)$,

The argument for Metatheorem I now proceeds as follows:
Suppose that there is such a formal system, K*. Let C be the conjunction of its axioms (exclusive of any axioms in UL(K*), K*'s underlying logic; we have no right to suppose that UL(K*) has only finitely many axioms). Then,

(1)	$\vdash_{K^*} C$	
(2)	$\Diamond Kn(C)$	(EL hypothesis.)
(3)	$\Diamond Kn(True_{L(K^*)}(C))$	(By (2) and L_2)
(4)	$\Diamond Kn(C \nvdash_{UL(K^*)} A \& \sim A)$	(By (3) and L_3)
(5)	$\Diamond Kn(\nvdash_{K^*} A \& \sim A)$	(Since C is conjunction of axioms of K* - UL(K*))

([1]) "$\rightarrow$" is here the symbol for logical implication. We can assume that an expression like $True_{L(K)}(X)$ is short for $True_\alpha(\beta)$ with α some unspecified expression denoting L(K) and β a (standard) name for X.

(6) "$\nvdash_{K^*}$A & ~ A"
is a mathematical truth (From (5), L_1, and the fact that this statement is assertable within arithmetic.)

(7) "$\nvdash_{K^*}$A & ~ A" asserts the consistency of K* and, as a statement of number theory, has a translate in L(K*), abbreviated "Con(K*)".

Hence (8) $\vdash_{K^*}$Con(K*) (By (5), (6), (7) and the EL hypothesis.)

But (8) contradicts Gödel's second incompleteness theorem. Hence there cannot be a formal system K* meeting the conditions of EL.

The arguments against EL_2 and EL_3 are similar. Each argument makes use of a principle much like L_3 to establish that the proposition ◇Kn (all axioms of K are true) entails that ◇Kn (K is consistent). Since the first claim can be verified in the case of systems of type EL_2 or EL_3 the supposition that there are such systems again contradicts Gödel's second incompleteness theorem.

3. *Challenging the argument*

First a preliminary worry about Hellman's argument. Hellman's presentation leaves it unclear whether he thinks that the operator *knowable* is a propositional operator, or an operator on sentences, or perhaps both (compare, for example, steps 3 and 4 of Hellman's argument). He most often writes as if it is an operator on sentences and statements (which presumably are also to be thought of a linguistic items), and so we should perhaps regard this as his considered view. But the position is a troublesome one. If Hellman does accept such a quotational account of knowability, and also accepts a number of minimal conditions on this operator, specifically the ones set out in Rich Thomason's [12], his position can surely be shown to be incoherent via the sort of argument Thomason develops. (Thomason's argument against such quotational accounts is an adaptation of Montague's well-known argument against syntactic treatments of modality: an argument, ironically, which is based on the possibility of fixed-point results of the kind Gödel himself established on the basis of the technique of Gödel-numbering.) Since Hellman accepts these minimal conditions of knowability (conditions such as (i) the truths of first-order logic are knowable, (ii) if S is knowable and $(S \supset T)$ is knowable, then T is knowable, at least for sentences S, T of elementary number theory), it is perhaps tempting to hypothesize that Hellman's argument against EL illustrates the incoherence of certain of Hellman's own presuppositions concerning knowledge rather than the incoherence of epistemological logicism itself, which is not wedded to a quotational account of knowledge of knowability ([2]) and whose principals like Frege indeed held strongly non-quotational views of epistemic operators.

The question is: even if Hellman accepts a quotational account of knowledge, does his argument actually rest on such an account? At times, the answer seems to be "yes". L_2, in particular, seems hard to motivate independently of a quotational account: construed as saying that the knowability of the proposition expressed by sentence X entails the knowability of the proposition expressed by $True_{L(K)}X$, L_2 no longer looks so obvious, for unlike knowledge of the proposition expressed by X knowledge of the proposition expressed by $True_{L(K)}(X)$ is knowledge about sentence X itself – quite a different matter. Perhaps L_2 can survive this criticism. It is, after all, a thesis about what it is possible to know, and not one about knowledge. But even if L_2 is to be rejected, Hellman's argument is not beyong rescuing. Suppose we characterize EL slightly differently, but I think no more contentiously, as the following claim: there is a formal system K such that for sentences S of L(K) it is knowable that S is a truth of L(K) iff S is a theorem of K. The following proof of metatheorem I bypasses L_2:

(1)	$\vdash_K C$	
(2)	$\Diamond Kn(True_{L(K^*)}(C))$	(new EL hypothesis)
(3)	$\Diamond Kn(\nvdash_{K^*} A \& \sim A)$	(by (2), L_3, and since C is a conjunction of (proper) axioms of K*)
(4)	$\Diamond Kn(True_{L(K^*)}(Con(K^*)))$	(by (3))
(5)	$\vdash_{K^*} Con(K^*)$	(new EL hypothesis).

This is essentially Hellman's proof, and hence it seems safe to assert that his complaint against epistemological logicism does not rest on a quotational view of knowability.

([2]) See also [1].

There should be doubts, nonetheless, even about the present proof. Consider the move from (2) to (3), which rests on L_3. L_3, I suggest, should be viewed as contentious, but for reasons that have to do with the nature of knowing rather than the nature of objects of knowledge. Let me explain why.

What warrants acceptance of L_3? Presumably the obvious fact that:

(a) For any system meeting EL and any sentence X in L(K),

$$\Diamond Kn(True_{L(K)}(X)) \rightarrow \Diamond Kn(True_{L(K)}(X) .\&.$$
$$(True_{L(K)}(X) \rightarrow X \nvdash_{UL(K)} A \& \sim A))$$

That is, if it is knowable that X is true in L(K), then so is the conjunction of this claim about X and the fairly trivial a priori claim that the truth of X implies the consistency of X: one's knowledge of the former claim does not, in this kind of case, upset the possibility of also having knowledge of the implication.

Does (a), by itself, suffice for L_3? Clearly not. To derive L_3 we need a further condition:

(b) $\Diamond Kn(True_{L(K)}(X) .\&. (True_{L(K)}(X) \rightarrow X \nvdash_{UL(K)} A \& \sim A)) \rightarrow \Diamond Kn(X \nvdash_{UL(K)} A \& \sim A)$

L_3 now follows from (a) and (b) by the transitivity of $\rightarrow$, and hence we have a proof of L_3 from apparently uncontroversial premises.

(a), indeed, appears quite unexceptionable. Is (b) similarly well-grounded? It certainly appears so. What licenses (b) seems to be the following schema:

(c) $\Diamond Kn(S \& (S \rightarrow T)) \rightarrow \Diamond Kn\ T$

and this schema is far weaker than most theses in epistemic logic. (c), in fact, is a very guarded version of the following schema:

(c_1) $Ka(S .\&. (S \rightarrow T)) \rightarrow Ka\ T$,

whose instances are all theorems of standard epistemic logics ("Ka" means "a knows......"). Now it is arguable that (c_1) should not, in fact, be tolerated in epistemic logic: drawing inferences takes time (as well as a modicum of rationality), and the ability to draw inferences instantaneously is not part of our ordinary concept of a knower (or even our concept of a *rational* knower). But this objection is an objection to (c_1) only, and not to the more guarded claim (c) which makes room for the fact that drawing a particular inference requires resources whose availability is not logically guaranteed.

(c_1) suffers from another defect, however, one which is not rectified by means of modalizing (c_1) to yield (c). The defect was first noted by Fred Dretske in "Epistemic Operators" ([3]). Consider the following examples. On the basis of ordinary perceptual evidence – noticing that it was a Parker pen I picked up from my pen case, and not, e.g., a ball-point pen – I know that I am now writing with my Parker pen. I also know that if this is the case then I am not writing with Jones' identical Parker pen, which he switched for mine when I was not looking. But this is not to say that I know that I am not writing with Jones' identical Parker pen, switched for mine when I was not looking. How could I be said to know this, not having tried and surely not being required, to rule out this (unlikely) possibility when I first acquired the belief that it was my Parker pen I was writing with? Here, then, we have a case of knowing some proposition S (I am writing with my Parker pen), knowing at the same time an entailment $S \rightarrow T$ (if I am writing with my Parker pen, then I am not writing with Jones' identical Parker pen, switched for mine when I was not looking), and not knowing, for lack of relevant evidence, proposition T. This kind of counterexample to (c_1) is also a counter-example to (c), note. It need not be true that T is even knowable in this case, for Jones might have God-like powers that prevent people from ever being able to detect the switch.

Robert Nozick provides us with another kind of example in *Philosophical Explanations* ([11]). I know that I am currently writing with my Parker pen, perhaps even that I am not writing with Jones' identical Parker pen. I also know that if I am really doing this then I can't be a brain in a vat on Alpha Centauri, being stimulated to have the beliefs and experienes I in fact do have. But I do not know, nor could I ever know, that I am not such a brain in a vat on Alpha Centauri. Nothing could ever be a relevant piece of evidence for this latter claim, since even if I were a brain in a vat being stimulated to have the experiences I in fact do have the world would appear no different from the way it actually appears to me.

Such arguments seem to cast doubt on both (c_1) and the weaker (c)(3) but particular instances of (c) and (c_1) might nonetheless still be logically true. Perhaps, for example, there are propositions T that any knower must know in order to satisfy some minimal rationality requirement, thus trivially validating instances of (c) and (c_1) whose consequents involve such propositions T. This sort of approach is useless, however, as far as justifying L_3 is concerned. Consistency claims simply do not have this kind of status. To justify L_3 for all, or almost all, sentences X, we must somehow give a general argument that shows that if the conjunction of the premises in an implicational inference are known, or are knowable, and if they have the form given by the antecedent of L_3, then the conclusion is known (knowable) as well.

I doubt that there can be such a general argument. Nozick, as it happens, must think otherwise. In Ch. 3 of *Philosophical Explanations*, he tries to explain the failure of (c) in terms of the thesis that knowledge is belief which tracks the truth. Specifically (4), A knows via method M that p iff (i) p is true, (ii) A believes that p, (iii) if p were not true, and a were to use M to decide whether or not p, then a would not believe, via M, that p, and (iv) if p were true and A were to use M to decide whether or not p, then a would believe, via M, that p. A knows that p, further-

(3) There are other ways of coping with the intuitions on which Dretske and Nozick rest their case – fortunately enough, given the many problems facing Nozick's account (see, e.g., [9]). Some would prefer to accept a view according to which knowledge is knowledge relative to a context, where a context may change if what is at issue changes. Thus if the possibility of my being a brain on Alpha Centauri is not at issue, then I may know, relative to such a context, that I am writing with a pen. But I will then also know, relative to that context, that I am not a brain on Alpha Centauri, i.e., the context-relative version of (c) holds. If, on the other hand, the context changes in a way that makes the possibility of my being a brain on Alpha Centauri an issue, the knowledge-claim need no longer hold. (I am indebted to M.J. Cresswell for originally drawing my attention to this alternative view. I develop the view further, using examples somewhat similar to the ones in the present paper, in my [8].)

This contextualist view is in many ways very attractive. I shall not here assess its merits, however (again, see [8]). The only point that needs to be made is that such an alternative does not save Hellman's argument against epistemological logicism, but in effect requires a radical reformulation of that doctrine. Whether, and how, Hellman's argument can then be reformulated to apply to this new version of epistemological logicism are questions I shall not consider here.

(4) [11], pp. 197. Nozick introduces a further condition on p. 182f, about the relative importance of various belief-acquiring methods when belief is overdetermined.

more, iff for some method M A knows via M that p. (iii) and (iv) are the tracking conditions that specify how A *would* believe if the proposition p *were* true/false. In the "brain in a vat" example (as in the example of Jones' identical Parker pen), condition (iii) fails, since even if I were a brain in a vat, using whatever method was available to determine whether or not this was really the case, I would still believe that I was not, since my set of beliefs and experiences would be no different from my actual set of beliefs and experiences. Clearly, however, such a use of tracking condition (iii) becomes impossible once the proposition p in question is necessarily true. In no counterfactual situation is a necessarily true proposition false. Hence (i)-(iv) above reduce to (i), (ii) and (iv) in the case of a necessarily true proposition p. Nozick deliberately does not advocate an alternative non-tracking account of our knowledge of propositions of this sort (5), and appears to accept the following principle:

(d) [T is necessarily true &

$$\Diamond \mathrm{Kn}(S \;.\&.\; (S \rightarrow T))] \rightarrow \Diamond \mathrm{Kn}\, T$$

Given that a consistency claim is non-contingent, schema (b) immediately follows.

I reject this reasoning. What seems to me wrong with it is precisely that the "tracking" account of knowledge is wrong for necessary propositions, although perhaps right for contingent propositions. From this point of view, an argument for (d) and (b) based on the "tracking" conception of knowledge must appear specious. A stronger claim can be made, in fact: some of the same considerations that show a schema like (c) to be incorrect when contingently true propositions T are in question also show it to be incorrect when necessarily true propositions are in question (schema (d)). The first piece of evidence for this is simply that the cognitive content of the proposition expressed by a sentence does not seem to be altered by prefixing to the latter the "actuality" operator which converts sentences expressing true propositions to ones expressing necessary truths (6): just as the contingent truth expressed by "I am not a brain in a vat on Alpha Centauri" is unknowable (by me), so is the necessary truth expressed by "I am not actually a brain in a vat on Alpha Centauri"

(5) [11], pp. 186-187.
(6) See, e.g., [2].

unknowable (by me). The very same reasons that make the one unknowable also make the other unknowable. But (d) clearly must deny this, and hence (d) is inadequate.

I shall not investigate possible ways of dealing with the criticism. Instead I shall outline a more telling criticism against (d), more telling because it does not depend on the peculiarities of the "actuality" operator and because it can be formulated as a direct criticism of (b). Suppose that I know a certain mathematical statement X to be true. Suppose that I also realize that if X is true then one can't derive an inconsistency from X using the resources of classical logic. I may put this knowledge together, and yet still not know that from X no inconsistency is derivable. I am, of course, committed to believing that no inconsistency is derivable from X, but that alone is scarcely sufficient for knowledge. (I am similarly committed to believing, on the basis of other things I know and believe, that I am not a brain in a vat in Alpha Centauri, but I do not thereby *know* that I am not a brain in a vat on Alpha Centauri.) The reasons for denying that I must know that X is consistent are similar to the reasons we gave earlier for denying that I must know that I am not a brain in a vat, or that it is not Jones's Parker pen I am writing with. In the first place, the evidence which justifies my acceptance of X and leads to my knowledge of X, will generally be evidence of a positive kind: evidence that there is a proof of X from other statements I know, or (in the case of fundamental axioms) evidence of the usefulness of incorporating X into our overall scientific theory of the world (cf. Quine), insight based on Gödelian intuition (7), or even (for some logistics) a special logical insight. The evidence for X may (for the Quinean position, say) sometimes include an argument to the effect that adding X to our theory of the world T yields a theory which is consistent relative to T (this doesn't yet show that X is consistent, of course, only that $T \cup \{X\}$, and hence X, is "at least as consistent as" T), but giving such an argument will seem necessary only when there is more than the customary degree of suspicion that adding X will do no harm. That we must have a consistency proof for X (let alone an absolute consistency proof for the ensuing new theory of the world $T \cup \{X\}$) before we can be said to know that X is true is not

(7) See [5], especially pp. 271-272. A recent defence of Gödel's epistemological platonism is found in [10]. In the light of Hellman's appeal to Gödel, Hellman more than others, might well have seen that (b) and L_3 are not self-evidently true.

likely to be a general evidential requirement, even for these positions ([8]), while to positions that favour the possibility of something like a Gödelian intuitive faculty for logico-mathematical knowledge, consistency proofs for mathematical statements are probably going to appear evidentially irrelevant as far as our knowledge of these statements is concerned.

We are now in a position to make the following point against (b) and hence (d). Suppose that I know that X is true. My evidence for X is also evidence for the claim that X is true: the evidential route to the former leads directly to the latter via a Tarski biconditional. But even though I may also realize that the consistency of X follows from the truth of X, I do not thereby know that X is consistent. I may have no definitive argument establishing the consistency of X, not even a moderately good probabilistic argument, and hence I can scarcely be said to *know* that X is consistent. This is not really surprising. After all, a consistency claim is an infinitistic claim, a claim to the effect that there is no proof from X to a statement of the form A& ~ A. Whatever evidence I have for statement X itself, that evidence is unlikely to yield an argument for an infinitistic claim about all possible proofs from X in some given logic. Furthermore, no such argument may be possible in the case of sufficiently powerful statements X: the consistency of X may be unknowable, that is, even when the truth of X is knowable (which is not to say that we have some reason to doubt that X is consistent, some reason for supposing that the question is an open one). In so far as this possibility is denied by (b) and (d), they are prima facie unacceptable.

In the case of the Parker pen and brain-in-a-vat examples there was another argument we were able to use: there we could say that even if I were a brain in a vat, or writing with Jones' identical Parker pen, I would still believe that I wasn't (that is, tracking condition (iii) fails). This way of putting the point about the lack of relevant evidence is not available to us in the present case, since there is no counterfactual possibility that a particular statement is inconsistent if it is, in fact, consistent. But we can perhaps make an analogous epistemic point. Suppose that I know X to be true. X is therefore consistent. But if X is sufficiently complex one can certainly imagine that X might one day be classified as inconsistent. Were X a statement that would, perhaps "at the end of enquiry",

([8]) Certainly one cannot imagine Quine, with his high disdain for the absolutely evident, the absolutely certain, insisting on such an evidential requirement.

be classified as inconsistent (perhaps because of the possibility of an evidentially powerful but nonetheless subtly fallacious proof of the inconsistency of X), this is not likely to affect my present belief that X is consistent, however. It is more likely than not that a "demonstration" of inconsistency of this kind would not emerge for a long time (after all, everything so far points to the consistency of X), and hence more likely than not that even if X were "inconsistent" in this epistemic sense I would still believe in the consistency of X (on the basis of my argument for the truth of X). So the epistemic analogue of tracking condition (iii) also fails in the present case.

4. A logicist defence of the challenge

There are at least two kinds of rejoinder to this attempt at establishing the possible unknowability of the consistency of knowable statements. The first and more general kind aims to defend (c), and proceeds by way of an argument for the claim that knowledge in the strict sense (strict knowability) must, after all, be closed under known (knowable) logical implication. The second kind aims to defend only some more limited closure principle, proceeding by way of an argument for some claim to the effect that knowledge (knowability) of the sort of statements involved in L_3 must, after all, be closed under known (knowable) logical implication, even if this doesn't hold – *pace* (c) – for all statements.

I shall not consider this second kind of rejoinder further. It seems to me evident that if the intuitions underlying the Dretske-Nozick examples do indeed show (c) to be suspect, then intuitions of a similar kind show (d) and (b) to be suspect as well.

What about the first kind of rejoinder, that the truth of (c) is somehow constitutive of knowledge proper, knowledge in the strict sense? On such an accoint, if I really cannot know that I an not a brain in a vat on Alpha Centauri, then I also cannot know propositions that demonstrably entail that I am not, e.g. that I am writing with my Parker pen. Now it is reasonable to argue that the underlying conception of knowledge sets unrealistically, if not impossibly, high standards for knowledge, but let us leave that criticism aside. A more telling point can be made by way of the observation that logicists themselves, unlike some other proponents of mathematics-as-a-unified-science, have all along had a particular interest in rejecting this high conception of knowledge, whether or not they have acknowledged this interest. It is a conception of knowledge that conflicts with a view about the nature of properly constituted mathematical knowledge as a priori or experience-independent ([9]). Surprisingly enough, given one's intuitive impression that logicists of all people should embrace "high" conceptions of knowledge, and that a priorism is a "high" conception of knowledge, the epistemological a priority of logicist truths stand threatened by (c). Here is why. Let "X" range over logicist theorems, and let "E" range over formulae of the form:

> "In reasoning to conclusion X, a was not affected by cognitive malfunctions (e.g., failure of memory, double vision when looking at certain symbols, etc.) at crucial stages of his argument."

Let "Ka" stand for "a knows ... a priori", and "◇Ka" for it is possible for a to know ... a priori ([10]). Consider the argument:

(i) $\Diamond Ka(KaX \mathbin{.\&.} (KaX \rightarrow E))$	(assumption)
(ii) $\Diamond Ka(KaX \mathbin{.\&.} (KaX \rightarrow E)) \rightarrow \Diamond KaE$	(an instance of the appropriate version of (c))
(iii) $\Diamond KaE$	(by MP from (i) and (ii)

The logicist, however, will surely reject (iii) if the proof of X is sufficiently complex. For such X, perhaps for all X, and with "a" replaced by a

([9]) For a recent characterization of a priority as an epistemic qualifier, that is, an attribute of knowledge, see [7]. Paul Benacerraf has suggested to me that Frege at least had a non-epistemic concept of a priori. In [4], sect. 4, Frege wrote:

> ... if... the proof of a proposition can be derived entirely from general laws, which themselves neither need nor admit of proof, then the truth is a priori.

My own view is that Frege held the Kantian view of a priority as an epistemic qualifier, and then simply supposed that a proposition characterized as in *Grundlagen* 4 was a priori knowable in this Kantian sense. Frege was too much of a Kantian to have introduced a new non-Kantian concept of the a priori, especially to have introduced it without proper acknowledgement or justification. I do not agree, in particular, that the change to a new non-epistemic concept of a priori was thought by Frege to be required for an appropriately de-psychologized approach to mathematical foundations.

([10]) We could presumably use these locutions to define "is a priori knowable" and "is a priori known", e.g. $\Diamond KnS$ could be seen as equivalent to $\Diamond(\exists a)KaS$.

term denoting a person, E is an *empirical* statement (you and I don't have an a priori guarantee that our mental functions are never impaired when trying to prove theorems!) Hence the logicist has his own reasons for rejecting (c) if he also believes himself to have strong reasons for accepting (i) in the case of some (sufficiently complex) X. But he seems certain to believe that he has such reasons. Take as value of "a" the logicist himself, and assume that he knows (the proposition expressed by) X through having constructed and/or surveyed a complex proof of X. He will then believe himself to know X a priori, and, in addition, he is bound to believe that his having constructed such a proof demonstrates that he knows X a priori: his reflexive recognition that his proof is a proof in a system meeting (EL) provides him with an a priori warrant for his claim that he knows X a priori ([11]). Furthermore, he will be able to know concurrently that his knowledge of X implies the truth of E: the falsity of E would mean the defeasibility of his justification for believing X, and hence the absence of knowledge of X. His knowledge of this implication he will also regard as a priori, being knowledge of a conceptual kind. Hence he is bound to accept that he has a priori knowledge of conjunction KaX .&. (KaX → E)), i.e. he is bound to accept an instance of (i). To resist the inference to the clearly false conclusion (iii), he must now give up (ii), and hence must give up (c).

Thus we have provided the logicist with internal reasons for acknowledging (c) as wrong if a priori knowability is the species of knowability in question. And since ◇Kn must be taken to be the a priori knowability operator in L_1 - L_3 if we are to capture such a logicist's intentions, the logicist must acknowledge that there is now a problem about the justification of (b) and hence of L_3. What, after all, if not the a priori validity

([11]) What if the logicist denies.

(*) for all logicist propositions S, KaS → ◇Ka(KaS)?

He will then need to decide what to give up in his understanding of a priori knowledge: (*) or the thesis ∀S, T ◇Ka(S & (S → T)) → ◇KaT (a version of (c)). Because thesis (c) already looks suspect for other reasons, the sensible plan, I think, is to reject the version of (c) cited, whatever else is done. And if the logicist accepts EL_3 in addition to EL_1, it is very clear that this is what he should do. In accepting EL_3, he accepts that he knows that his logicist system is a logicist system, knowlegde which he is bound to regard – how else? – as a priori. Such a logicist will certainly assent to (*).

of the argument form: S, S $\rightarrow$ T $\therefore$ T, gives (b) and hence L_3 its strong air of plausibility? The logicist thus seems to lack a clear reason for accepting L_3, and hence cannot readily be accused of internal inconsistency should he venture the denial of L_3.

But why should the logicist who rejects (ii), and thereby the appropriate version of (c), not simply be charged with *ad hoc* manoeuvring in the face of yet another problem facing familiar forms of epistemological logicism, the problem exposed in argument (i)-(iii) above? And what can he now possibly *mean* by knowledge in the light of such manoeuvring? What, in detail, might a reasonable theory of knowledge along these lines look like? First questions first. The logicist shouldn't be charged with *ad hoc-ery* because his rejection of (ii) receives independent confirmation in the considerations of evidential relevance that also motivated our earlier argument against (c) and subsequently (b) and L_3. We stressed, in that argument, that whatever evidence the logicist or anyone else has for accepting the statement $\text{True}_{L(K)}(X)$ this evidence is not likely to bear significantly on the question of the consistency of X. In the case of (ii), we can similarly point out that whatever a priori evidence the logicist or anyone else has for the proposition that he knows X and that his knowing X logically implies E, this evidence is not likely to bear significantly on the question of the truth of E itself. E focuses on the possible existence of cognitive malfunctions, after all, while the former focuses on logico-mathematical proofs of a certain kind. For the logicist, in fact, the gap between the two propositions as far as relevant evidence is concerned must appear extreme, since he regards one proposition as a priori and the other, E, as a posteriori.

The second question is much more difficult, and I shall not try to develop an answer here. It is possible that at the end of the day the logicist will be left with no leg to stand on, since there might not be any reasonable account of knowledge embracing the suggestion above. But whether that is so or not is surely a further question, one that Hellman and other detractors have yet to address.

University of Auckland
The Australian National University

Frederick W. KROON

BIBLIOGRAPHY

[1] M.J. Cresswell, "Quotational Theories of Propositional Attitudes", *Journal of Philosophical Logic* 9 (1980): 17-40.

[2] M. Davies and L. Humberstone, "Two Notions of Necessity", *Philosophical Studies* 31 (1980).

[3] F. Dretske, "Epistemic Operators", *Journal of Philosophy* 67 (1970): 1007-1023.

[4] G. Frege, *Die Grundlagen der Arithmetik* (Breslau, 1884).

[5] K. Gödel, "What is Cantor's Continuum Problem?" in *Philosophy of Mathematics: Selected Readings*, ed. P. Benacerraf and H. Putnam (Prentice-Hall, 1964).

[6] G. Hellman, "How to Gödel a Frege-Russell: Gödel's Incompleteness Theorems and Logicism", *Nous* 15 (1981): 451-468.

[7] P. Kitcher, "A Priori Knowledge", *Philosophical Review* 89 (1980): 3-323.

[8] F. Kroon, "Sceptical Intuitions and Philosophical Explanations", *The Philosophical Quarterly* (1986).

[9] S. Luper Foy, ed., *The Possibility of Knowledge* (Rowman & Littlefield, 1987).

[10] P. Maddie, "Perception and Mathematical Intuition", *Philosophical Review* 89 (1980).

[11] R. Nozick, *Philosophical Explanations* (Harvard University Press, 1981).

[12] R. Thomason, "Indirect Discourse is not Quotational", *Monist* (1977): 340-354.