Synthese (2010) 176:177–225
DOI 10.1007/s11229-009-9490-3

From the Knowability Paradox to the existence of proofs

W. Dean · H. Kurokawa

Received: 1 July 2008 / Accepted: 27 January 2009 / Published online: 19 May 2009

Abstract The Knowability Paradox purports to show that the controversial but not patently absurd hypothesis that all truths are knowable entails the implausible conclusion that all truths are known. The notoriety of this argument owes to the negative light it appears to cast on the view that there can be no verification-transcendent truths. We argue that it is overly simplistic to formalize the views of contemporary verificationists like Dummett, Prawitz or Martin-Löf using the sort of propositional modal operators which are employed in the original derivation of the Paradox. Instead we propose that the central tenet of verificationism is most accurately formulated as follows: *if* φ *is true, then there exists a proof of* φ. Building on the work of Artemov (Bull Symb Log 7(1): 1–36, 2001), a system of explicit modal logic with proof quantifiers is introduced to reason about such statements. When the original reasoning of the Paradox is developed in this setting, we reach not a contradiction, but rather the conclusion that there must exist non-constructed proofs. This outcome is evaluated relative to the controversy between Dummett and Prawitz about proof existence and bivalence.

Keywords Knowability Paradox · Fitch · Verificationism · Intuitionistic logic · BHK interpretation · Existence predicate · Logic of proofs · Potential proof · Bivalence

W. Dean · H. Kurokawa
Graduate Center, City University of New York, New York, USA

W. Dean
e-mail: wdean@gc.cuny.edu

H. Kurokawa
e-mail: hkurokawa@gc.cuny.edu

W. Dean (✉)
Department of Philosophy, University of Warwick, Coventry, UK
e-mail: w.h.dean@warwick.ac.uk

1 Introduction

1.1 The Knowability Paradox

The Knowability Paradox purports to show that from the controversial but not patently absurd hypothesis that all truths are knowable, it is possible to derive the implausible conclusion that all truths are known. These statements can be expressed schematically as follows:

(KP) For all φ, if φ is true, then it is possible to know φ.
(Om) For all φ, if φ is true, then φ is known.

(KP) is traditionally referred to as the *Knowability Principle*. It is conventionally formalized in a propositional modal language $\mathcal{L}_{LK}$ containing operators $\Diamond$ (*it is possible that*) and K (*it is known that*). In this language, (KP) takes the form

(KP1) For all φ, $\varphi \rightarrow \Diamond K\varphi$.

(Om) is accordingly formalized in $\mathcal{L}_{LK}$ as

(Om1) For all φ, $\varphi \rightarrow K\varphi$.

On the interpretation of the modal operators just offered, (Om1) is presumably false. For surely as a matter of empirical fact, we are non-omniscient in the sense that

(NonOm) For some ψ, ψ is true but not known.

This claim is naturally rendered in $\mathcal{L}_{LK}$ as

(NonOm1) For some ψ, $\psi \wedge \neg K\psi$.

Suppose we now assume that the knowledge operator K is factive (i.e. $K\varphi \rightarrow \varphi$) and distributes over conjunction (i.e. $K(\varphi \wedge \psi) \rightarrow (K\varphi \wedge K\psi)$) and that provability implies necessity (i.e. $\vdash \varphi \;\therefore\; \vdash \Box\varphi$). If we now take (NonOm1) as a premise for reductio for some fixed ψ, then the corresponding instance of (Om1) can be derived via the following argument:

(1)	$\psi \wedge \neg K\psi$	NonOm1
(2)	$(\psi \wedge \neg K\psi) \rightarrow \Diamond K(\psi \wedge \neg K\psi)$	KP1 (taking $\varphi \equiv \psi \wedge \neg K\psi$)
(3)	$\Diamond K(\psi \wedge \neg K\psi)$	(1), (2)
(4)	$K(\psi \wedge \neg K\psi)$	premise for reductio
(5)	$K\psi \wedge K\neg K\psi$	distributivity of K
(7)	$K\psi$	(5)
(8)	$K\neg K\psi$	(5)
(9)	$\neg K\psi$	factivity for K
(10)	$\neg(K(\psi \wedge \neg K\psi))$	(4), (5)–(8)
(11)	$\Box\neg(K(\psi \wedge \neg K\psi))$ $=_{df} \neg\Diamond(K(\psi \wedge \neg K\psi))$	necessitation (10)
(12)	$\neg(\psi \wedge \neg K\psi)$	(1), (3)–(11)

(13) $\psi \rightarrow \neg\neg K\psi$ (12) (intuitionistic logic)

(14) $\psi \rightarrow K\psi$ (13) (classical logic)

The observation that all instances of (Om1) can be derived from the assumption of (KP1) in the manner just described has come to be known as the Knowability Paradox. It is conventionally taken to be paradoxical in virtue of the fact the while (KP) is not clearly false, it entails an apparent falsehood on the basis of what appear to be weak background assumptions about the behavior of the knowledge and necessity operators. As such, most modern commentators view the Paradox as casting into doubt any view which embraces any principle connecting truth and knowability resembling (KP).

This is not, however, the context in which the significance of statements of the form (NonOm1) was observed. Fitch (1963) attempted to formulate an analysis of the notion of value which (for reasons which need not concern us here) would have been trivialized by the existence of unknowable truths. Fitch noted that the assumption of (NonOm1) for a particular proposition ψ implies the existence of a truth χ which cannot be known. This follows by taking $\chi \equiv \psi \wedge \neg K\psi$ and carrying through steps (4)–(11) in the preceding derivation. Fitch viewed this derivation not as a reductio of (KP), but rather as a demonstration that in the presence of the assumptions about the K and $\Box$ operators mentioned above, the existence of an unknown truth entails the existence of a necessarily unknown (or *unknowable*) truth.[1]

We will suggest below that the derivation (4)–(11) anticipates what we will present as a positive result about the existence of non-constructed proofs. This is not, however, the manner in which the complete derivation (1)–(14)—which we will subsequently refer to as the *Fitch derivation*—is viewed in the contemporary literature on the Knowability Paradox. This can be taken to begin with Hart and McGinn's (1976) rediscovery of this derivation and their subsequent recasting of it as an antimony for a view they refer to as *verificationism*. As understood by Hart and McGinn, this term describes a class of views according to which the notion of truth is subject to a so-called *epistemic constraint*—i.e. a principle which states that in order for a statement to be true, it must satisfy a criterion which makes substantial use of epistemic notions such as verification, knowledge or provability. We will follow Dummett's (e.g. Dummett 1982) terminology in referring to an opponent of verificationism as a *realist*—i.e. a theorist who holds that the truth value of a statement is fixed independently of our capacity for knowing or verifying it.

Hart and McGinn suggest that (KP) can be taken to be a schematic formulation of verificationism. Although we will take issue with this claim below, a statement closely resembling this principle can indeed be found in Dummett's (1993) "What is a theory of meaning? (II)." Hart and McGinn thus argue that the Fitch derivation can be taken to show that verificationists must also accept (Om). But on the interpretation offered above, this principle will presumably be rejected by both realists and verificationists

[1] Fitch attributes the foregoing statement to an anonymous referee, who subsequent investigation has revealed to be Alonzo Church. Documentary evidence recently discovered by Salerno (2008) also suggests that the proof Church had in mind was essentially the one given here as (4)–(11). We note in passing that the result stated in the text appears to be somewhat weaker than what this derivation actually establishes. For note that since assumption (4) is discharged at line (10), (11) does not require the premise that some truth is unknown.

alike. On this basis, Hart and McGinn propose that the Fitch derivation ought to be taken as a reductio of verificationism.

1.2 An approach to the Paradox

We are not now and never have been verificationists. Like most contemporary commentators on the Knowability Paradox, our stake in its resolution is thus indirect. Unlike many commentators, however, we are willing to take up the cause of intuitionism. The potential connection between the Paradox and intuitionism derives not only from the fact that virtually all contemporary proponents of verificationism are also proponents of intuitionistic logic and mathematics, but also from the observation that intuitionists have historically sought to define truth in terms of epistemic notions in a manner which is at least reminiscent of (KP). Standard treatments of intuitionistic logic (e.g. Troelstra and van Dalen 1988; Dummett 1977) thus often begin by laying down that the assertion of a logically unstructured mathematical statement P should be understood to be equivalent to the claim that we are able to prove P. It is thus accurate to formulate the fundamental intuitionistic notion of truth as

(1) P is true if and only if P is verifiable.

subject to the understanding—which we will argue below is shared among practicing verificationists—that to say P is verifiable is to say that we can construct a proof of P.

Our concern with the Knowability Paradox arises not from the fear that it undermines a specific formulation of verificationism such as Dummett's anti-realism, but rather that it might infect the logical substructure of an entire species of such views.[2] To appreciate why this might be a concern, it is important to realize not only that (1) serves to express the sort of epistemic limitation which intuitionists wish to impose on truth, but also that this principle is traditionally taken to serve as the basis for understanding the meanings of the intuitionistic connectives. In particular, it is generally accepted that a proper intuitionistic interpretation of the logical connectives must conform to (1) in the sense that it must specify the conditions under which a complex statement is verifiable in terms of the conditions under which its constituents are verifiable.

Within the context of the now standard Brouwer–Heyting–Kolmogorov [BHK] interpretation of the intuitionistic connectives, the notion of verification is analyzed in terms of the notion of *constructive proof*. The exact nature of such proofs has, of course, long been a matter of controversy within intuitionism (cf., e.g., Weinstein 1983). However, the BHK interpretation itself merely seeks to specify what would constitute a proof of a logically complex statement in terms of the so-called *proof conditions* of its constituents. This gives rise to the familiar inductive characterization of such conditions whose clauses are typified by

[2] In particular, our own interests lie with the class of proposals known as *proof theoretic semantics* (cf., e.g., Kahle and Schroeder-Heister 2006) which have grown out of attempting to make Dummett's general views about meaning more precise using technical apparatus from contemporary proof theory. Since the literature of the Knowability Paradox only interfaces with these developments in a peripheral manner, we have opted to keep the current discussion as general as possible.

(BHK$_\wedge$) a proof of $\varphi \wedge \psi$ is constituted by a proof of φ and a proof of ψ
(BHK$_\rightarrow$) a proof of $\varphi \rightarrow \psi$ is constituted by a method which allows us to transform any proof of φ into a proof of ψ

By giving clauses of this sort for each of the propositional connectives and quantifiers, the BHK interpretation may be taken as specifying the meaning of a logically complex statement in terms of the proof conditions of its constituents. It thus follows that if intuitionistic truth is to be understood as verifiability, and verifiability is to be understood in terms of constructive provability, then for an arbitrary statement φ, intuitionists will accept the principle

(KP2) φ is true if and only if φ is provable.

where the sense of the right-hand side is fixed according to the structure of φ by the BHK interpretation. However, since provability can be rendered in terms of proof existence, we will argue below that the views of contemporary verificationists are most faithfully represented by a principle resembling

(KP3) For all φ, if φ is true, then there exists a proof of φ.

It might first appear that (KP1) and (KP3) correspond to very different means of expressing an epistemic constraint on truth. And as such, one might reasonably wonder what relationship the latter schema bears to the Knowability Paradox in both a rhetorical and a logical sense. With respect to the former issue, it will be our first goal in this paper to show that (KP3) more accurately expresses the views of contemporary verificationists than does (KP1). As we have already pointed out, this is indeed a sensitive textual issue because a statement closely resembling (KP) does appear in Dummett's (1993) "What is a theory of meaning? (II)." But in attempting to assess whether this statement can reasonably be taken to be a paradigmatic expression of the views of contemporary verificationists, it is reasonable to look more broadly at the writings not only of Dummett but also of the other self-describing verificationists Dag Prawitz and Per Martin-Löf. Once this step is undertaken, it becomes apparent that these theorists customarily elect to express their views about truth using the notions of proof and existence as opposed to those of knowledge and possibility. As we will argue in the next section, there is thus a strong textual basis for preferring (KP3) over (KP1) as a means of schematizing the views of practicing verificationists.

If verificationism is indeed imperiled by concerns related to those raised by the Knowability Paradox, this would be because it is possible to formulate an analogous paradox premised on (KP3). However, if the notion of *knowability* in terms of which (KP) is conventionally formulated is understood in terms of proof existence in the manner suggested by (KP3), we must also ask how the notion of *knowledge* which figures in (NonOm) should be understood. We will argue below that relative to an anti-realist notion of truth, the failure of omniscience is most naturally understood in terms of the existence of true but *unproven* statements. The existence of such statements can be expressed as follows:

(NonOm2) For some ψ, ψ is true and no proof of ψ has been constructed.

(NonOm2) will presumably be accepted by realists not only because of the existence of formally undecidable statements, but also for the more mundane reason that they will likely hold that there are many decidable mathematical statements which, while true, remain as yet underived. However, gauging the likely attitude of intuitionists towards this statement is less straightforward. One might initially think that acknowledging the existence of true but unproven statements appears to stand in direct conflict with Brouwer's dictum that "there are no unexperienced truths" (Brouwer 1983). But it is often acknowledged that intuitionistic truth is a subtle notion and that the founding figures of intuitionism may not have agreed about its proper analysis.[3] Taking this into account, we will argue below that there is a straightforward argument based on assumptions that would be accepted by most intuitionists—and in particular the theorists with whom we will be concerned—to the effect that (NonOm2) ought to be accepted.

The question thus naturally arises whether, when taken in conjunction with (KP3), (NonOm2) entails any statement parallel to (Om) which is in conflict with common intuitions about our everyday concepts of truth and knowledge. In the framework which we will develop below, such a statement would take the form

(Om2) For all φ, if φ is true, then we have constructed a proof of φ.

It will be our second substantial goal to show that no statement of this kind is derivable from (KP3) and (NonOm2) via any means by which either a realist or a verificationist would consider legitimate. In particular, we will argue that there is no way to reconstruct the reasoning of Fitch's original derivation to derive (Om2) from (KP3) and (NonOm2). We will also provide a consistency proof which demonstrates that there is no other means of deriving a contradictory statement from these principles.[4]

We will endeavor to do this in a way which is as precise and rigorous as possible. Since the reasoning of the Fitch derivation is reasonably involved, we will have to find a means of formalizing (KP3) and (NonOm2) in a language in which such principles can be stated directly. This task poses two distinct challenges. On the one hand, there is no way of formulating these principles in a conventional modal language such as $\mathcal{L}_{LK}$ as this language contains no means of quantifying over proofs. And on the other, once the relationship between intuitionism and verificationism has been highlighted, one is immediately drawn to ask whether the use of classical logic at the last step of the original derivation of the Paradox is legitimate. We will now discuss both of these matters before turning to an in-depth analysis of the views of the verificationists and the technical exposition of the system we will use to represent them.

---

3 Cf., e.g., Raatikainen (2004) for discussion of this point.

4 Our closest anticipators in proposing such a resolution to the Paradox are Cozzo (1994) and Hand (2003, 2008). We will consider Hand's proposal explicitly in Sect. 3.1. However, since there will be no convenient point to return to Cozzo's proposal, we will simply note in passing that on the basis of concerns similar to those we will discuss in Sect. 2, Cozzo argues that (KP) should be formalized using a quantifier over what he refers to as *ideal verifications*. The (KP)-like principle which he formulates thus mirrors our (KP3) in form. However, Cozzo also elects to treat knowledge as an unanalyzed propositional operator as opposed to analyzing it in terms of our failure to have constructed a proof of ψ in the manner of (NomOm2). For this reason, it is difficult to directly compare his approach to the Paradox to the one adopted here.

1.3 Knowledge, proof and explicit modal logic

Another point we wish to highlight in this paper is that in order to fully understand the significance which the Knowability Paradox holds for verificationism, it is necessary to analyze the particular notions of knowledge and knowability relative to which practicing verificationists will assent to (KP) and (NonOm). We have already taken a step in this direction by proposing that (KP3) and (NonOm2) better approximate the sense which such theorists will assign to these principles than do (KP1) and (NomOm1). Implicit in these proposals are the claims that verificationists will understand the notions of knowledge and knowability as follows:

(2) (a) φ is known if and only if a proof of φ has been constructed
(b) φ is knowable if and only if there exists a proof of φ

(2a, b) may seem unexpected as analyses of the familiar epistemic concepts which the Knowability Paradox is conventionally framed as drawing into conflict. However, subject to their assent to (KP) itself, it is not hard to see why these analyses represent the perspective on knowledge and knowability which the verificationists are likely to hold.

For note that there is no reason to think that verificationists will not agree with many realists in holding that the everyday notion of knowledge can be analyzed in the traditional manner as a species of justified, true belief. Like a realist, a verificationist may acknowledge that such an analysis is incomplete in the sense that it fails to explain why certain intuitively admissible forms of justification are sufficient to produce knowledge while others are not. However, unlike many realists, verificationists will presumably be in possession of a reasonably fleshed out analysis of justification in the form of their preferred notion of constructive proof.

It is typically supposed that constructive proofs satisfy the following conditions:

(I) A constructive proof is a structured entity $\mathcal{D}$ constituted by a sequence (or tree) of statements or propositions, one of which is identified as its conclusion $Concl(\mathcal{D})$.
(II) Such proofs may be constructed by epistemic agents and if an agent i constructs a proof $\mathcal{D}$ such that $Concl(\mathcal{D}) = \varphi$, he is thereby justified in believing φ.
(III) If there exists a constructive proof $\mathcal{D}$ such that $Concl(\mathcal{D}) = \varphi$, then φ is true.

If we take these properties of constructive proofs into account, then it is not hard to see why verificationists will be likely to accept the analyses given in (2a, b). For recall that the Knowability Paradox turns on the fact that to say that φ is *knowable* is equivalent to saying that φ is possibly known. But now note that according to the justified-true-belief analysis of knowledge, to say that φ is possibly known is equivalent to saying that there possibly exists some sort of evidence which would serve to justify φ. According to the verificationists, the notion of evidence should be understood in terms of that of constructive proof. And from this it follows that to say φ is knowable for a verificationist is equivalent to saying that there exists a constructive proof with conclusion φ, wherein the existential quantifier is taken to range over not only proofs which are actual in the sense of having been constructed, but also over those which are merely potential in the sense that they could be constructed.

As we discussed above, there presumably are statements ψ which are knowable in the sense just described but which we do not know to be true. Note, however, that if we have constructed a proof of ψ, by (II) we would be justified in believing it. And since by (III), we would then also be assured that ψ was true, the justified-true-belief analysis would then imply that ψ is known. As such, the fact that ψ is not known must presumably be explained in terms of the fact that no proof of ψ is available to us. Since proofs are the sort of things we construct, this in turn must be explained by the fact that we have not constructed a proof of ψ. And from this, it follows that for a verificationist to accept that ψ is known, it is sufficient to require both that there is a proof of ψ and also that this proof has been constructed.

The foregoing considerations do not, of course, demonstrate that (2a, b) are unassailable as analyses of our everyday concepts of knowledge and knowability. For note that it follows by condition (III) that since the existence of a constructive proof of φ implies that φ is true, such a proof serves as an indefeasible form of evidence for φ. If it is required that in order for an agent i to know φ, i must grasp such a proof, it follows that verificationists will be unable to account for cases in which we might credit i with knowing φ on the basis of a form of evidence which does not possess this property. This will paradigmatically be the case, for instance, when φ is an empirical proposition, and the justification which i possesses for φ is either premised on other empirical hypotheses or perceptual judgements.

(2a) thus highlights how wide the rift is between the accounts of knowledge and justification which are likely to be favored by verificationists and those employed in mainstream epistemology. For as we will see in the next section, contemporary verificationists do not think of (KP) as an isolated principle, but rather view it as a consequence of the view that a theory of meaning should be modeled on the sort of proof theoretic explanations of the intuitionistic connectives discussed above. The notion of proof which is most directly applicable to such an interpretation is that of a formal derivation in intuitionistic logic or mathematics. If the verificationists wish to formulate a theory of meaning which is applicable to empirical statements, they will presumably have to broaden their understanding of constructive proof to include demonstrations which can verify empirical statements.

The question of how an appropriate notion of empirical demonstration might be extrapolated from the traditional notion of constructive proof has been discussed by, e.g., Dummett (1993, 1991) and Prawitz (1998c). However, without a fully fleshed out account of this kind in hand, it is difficult to ascertain whether (2a) could ever serve as the basis for an analysis of knowledge which is in general accord with everyday intuitions. With that said, however, the Knowability Paradox itself threatens to undermine verificationism on the basis of only a single instance of (NomOm) in which ψ may be either an empirical or non-empirical proposition. And thus since the problem presented by the Paradox does not appear to turn on the general applicability of (2a) to empirical statements, we will henceforth ignore the problems our apparent knowledge of such statements raises for verificationism.

If we accept that (2a, b) serve at least as general templates for the form which the verificationists' analyses of knowledge and knowability must take, it follows that any formal language in which we might hope to represent their views must itself allow us to speak directly of proofs and the justificatory role they bear to sentences. The language we will adopt to formalize (KP3) and (NonOm3) is that of a member of a class of systems known as *explicit modal logics*. In particular, we will introduce a system which descends from Artemov's (2001) Logic of Proofs [LP] which was itself inspired by traditional Provability Logic (in the sense of Smorynski 1985 and Boolos 1992). Systems of the latter sort possess a single modal operator $\Box$ which is intended to interpret the provability predicate of a first-order arithmetic system T such as Peano arithmetic [PA]. More precisely, the meaning of $\Box$ in the context of classical Provability Logic is fixed by giving a mapping $(\cdot)^*$ from the modal language into that of T which assigns a fixed arithmetic statement P^* to each atomic formula P, commutes with connectives, and is such that

(3) $(\Box\varphi)^* = \exists x Proof(x, \ulcorner\varphi^*\urcorner)$[5]

where $Proof(x, y)$ is a proof predicate for T.

It is well known that if $T = \mathsf{PA}$, then the set of modal statements provable in T for all choices of $(\cdot)^*$ corresponds to the Gödel-Löb logic GL. Although GL has occasionally been discussed as an epistemic logic itself (cf., e.g., Binmore and Shin (1993)), it is generally thought to be ill-suited to this purpose. For it may readily be shown that GL is inconsistent with the factivity (or "reflection") axiom $\Box\varphi \rightarrow \varphi$. This axiom is assumed to hold for the knowledge operator K employed in the Fitch derivation and is also generally taken to express an intuitively valid principle about informal provability. However, it is not the specific axioms of GL which serve to motivate LP, but rather the suggestion that provability can be treated as a modal operator. This observation is generally attributed to Gödel, who proposed a means of interpreting intuitionistic propositional calculus [IPC] in a language $\mathcal{L}_{\mathcal{B}}$ containing the single propositional operator $\mathcal{B}$ subject to the interpretation

(4) $\mathcal{B}\varphi \rightleftharpoons \varphi$ is informally provable[6]

Gödel defined a mapping $(\cdot)^g$ from the language of IPC into $\mathcal{L}_{\mathcal{B}}$ such that for all φ, φ is derivable in IPC if and only if φ^g is derivable in the classical modal logic S4.[7]

Since S4 contains the factivity axiom, it is reasonable to think that the Gödel interpretation might be used to reconstruct the original derivation of the Knowability Paradox. For note that relative to (2b), $\mathcal{B}\varphi$ can be taken to analyze "φ is knowable." If we additionally propose that (KP) should be understood as (KP2), then $\mathcal{L}_{\mathcal{B}}$ provides a means of directly formalizing this statement as $\varphi \rightarrow \mathcal{B}\varphi$. However, unlike the language $\mathcal{L}_{LK}$ used in the original derivation, $\mathcal{L}_{\mathcal{B}}$ contains only a single operator. And thus it appears that in this language we are left without a means of distinguishing between knowledge and knowability.

[5] Here and below we omit overlines on gödel numbers of formulas to improve readability.

[6] The caveat "informal" is necessary because the $\mathcal{B}$ operator cannot be taken to represent provability in any system containing even a small fragment of PA. This follows immediately from the fact that since $\mathcal{B}\varphi \rightarrow \varphi$ is an axiom of S4, S4 $\vdash (\mathcal{B}\bot \rightarrow \bot) \equiv \neg\mathcal{B}\bot$. If it were possible to interpret $\mathcal{B}$ as $Bew_T(x)$ for any system T containing a sufficiently large fragment of PA, it would then follow that T proved its own consistency statement $\neg Bew_T(\ulcorner\bot\urcorner)$, in violation of Gödel's Second Incompleteness Theorem. An analogous caveat will apply to the system QLPE which we will employ below—cf., footnote 36.

[7] The left-to-right direction is credited to Gödel (1986). The right-to-left direction is due to McKinsey and Tarski (1948).

On the basis of the foregoing analysis of the difference between knowledge and knowability, this failure can be attributed to the fact that while $\mathcal{B}\varphi$ is intended to express a form of implicit quantification over proofs, this quantification is not expressed in the language $\mathcal{L}_{\mathcal{B}}$ itself. One of the hallmarks of explicit modal logics such as LP is that they employ a form of modality which directly expresses the relationship between a proof and the statement it serves to verify. So as opposed to a single modality like $\Box$ or $\mathcal{B}$, the language of LP possesses an infinite family of so-called *explicit modalities* t_1:, t_2:, ... Herein the t_is are potentially complex expressions known as *proof terms* which can figure in basic modal statements of the form $t : \varphi$. The intended interpretation of such a statement is

(5) $t : \varphi \rightleftharpoons t$ is a proof of φ

LP was originally provided with an arithmetic interpretation $(\cdot)^+$ whereby proof terms are interpreted as codes of formal proofs and the ":" operator is treated as expressing the relation which obtains between (the code of a) proof t and (the code of a) statement φ just in case the former is a proof of the latter. In particular, $(\cdot)^+$ maps LP statements of the form $t : \varphi$ to arithmetical statements of the form $Proof(\overline{t^+}, \ulcorner\varphi^+\urcorner)$. Relative to this class of interpretations, it is possible to prove an arithmetic completeness theorem for LP which is similar in form to that alluded to above for GL. However, one important consequence of the fact that statements in the range of $(\cdot)^+$ are quantifier-free is that the set of modal principles which are arithmetically valid (i.e. whose images are provable for all choices of $(\cdot)^+$) is different from those which are arithmetically valid with respect to all GL interpretations $(\cdot)^*$. In particular, the arithmetic interpretations of all instances of the so-called *explicit reflection principle* $t : \varphi \rightarrow \varphi$ are provable in PA and are accordingly adopted as axioms of LP. A consequence of this is that although the systems of explicit modal logic which we will be discussing have their intellectual roots in provability logic, their axioms more resemble those of S4 than GL.

The system which we will ultimately introduce to formalize (KP3) and (NonOm2) is a minor modification of a quantified version of LP—known as QLP—which was introduced by Fitting (2004, 2006). This system possesses not just explicit modalities, but also proof quantifiers of the form $(\exists x)x : \varphi$ with the intended interpretation

(6) $(\exists x)x : \varphi \rightleftharpoons$ there exists an x such that x is a proof of φ

The addition of quantifiers to the language of LP allows us to straightforwardly formalize (KP3) in QLP as

(KP4) For all φ, $\varphi \rightarrow (\exists x)x : \varphi$.

However, the addition of proof quantifiers alone is not sufficient to distinguish between statements which are provable and those whose proofs have been constructed. In order to express this distinction, we will add to the language of QLP a predicate $E(x)$ of proofs with the intended interpretation

(7) $E(t) \rightleftharpoons$ the proof denoted by t has been constructed

Statements of the form $E(t)$ will serve as well-formed formulas in a system called QLPE which we will introduce below. In the language of this system, (NonOm2) can naturally be formalized as

(NonOm3) For some ψ, $\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]$.

We will defer the formal exposition of QLPE until Sect. 3 where we will pursue the two technical goals announced above—i.e. 1) to show that the QLPE formalization of (Om2) (i.e. (Om3) For all $\varphi, \varphi \rightarrow (\exists x)[E(x) \wedge x : \varphi]$) cannot be derived from (KP4) and (NonOm3); 2) to show that the latter two principles are formally consistent. We will argue that taken together these results demonstrate that the Knowability Paradox poses no threat to at least one plausible means of formulating verificationism. This caveat is necessary for several reasons. For although we will argue that the formalizations we propose represent reasonable reconstructions of the core tenets of contemporary verificationism, the proponents of this view are not in complete agreement about all the issues which pertain to the proper formalization of these principles. These differences do not arise at the level of their endorsement of a principle like (KP3) which attempts to analyze truth in terms of proof existence. Rather, as we will see in Sect. 2, they come out with respect to the significance which the verificationists assign to the notion of constructive proof itself.

One consequence of the formalization we will provide is that it allows us to show how it is possible to derive from (KP4) and (NonOm3) the conclusion that there exist non-constructed proofs. The status of such "potential proofs"—i.e. derivations which have not been carried out by any actual epistemic agents but which are potentially constructible in the sense of falling under some general characterization of constructive proof—has been debated in the literature of contemporary verificationism. Although the notion of potential proof has been embraced by Prawitz and Martin-Löf, the status of such entities is contentious. In particular, Dummett has repeatedly suggested that the acknowledgement that there exist merely potential proofs represents a substantial concession to realism. In Sect. 4 we will endeavor to show that acknowledging the existence of such proofs need not be taken to have the consequences which are of concern to Dummett. As such, we believe that the real significance of the Knowability Paradox is not to undermine verificationism, but rather to illuminate a way in which this view is committed to the existence of potential proofs.

Before embarking on these tasks in earnest, it will finally be useful to highlight one other aspect of our approach to the Paradox which differs from that of some other commentators. As we mentioned in passing above, the original Fitch derivation employs classical logic in the transition from step (13) to (14). From the time of its rediscovery by Hart and McGinn, this derivation has repeatedly been used as a means of critiquing verificationism. But since verificationists generally reject classical logic, it is natural to ask whether the very theorists against whom these critiques are targeted must accept the conclusion of the Paradox. This issue has influenced the literature on the Paradox in a variety of ways. For instance, in his original article, Hart (1979) acknowledges that a case can be made that intuitionists should simply accept the principle (Om) which he otherwise argues should be taken as a reductio of verificationism.[8] Hart remarks

[8] For note that if we interpret the conditional in (KP1) according to ($BHK_{\rightarrow}$), then (Om1) appears to report the innocuous finding that were to possess a proof of φ then we could uniformly construct a proof of $K\varphi$ (presumably by reflecting on our original proof). This observation suggests that intuitionists possess a natural response to the Paradox—i.e. they should simply accept (Om) on the fact that it reports a statement which they should be willing to accept on its own lights. We will henceforth refer to this as the *bullet biting*

that he finds this resolution untenable since, understood informally, (Om) expresses a principle which is radically out of keeping with our ordinary intuitions about knowledge. As such, he proposes that the Paradox should be taken as a genuine antinomy for verificationism.

Williamson (1982), on the other hand, attempts to intervene on behalf of the verificationists by arguing that they should reject the classically mediated step. This would save them from having to accept (Om) at the expense of accepting step (13) (i.e. $\psi \rightarrow \neg\neg K\psi$). However a number of theorists have replied to Williamson by arguing either that (13) is already an embarrassing consequence of verificationism or that it is possible to formulate a related paradox based on the existence of undecided statements which is based on intuitionistically valid reasoning (cf., e.g., Percival (1990)).

Our ultimate view will be that the debate over the use of classical logic in the Fitch derivation represents something of a red herring. For in order for the Paradox to be deployed by realists as an argument against verificationism, it must presumably be formulated in a language which the realists can themselves understand. However, it is open to a realist to claim that he does not understand the meanings of the intuitionistic connectives. This consideration appears to favor the adoption of a classical background logic for use in assessing the consequences of (KP) and (NonOm). Verificationists can, of course, reply that they are analogously unable to understand the classical meanings of negation and implication which underwrite the legitimacy of the inference from (13) to (14). And for this reason, it should be left open for them to adopt the bullet biting resolution described in footnote 8 for accepting (Om). With this said, however, we are in agreement with Hart that in taking this step, verificationists put themselves in the position of embracing a principle which is in strong conflict with our everyday intuitions. In particular, such a step may seem overly drastic, as it appears to undermine the verificationist's ability to express his views to the realist using a common language.

There is a long tradition of providing classical interpretations of intuitionistic logic and mathematics in a language which classical theorists can understand. It is often proposed that such interpretations can be used as a means of overcoming communication breakdowns of the sort we have just described. In fact (KP3) itself can also be understood as part of an attempt to interpret the notion of intuitionistic truth in terms which a realist can presumably understand—i.e. those of proof and existence. This analysis may be partial since the significance which an intuitionist attaches to the occurrence of "there exists" in (KP3) may itself differ from that which it is assigned by a realist. However, we believe that verificationists should not balk at the application of classical logic to statements in the language of QLPE for purposes of reasoning about the Paradox. For such a language need not be understood as the native tongue of a verificationist, but rather that of an idiom which he is temporarily adopting in order to thwart a threat which has been imposed from outside.[9]

Footnote 8 continued
resolution of the Paradox. Hart's negative appraisal notwithstanding, some commentators have proposed that intuitionists ought to adopt such a resolution—e.g. Martino and Usberti (1994).

[9] Our ultimate view on these matters is in fact slightly more complex. We will return the significance of formulating QLPE using intuitionistic logic in Sect. 4.2.

2 Neo-verificationism, knowledge and proof

Our primary goal in this section is to convince the reader that our proposed interpretations of the premises of the Knowability Paradox—i.e. (KP4) and (NonOm3)—have a strong basis in the writings of contemporary theorists who describe themselves (or are often described) as verificationists. As we have mentioned above, we will take Michael Dummett, Dag Prawitz and Per Martin-Löf to be the paradigmatic exemplars of this class. These theorists have all famously endorsed versions of an epistemic constraint on truth in the sense explained above. However, the application of the term "verificationism" to describe their views is at least slightly problematic due to the prior application of this term to describe the views of Ayer and the theorists of the Vienna Circle.[10] To avoid any confusion on this issue, we will henceforth adopt the term "neo-verificationism" not only to differentiate the views of Dummett, Prawitz and Martin-Löf from these earlier theorists, but also to demarcate several positive factors which serve to distinguish their views.[11]

Each of the neo-verificationists can be taken to be a proponent of the view which is often described as semantic anti-realism. This view can be characterized as the denial of realism as formulated above—i.e. as the view that the truth of a statement cannot be characterized in a manner which transcends any means by which it can potentially be verified. Dummett (e.g. Dummett (1978)) has repeatedly cited Wittgenstein's slogan "meaning is use" as a motivating factor for adopting such a position. For if the meaning we assign to a statement is taken to be significantly constrained by the occasions on which it may be legitimately asserted, then it becomes inadmissible to analyze its meaning in terms of truth conditions whose satisfaction cannot be ascertained in all cases. According to this view, the problem with traditional analyses of meaning in terms of truth conditions is exemplified by undecidable statements—i.e. ones for which we have no means of determining whether their truth conditions are satisfied or fail to be satisfied. For in such cases, it seems that we can have no justification for either asserting such a statement or its negation.

Perhaps the most significant feature which the neo-verificationists have in common is their shared view that this problem may be overcome by associating the meaning of a statement not with its classical truth conditions but rather with some notion of a verification condition. In this setting, the notion of a verification is taken to generalize that of a constructive proof which is also "direct" in (roughly) the sense that its combinatorial structure reflects the logical structure of its conclusion. The notion of a direct verification is paradigmatically formalized via the notion of a canonical or normal proof in a Gentzen-style natural deduction system. In this context, the introduction rules of such a system are taken as definitionally giving the meaning of the logical constants, while the elimination rules are justified on the basis of their structural features relative to the introduction rules. For present purposes, it will be this shared intellectual legacy—which notably differs from that of the early twentieth century verificationists both with respect to its connection with intuitionism and also with subsequent work

[10] Martin-Löf (1996, p. 34) has been most explicit about drawing attention to the historical basis (and ultimate clash) of these applications of the term "verificationism."

[11] This term has also been employed by Usberti (1995) in a similar manner.

in proof theory in the tradition of Gentzen—which we will take to unify the views of the neo-verificationists.

Although it is the neo-verificationists who are the presumptive targets of the Knowability Paradox, the literature on this subject interfaces only indirectly with the aspects of their views we have just described. In this latter-day setting, the paper "What is a theory of meaning? (II)" Dummett (1993) can be taken as the canonical exposition of the neo-verificationist position.[12] This paper contains the following formulation of an epistemic constraint on truth:

(K) If a statement is true, then it must be in principle possible to know that it is true. (Dummett 1993, p. 61)

Principle (K) clearly proposes to analyze truth in terms of the notions of knowledge and possibility. And thus this statement appears to provide a clear basis for schematizing neo-verificationism by using propositional knowledge and possibility operators of the sort employed in (KP1).

In order to understand the role which (K) plays with respect to Dummett's views as a whole, it should be kept in mind that he repeatedly cites the intuitionistic understanding of the logical connectives as a paradigm for the form which a theory of meaning should take.[13] On this basis, Dummett proposes within Dummett (1993) itself that to grasp the meaning of an arbitrary statement is to grasp what would count as a proof of it:

> The intuitionistic explanations of the logical constants provide a prototype for a theory of meaning in which truth and falsity are not the central notions. The fundamental idea is that a grasp of the meaning of a mathematical statement consists, not in a knowledge of what has to be the case . . . for the statement to be true, but in an ability to recognize, for any mathematical construction, whether or not it constitutes a proof of the statement; an assertion of such a statement is to be construed, not as a claim that it is true, but as a claim that a proof of it exists or can be constructed. (Dummett 1993, p. 70)

This passage is a typical illustration of the way in which the neo-verificationists propose to ground a general theory of meaning in an analysis of the relationship between a statement and the proofs which serve to verify it. General formulations of this view inevitably involve quantification over proofs in the sense that in order to grasp the meaning of φ, it is asserted that we are able to ascertain for *any* proof whether it constitutes a proof of φ. The prior passage appears to suggest that such quantification can be used to eliminate the need for the notion of truth within a theory of meaning.

[12] Dummett is the rhetorical target of the paper (Hart 1979) in which Hart rediscovers (and repurposes) the Fitch derivation. And although Dummett (1993) is not explicitly cited as the source for the view which Hart seeks to critique, it contains a discussion of verificationism which is similar to the view described in this paper.

[13] The neo-verificationists all appear to think that the analysis of meaning based on canonical proof described above should be thought of as a further refinement of ideas already implicit in the BHK interpretation—cf., Dummett (1978) and Prawitz (1977). However, since the specific features which they assign to canonical proofs will not play a substantial role here, we will generally not bother to issue caveats which flag the (often very complex) ways in which this notion contributes to their views.

However at other times Dummett argues instead that truth should be *defined* in terms of proof—e.g.

> A sentence may then be said to be true if and only if there exists a construction that constitutes a proof of it. (Dummett 1982, p. 234)
> [A]n understanding of a mathematical statement consists in the capacity to recognize a proof of it when presented with one; and the truth of such a statement can only consist in the existence of such a proof. (Dummett 1977, p. 4)

In drawing attention to these passages, we do not intend to challenge the claim that the sort of epistemic constraint which Dummett and the other neo-verificationists seek to impose on truth can be legitimately expressed through a statement like (K). However, Dummett's understanding of this principle is clearly informed by his other views about the role which the notion of proof should be assigned within a theory of meaning. These, in turn, are informed by the view that such a theory should be modeled on the notion of truth employed within intuitionistic logic and mathematics. And for this reason, there seems to be little motivation for thinking that a principle like (K) involving the unexplicated notion of knowability ought to be taken as the fundamental means of expressing how an epistemic constraint on truth ought to be formulated.

If we look more broadly at the writings of the other neo-verificationists, it is easy to find evidence not only for this conclusion, but also for the view that proofs are taken to play the epistemic role of verifications within a theory of meaning. Both tendencies are clearly evident in the following passage from Martin-Löf:

> A is true is taken to mean that there exists a proof of A, a proof which need not necessarily be direct or canonical. The term proof is of course synonymous with verification here. This definition of the notion of truth of a proposition reduces it to two notions, namely the notion of proof or verification and the notion of existence, and it is because of this that it is very natural to use the term verificationism in connection with the theory of meaning for intuitionistic logic: the term verificationism is used to stress the fact that the notion of truth is not taken as a primitive notion, like a truth conditional theory of meaning, but is rather defined in terms of an underlying notion of verification by the principle that A is true if there exists a proof of A. (Martin-Löf 1995, p. 191)

The following passage, which appears as part of an extended commentary on Dummett's (K), is equally explicit in its equation of truth with provability:

> So, we must ask ourselves, "How is the notion of truth of a proposition to be defined?" This is precisely the problem of Dummett that I started by quoting, namely, of how the notion of truth, within a theory of meaning in terms of verification, should be explained. The answer is most simply given in the form of equations

$$\begin{aligned} A \text{ is true} &= \text{there exists a proof of } A \\ &= \text{a proof of } A \text{ can be given} \\ &= A \text{ can be proved} \\ &= A \text{ is provable (Martin-Löf 1998, pp. 111–112)} \end{aligned}$$

in which the equality sign signifies sameness of meaning.[14]

Although the views of Prawitz and Martin-Löf differ in many respects, Prawitz is also quite explicit in his endorsement of the view that truth should be defined in terms of proof existence. The following statement is characteristic:

> Once we accept the notion of provability as legitimate, it is hardly controversial within verificationism that the truth of a proposition is to be identified with provability or existence of proofs. . . . [P]roofs as here understood are something that in principle can be known by us, and hence there is no talk about in principle unknowable proofs. (Prawitz 1998d, p. 48)

In order to appreciate the relationship between a statement such as this and Dummett's (K), a bit more must also be said about the difference between how Prawitz and Martin-Löf view the notion of proof itself. For Martin-Löf it is judgements which are the fundamental epistemic concept, while proofs derive their epistemic significance only from their relation to judgements. For Prawitz, on the other hand, proofs themselves are epistemic in nature in the sense that they are intrinsically the sorts of entities which can potentially be grasped by us—a point which he goes out of his way to emphasize in the prior passage.

The foregoing passages are typical of many others in the writings of the neo-verificationists in which the notion of truth is analyzed in terms of proof existence. We take them to illustrate not only that the principle we have labeled as (KP3) reflects the views of these theorists, but also that they are in agreement that such a principle is to be taken as having the same general conceptual significance as (KP). But it should also be noted that from a purely logical perspective (KP) and (KP3) appear quite different. For while (KP) uses the combination of propositional operators "possibly known" to express knowability, (KP3) achieves this effect by using a single quantifier over proofs. Accepting a principle of this nature would generally be regarded as coming with a commitment to regarding proofs as objects. Such a commitment also appears to be

---

[14] In order to appreciate the significance which Martin-Löf assigns to these identities, it is important to keep in mind that he assigns epistemic significance only to what he refers to as judgements and not to proofs or propositions themselves. Martin-Löf takes a judgment to be an entity which is simultaneously an epistemic act and an object of such an act—a view which finds expression in statements such as "the truth of a judgement is simply defined as [its] knowability" (Martin-Löf 1998). So, if A denotes a proposition, and we consider the judgement expressed by "A is true," we reach Martin-Löf's preferred reformulation of Dummett's (K):

> (K$'$) If a judgment of the form "A is true" is correct, then the proposition A can be known to be true.

If we now think of (K$'$) relative to the foregoing equations, its conceptual relationship to (K) is apparent. For since the truth of A can be equated with the existence of a proof with conclusion A, (K$'$) licenses the inference from the correctness of the judgment of "A is true" to knowledge of the correctness of the judgment "There exists a proof of A."

implicit in (NonOm2) which, in addition to quantifying over proofs, also presumes that it is legitimate to predicate of them the property of having been constructed. Although (KP4) and (NomOm3) represent natural formalizations of these statements in a language with proof quantifiers, adopting these statements as expressions of (KP3) and (NonOm2) only intensifies the need for adopting an objectual interpretation of proof quantification.

Given that we have already seen that the roots of neo-verificationism lie in intuitionistic mathematics, we must also confront the following questions in the course of justifying the adoption of (KP3) and (NonOm2):

(1) Is it consistent with the general tenets of intuitionistic logic and mathematics to regard proofs as objects in a manner which legitimates the use of proof quantifiers?
(2) Is it similarly admissible to distinguish between those proofs which have actually been constructed and those which have not been constructed and thus exist "merely potentially"?

These questions touch on some of the most difficult and controversial issues concerning the interpretation of intuitionistic logic and mathematics. As such, we can hardly hope to provide satisfactory general answers here. What we will now attempt to do, however, is to briefly describe how the views of the neo-verificationists on these matters bear on the legitimacy of adopting (KP3) and (NonOm2) in order to reason about the Knowability Paradox.

In order to address the first question, it is necessary to briefly reprise several historical aspects of intuitionism. Brouwer, it will be recalled, considered mathematics to be a "languageless activity of the mind." This activity consists in the performance of mental constructions, which can constitute both mathematical objects and also proofs that the objects given by certain other constructions have various properties. This sort of mental activity is taken to precede formal logic, whose laws Brouwer held are isolated only by subsequent reflection on the general form of mental constructions. It was Heyting (1974) who first suggested that the principles of intuitionistically valid reasoning could be axiomatized, leading to such systems as intuitionistic predicate calculus [IQC] and Heyting arithmetic [HA]. Although he suggested no precise semantics for these calculi, he was the first to propose the sort of proof-based explanation of the connectives mentioned above.

Heyting's informal explanation eventually evolved into the BHK interpretation discussed in Sect. 1. It is unclear, however, whether Heyting himself took his description of a proof-based interpretation to constitute a genuine analysis of the meanings of the connectives as opposed to merely an informal explication. In addition to this, there is also some reason to suspect that Heyting would not have been entirely at home with treating proofs as objects. For, like Brouwer, Heyting held that proofs are constructions. And it is unclear whether he took a construction to be more like an object, or more like an act or process which results in a mental entity.[15]

[15] For an extended discussion of this issue see Sundholm (1983). Such a distinction between constructions as acts and constructions as objects influenced the later work of Scott (1970). However, Sundholm also notes that the object view has dominated over the act view in subsequent work within intuitionistic mathematics.

Modern expositions of the BHK interpretation often implicitly treat proofs as objects at least to the extent that they introduce schematic terms to denote proofs and operations on them. Beyond this, however, the "proofs as objects" view is explicitly adopted within several of the systems which attempt to formalize Heyting's interpretation, perhaps most notably Kreisel's *theory of constructions* (1962). This system corresponds to a formal calculus in which propositions, proofs and operations on proofs (i.e. constructions) are each assigned a syntactic representation. And although Kreisel claims this system to be "logic- and type-free", not only the language of this theory but also its subsequent use by other theorists are all suggestive of the fact that it is indeed committed to regarding proofs as abstract, structured objects.[16]

This perspective also seems to be paradigmatic of the "propositions-as-types" (or "formulas-as-types") approach to intuitionism to which both Prawitz and Martin-Löf have contributed significantly. According to Martin-Löf (1998), a type is prescribed by what must be done in order to construct an object of that type. For example, the type $\mathbb{N} \rightarrow \mathbb{N}$ is constituted by a general specification of what it is to be a function from natural numbers to natural numbers. Propositions are accordingly taken to comprise a special type $Prop$ which is constituted by means of verifying (i.e. proving) that a statement is true. In this context it is also standard to think of a given proposition as determining a type in accordance with its logical structure. This idea can be made precise by first identifying what we would conventionally describe as the proposition expressed by a sentence φ as the class of (canonical) natural deduction proofs with conclusion φ. If we then apply the Curry–Howard isomorphism, it becomes natural to think of a proposition as a type itself whose members are its (canonical) proofs. On the basis of this sequence of analogies, it has become conventional within the propositions-as-types tradition to speak of proofs as objects (qua members of the propositions which they verify).[17]

Dummett's views on the objectual status of proofs are less clear. This is in part due to the fact that his discussions often hew more closely to the historical consideration of intuitionism than do those of Prawitz and Martin-Löf. However, many of the pronouncements he has made on other issues appear to be motivated by a fear that regarding proofs as abstract entities will entail that truth is bivalent, an outcome which he has repeatedly argued would signify a collapse into realism. One of our goals below will be to show that such a fear is ill-founded. However, we must first attempt to understand the basis of his worries. And in order to do this, it will be useful to turn

[16] Such a conclusion is, for instance, already suggested by Kreisel's use of a formal predicate $\Pi(c, A)$ to which he assigned the interpretation "c is a proof of A." Subsequent work on the theory of constructions by Goodman (1970) led to the recognition that in order to ensure the consistency of Kreisel's theory, it is necessary to regard constructions (which are identified with proofs in the theory) as inhabiting stratified domains, the lowest of which consists of the rules of the original system and operations on them while higher levels are formed by reflecting on lower ones, conceived as completed domains. This picture is again suggestive of the "proofs as objects" view. Sundholm (1983) arrives at the same conclusion on the basis of more general considerations about the role which constructions play within intuitionism.

[17] The foregoing argument represents our attempt to summarize considerations which have evolved significantly since the seminal works in which Prawitz and Martin-Löf first attempted to make a connection between proof theory and a theory of meaning (e.g., Prawitz 1980a; Martin-Löf 1984). Independently of our recounting, however, it is clear that both Prawitz and Martin-Löf now frequently speak of proofs as objects—cf., e.g., Martin-Löf (1991) and Prawitz (1998a).

to the second question posed above—i.e. does it make sense to distinguish actual from potential proofs?

Prawitz and Martin-Löf have discussed the need for drawing such a distinction on a variety of occasions. Both theorists generally invoke the notion of potential proof in the context of arguing that mathematical truth is tenseless—i.e. that mathematical statements have truth values independently of our having proved them and thus do not "become" true only when then have been derived. Prawitz has put this point as follows:

> Intuitionistic philosophers sometimes use true as synonymous with truth as known, but this is clearly a strange and unfortunate use. We need a notion of truth where, without falling into absurdities, we may say, e.g., that there are many truths that are not known today. (Prawitz 1980b, p. 8)

The necessity for distinguishing between actual and potential proofs comes from the need to reconcile this view with the position that the truth of a statement implies its provability. Prawitz describes how this can be accomplished as follows:

> [I]t is here convenient to distinguish actual and potential existence . . . We can then say that the correctness of an assertion requires the actual existence of a proof, while the truth of the asserted proposition requires only (and is identical with) the potential existence of a proof of the proposition.[18](Prawitz 1998d, p. 48)

The picture which emerges from these passages is one according to which proofs are conceived as being brought into actual existence by the activity of mathematical agents, while at the same time acknowledging that such entities exist independently of their activities (i.e. in some "potential" sense). This picture leaves open the exact manner in which the distinction between actuality and potentiality should be understood. The problem comes out most clearly when we consider statements of the form "a proof of φ has not actually been constructed" which are relevant to the interpretation we wish to assign to (NonOm). For note that such a statement appears to be contrastable both with "a proof of φ *might* be constructed" and also "a proof of φ *will* be constructed."[19] We will propose a tentative means of making sense of the neo-verificationists' use of these terms in Sect. 4.1 below.

One is also naturally led to ask whether the picture just described forces us to adopt the view that there is a fixed class encompassing all potential proofs which must be thought of as a "completed totality"—i.e. a determinate (and presumably) infinite class

[18] Several similar passages may be found in Martin-Löf—e.g.:

> That proposition A is *actually true* means that A has been proved, that is that a proof of A has been actually constructed, which we can also express by saying that A is known to be true. . . . [T]o say that A is *potentially true* is to say that A can be proved, that is that a proof of A can be constructed, which is the same as to say, in usual terminology, simply that A is true. (Martin-Löf 1991, p. 141)

[19] We may additionally note in this regard that Prawitz (1987) often uses expression like "tenseless" to qualify what he intends by the notion of "potential existence." It thus follows that if actual existence contrasts with potential existence, then to say that a proof exists not just potentially but actually as well, is presumably to say that it exists in time.

of (presumably) abstract objects. Such a conclusion might appear counter to the spirit of verificationism because of its apparent resemblance to the platonistic view that, e.g., the natural numbers form a set of abstract objects independently of our having constructed (representations of) them. If we additionally adopt the view that the falsity of a statement is equivalent to the non-existence of a proof, then it appears to follow from this view that all statements are determinately either true or false. Since such a conclusion could then be used to justify the law of the excluded middle for arbitrary statements, it will presumably be unacceptable to the neo-verificationists for, as we have seen, they are all proponents of intuitionistic logic.

Dummett has voiced both of these concerns. On the one hand:

> We can introduce such a notion only by appeal to some platonistic conception of proofs as existing independently of our knowledge, that is, as abstract objects not brought into being by our thought. But if we admit such a conception of proofs, we can have no objection to a parallel conception of mathematical objects such as natural numbers, real numbers, metric spaces, etc. (Dummett 1982, p. 258)

And on the other:

> There is a well-known difficulty about thinking of mathematical proofs—and, equally, of verifications of empirical statements—as existing independently of our hitting on them, which insisting that there are proofs we are capable of grasping or of giving fails to resolve. Namely, it is hard to see how the equation of the falsity of a statement (the truth of its negation) with the non-existence of a proof or verification can be resisted: but, then, it is equally hard to see how, on this conception of the existence of proofs, we can resist supposing that a proof of a given statement determinately either exists or fails to exist. We shall then have driven ourselves into a realist position, with a justification of bivalence. (Dummett 1987, p. 285)

Dummett is able to reject the notion of potential proof because, at least at times, he has considered adopting the view that mathematical statements are significantly tensed (e.g. Dummett 1978). Such a view is not only in direct conflict with those of Prawitz and Martin-Löf, but it has a variety of other controversial consequences.[20] With this said, however, our ultimate contention will be that he is mistaken in deriving both of the conclusions expressed in the foregoing passages—i.e. that accepting the notion of potential proof forces us to accept a platonistic view of mathematical objects or that truth is bivalent. However, we will leave these matters for Sect. 4.2 and focus now on an independent argument which we believe shows that the notion of potential proof is in good standing intuitionistically.

Our argument is grounded in the notion of decidability. Within intuitionism, a statement ψ is traditionally defined to be decidable if $\vdash_I \psi \vee \neg\psi$ where $\vdash_I$ denotes provability in an appropriate intuitionistic formal system (e.g. IQC or HA). This notion can be extended to an arbitrary predicate $P(x)$ over a domain D which is also decidable (i.e. one for which $\vdash_I \forall x(x \in D \vee x \notin D)$. In particular, we say that $P(x)$

[20] E.g. that prior to Wiles' proof in 1995, the Fermat Conjecture was not true.

is decidable just in case for all $a \in D, \vdash_I P(\overline{a}) \vee \neg P(\overline{a})$ where $\overline{a}$ is a name for a. The paradigmatic example of a decidable domain is the natural numbers. Paradigmatic examples of decidable predicates include *x is greater than 17*, $Even(x)$ and $Prime(x)$. For when the definition of each of these predicates is written out as an open sentence $P(x)$ in the language of HA, it may be shown that $\mathsf{HA} \vdash P(\overline{n}) \vee \neg P(\overline{n})$ for each natural number n.

Suppose we fix the example $P(x) = Prime(x)$. We then have

(8) For every natural number n, $\mathsf{HA} \vdash Prime(\overline{n}) \vee \neg Prime(\overline{n})$.

Since HA satisfies the disjunction property, we additionally have

(9) For every natural number n, $\mathsf{HA} \vdash Prime(\overline{n})$ or $\mathsf{HA} \vdash \neg Prime(\overline{n})$.

Derivations within HA are presumably constructive proofs, *par excellence*. And thus when formulated in the informal language of (KP3), (8) implies

(10) For every natural number n, there is a proof of $Prime(\overline{n})$ or there is a proof of $\neg Prime(\overline{n})$.[21]

Discussions of decidable predicates occur throughout standard expositions of intuitionism (notably including Dummett 1977). And hence there seems to be no reason to think that either (8) or (9) would be regarded as contentious by even the most orthodox intuitionists.

Note, however, that the following statement also seems uncontentious

(11) There exists n such that we do not know either $Prime(\overline{n})$ or $\neg Prime(\overline{n})$.

The existential quantifier in (11) can be read constructively in the sense that we can, at any given time, find an explicit value for n which makes the matrix of this statement true—e.g. as of June 2008, we can take $n = 2^{41,854,567} - 1$. Note, however, that it follows from this that we cannot have constructed a proof of either $Prime(\overline{n})$ or $\neg Prime(\overline{n})$. For if we had, then surely we would know one statement or the other. And thus it appears to follow from (10) that even the most orthodox intuitionists must acknowledge the existence of non-constructed or "merely potential" proofs.

The foregoing considerations suggest that the distinction between actual and potential proofs is not only intuitionistically legitimate but necessary in order to account for obvious facts about our epistemic situation within mathematics. We may additionally note, however, that (9) also appears to commit us to the existence of infinitely many potential proofs. Since at any given time, the sum of our mathematical activity can presumably have led to the construction of only a finite number of proofs, this suggests that we must accept the existence of proofs which have not been constructed. Note, however, that (9) only commits us to the existence of infinitely many proofs in the sense that it ensures that for every natural number there is a proof of $Prime(\overline{n})$ or $\neg Prime(\overline{n})$. There has generally been consensus among intuitionists that the natural numbers can legitimately be treated as a *potential infinity*—i.e. a class of entities given by a construction which can be indefinitely repeated. What follows from (10)

[21] In fact, (10) can be proven internally in HA—i.e. $\mathsf{HA} \vdash Bew_{\mathsf{HA}}(\ulcorner \forall x[Bew_{\mathsf{HA}}(\ulcorner Prime(\dot{x})\urcorner) \vee Bew_{\mathsf{HA}}(\ulcorner \neg Prime(\dot{x})\urcorner)]\urcorner)$. Such a principle thus appears to be an undeniable consequence of principles universally accepted within intuitionism.

is that for each n, there exists a uniform method of constructing a proof $t(n)$ whose conclusion is either $Prime(\overline{n})$ or $\neg Prime(\overline{n})$. But this appears only to commit us to regarding potential proofs as corresponding to a potential infinity.

We will take the foregoing considerations to provide intuitionistic bona fides not only for distinguishing between actual and potential proofs but also for acknowledging that the latter category must be inhabited—i.e. that there exist potential proofs which have not been constructed. On this basis, the example of pairs of statements like $Prime(\overline{n})$, $\neg Prime(\overline{n})$ may additionally be used to make a related point which is of direct relevance to the Knowability Paradox. For as we have argued, we know (and in fact can prove—cf., footnote 21) that one or the other of these statements must be provable, despite the fact that neither has been proven. As such, it follows that there is no reason why the neo-verificationists should resist the conclusion that one of these statements is *true* (although of course we do not currently know which). If we now take χ to be whichever one of $Prime(\overline{n})$, $\neg Prime(\overline{n})$ is provable (and hence true), it follows that neo-verificationists should have no difficulty in accepting the conjunction

(12) χ and no proof of χ has been constructed.[22]

But this, of course, is just an instance of (NonOm2).[23] As we observed in Sect. 1, one might naturally think that there are manifest reasons to be suspicious of such a statement from the standpoint of intuitionism. But it now appears that the intuitionistic admissibility of statements like (12) is a straightforward consequence of the traditional understanding of decidable predicates coupled with the observation that we can only have constructed a finite number of proofs at any given time.[24]

[22] It might be objected here that since no proof of either $Prime(\overline{n})$ or $\neg Prime(\overline{n})$ has been constructed, an intuitionist will not be able to assert a concrete instance of (NonOm3) but only the intuitionistically weaker statement

(NonOm2d) For some ψ, $\psi \vee \neg\psi$ and no proof of either ψ or $\neg\psi$ has been constructed.

We have presented a case for the intuitionistic legitimacy of (NonOm2) itself mainly to preserve parity with the original exposition of the Paradox. However the current observation might be taken to weaken the claim that intuitionists are committed to accepting (NonOm1) in the original derivation of the Paradox and thereby also the conclusion (Om). But it may be shown in a manner similar to that employed by Percival (1990) that a version of the Paradox survives even if we weaken (NonOm1) to

(NonOm1d) For some ψ, $(\psi \vee \neg\psi) \wedge (\neg K\psi \wedge \neg K\neg\psi)$.

This observation should be compared with that given in footnote 29 below which formalizes the fact that the existence of non-constructed proofs follows from (KP4) by intuitionistically valid reasoning even if we assume (NonOm2d) instead of (NonOm2).

[23] Although Dummett has at times *rejected* the principle (NonOm), it is also notable that there are several more recent occasions where he seems to accept (NonOm2)—e.g. "[A] statement that a certain large number is prime is decidable, and may, when we apply the decision procedure, turn out to be true ... Hence, if it would turn out that the number is prime, the statement that it is prime is, on the definition I gave, true, even though we have at present no proof that it is, and may never have one, though we possess what is in fact an effective means of constructing one." (Dummett 1998, p. 123)

[24] We also take the foregoing considerations to constitute an argument which the neo-verificationists might use in order to reject the necessity of accepting what we called the bullet-biting resolution of the Knowability Paradox in footnote 8. On our interpretation, the occurrence of $K\varphi$ in (Om) ought to be rendered as "we have constructed a proof of φ." But the situation just described appears to be one in which a verificationist will acknowledge that there are provable (and hence true) statements whose proofs have not been constructed.

So goes our case for the use of (KP3) and (NonOm2) to schematize the views of the neo-verificationists within natural language. What we have left to do is to motivate the transition from the informal statement of these principles to their regimentation in the language of QLPE as (KP4) and (NonOm3). The legitimacy of this step will depend not just on the intended interpretation of its language sketched in Sect. 1, but also on how faithfully the axioms and rules of this system reflect this interpretation. We motivate the axioms of this in Sect. 3 on the basis of their relationship to the Gödel provability interpretation discussed above. In Sect. 4.1, we will discuss a potential precedent for our inclusion of the predicate $E(x)$ in the language of QLPE from within traditional intuitionistic mathematics. And in Sect. 4.2, we will discuss what turns on our decision to base this system on classical as opposed to intuitionistic logic.

3 Reconstruction and consistency

In the previous section we argued that (KP3) and (NonOm2) represent plausible reconstructions of how the neo-verificationists are likely to understand the epistemic principles on which the Knowability Paradox is premised. It remains to be shown that these principles do not entail any consequences which can reasonably be taken to imperil neo-verificationism. We will pursue this goal here in four stages: (1) we will formally present the language of QLPE in which our proposed formalizations (i.e. (KP4) and (NonOm3)) are stated; (2) we will propose an axiom system for reasoning about statements in this language; (3) we will argue that no statement analogous to (Om) can be derived in QLPE from (KP4) and (NonOm3) in a manner which mimics Fitch's original derivation; and (4) we will then demonstrate that these principles do not conflict in any other way by presenting a semantics for QLPE relative to which they can be shown to be formally consistent.

3.1 QLPE and the reconstruction of the Fitch derivation

As we mentioned in Sect. 1, QLPE is a form of explicit modal logic which descends from Fitting's (2004, 2006) system QLP, which itself descends from Artemov's (2001) system LP. In the interest of concision, we will present the language $\mathcal{L}_E$ of QLPE directly.

Definition 3.1 The class Term_E of *proof terms* $\mathcal{L}_E$ is specified by the grammar

$$t := x_i \mid a_i \mid !t \mid t_1 \cdot t_2 \mid t_1 + t_2 \mid (t\forall x)$$

$x_1, x_2, \ldots$ are known as *proof variables*, $a_1, a_2, \ldots$ as *axiom constants* and $!, \cdot, +$ and $(t\forall x)$ as *proof operations* respectively called *proof checker* (unary), *application* (binary), *sum* (binary) and *uniform verifier* (binary).

Footnote 24 continued
And from this it follows that if we analyze knowledge via (2a), there appears to be no reason why we should not accept the existence of true but unknown statements.

Definition 3.2 For $t \in \text{Term}_E$, the class Form_E of formulas of $\mathcal{L}_E$ is specified by the grammar

$$\varphi := P_i \perp \mid E(t) \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2 \mid \neg\varphi_1 \mid t : \varphi_1 \mid (\forall x)\varphi_1 \mid (\exists x)\varphi_1$$

The free variables of $\varphi \in \text{Form}_E$ are defined in the standard manner of first-order logic. The class Sent_E of sentences of $\mathcal{L}_E$ are accordingly defined as those formulas which contain no free variables.

We will see in a moment that there is a strong sense in which proof terms can be taken to directly reflect the structure of derivations in a Hilbert-style system for propositional modal logic. It is thus reasonable to think of these terms as both denoting formal proofs and also directly displaying their structure. However, the semantics we will propose for QLPE allows for somewhat more flexibility in interpreting proof terms. And thus for purposes of this section we will adopt the convention of speaking of the denotation of a proof term as a *verification* rather than a proof.[25]

We now introduce a Hilbert-style axiom system for $\mathcal{L}_E$ in two stages, first introducing the axioms and then the rules.

Definition 3.3 The axioms of QLPE are as follows:

LP1 all tautologies of classical propositional logic
LP2 $t : (\varphi \rightarrow \psi) \rightarrow (s : \varphi \rightarrow t \cdot s : \psi)$
LP3 $t : \varphi \rightarrow \varphi$
LP4 $t : \varphi \rightarrow !t : t : \varphi$
LP5 $t : \varphi \rightarrow t + s : \varphi$ and $s : \varphi \rightarrow t + s : \varphi$
QLP1 $(\forall x)\varphi(x) \rightarrow \varphi(t)$, for any proof term t that is free for x in $\varphi(x)$.
QLP2 $\varphi(t) \rightarrow (\exists x)\varphi(x)$, for any proof term t that is free for x in $\varphi(x)$.
QLP3 $(\forall x)(\psi \rightarrow \varphi(x)) \rightarrow (\psi \rightarrow (\forall x)\varphi(x))$, where x does not occur free in ψ.
QLP4 $(\forall x)(\varphi(x) \rightarrow \psi) \rightarrow ((\exists x)\varphi(x) \rightarrow \psi)$, where x does not occur free in ψ.
UBF $(\forall x)t : \varphi(x) \rightarrow (t\forall x) : (\forall x)\varphi(x)$

Before specifying the rules of QLPE, we must first introduce several auxiliary definitions. A *constant specification* is a mapping $\mathcal{C}$ from proof constants to sets of formulas (possibly empty). Informally speaking, a constant specification is a way of

[25] Having now defined the formal language of QLPE, we wish to draw two quick contrasts between this language and that in employed in Martin-Löf's Intuitionistic Type Theory [ITT] (Martin-Löf 1984). The first concerns the presence of the ":" symbol in both systems. As we have seen, in explicit modal logic this symbol is used to express the relation between proofs and formulas. However, in ITT, : is used to express typing judgements. This means that although these systems contain similar looking formulas—e.g. $t \cdot s : \varphi$ versus $t(s) : \varphi$—such formulas are assigned very different interpretations. One characteristic ramification of the different uses of : is that in the context of LP and related systems which derive from modal logic, it is allowable to iterate the application of : leading to formulas such as $!t : t : \varphi$ (which expresses that $!t$ is a proof of $t : \varphi$) whereas no such iteration is allowed in ITT. This contrast also seems to contribute to the fact that quantification over proofs cannot be expressed in the language of ITT. For in this system, the formal existential quantifier is defined in terms of a dependent typing judgement whereas the concept of existence is analyzed in terms of the habitation of a type.

assigning proof constants as names for unstructured verifications for particular statements. We say that a formula φ *has a proof constant* with respect to the specification $\mathcal{C}$ if $\varphi \in \mathcal{C}(c)$ for some proof constant c. A constant specification $\mathcal{C}$ meets the *free variable condition* if whenever $\varphi(x_1, \ldots, x_n) \in \mathcal{C}(c)$ and $y_1, \ldots, y_n$ are variables which do not occur in $\varphi(x_1, \ldots, x_n)$, then $\varphi(y_1, \ldots, y_n) \in \mathcal{C}(c)$. $\mathcal{C}$ is *axiomatically appropriate* if for all (and only) instances φ of the axioms given in Definition 3, there exists a proof constant c such that $\varphi \in \mathcal{C}(c)$. We henceforth assume that we are working with a fixed axiomatically appropriate constant specification $\mathcal{C}$ meeting the free variable condition.

Definition 3.4 The rules of QLPE may now be given as follows:

RLP1 modus ponens
RLP2 If φ is an axiom and $\varphi \in \mathcal{C}(c)$, then $\vdash c : \varphi$.
RQLP1 If $\vdash \varphi(x)$, then $\vdash (\forall x)\varphi(x)$.

Since the system just described may seem obscure to readers unfamiliar with explicit modal logic, several observations are in order. As mentioned in Sect. 1, logics in the family of LP may be viewed as explicit analogs of S4 in a sense which is made precise in Artemov (2001) (see also footnote 27 below). Axioms LP2, LP3 and LP4 are analogous to the S4 axioms K (normality), T (reflection) and 4 (transitivity). In the context of S4, all of these axioms appear to express valid principles when the operator is viewed as expressing informal provability in the sense of the Gödel interpretation (4). For instance, relative to this interpretation, K expresses that if an implication is provable, then if its antecedent is provable, its consequent is provable as well. Axiom LP2 expresses an analogous fact about the relation between proofs and statements expressed by $t : \varphi$—i.e. if t denotes a proof of an implication and s denotes a proof of its antecedent, then the result of applying t to s denotes a proof of its consequent. Similar justifications can be given for axioms LP3 and LP4.

The quantifier axioms QLP1-QLP4 correspond to a standard set of axioms for a Hilbert-style axiomatization of first order logic. These statements can also be seen to express intuitively valid principles about the relationship between proofs and statements. For instance, QLP2 is justified on the basis of the fact that if some explicit proof term t denotes a verification of φ, then there must exist a verification of φ. The axiom UBF (which is short for Uniform Barcan Formula) expresses the fact that if for all x, t uniformly serves to justify $\varphi(x)$, then there is a proof denoted by the uniform verifier term $(t\forall x)$ which serves to justify $\varphi(x)$. This too is arguably a valid principle about the concept of informal provability. However, the role of UBF in our current exposition will be primarily instrumental in the sense that its inclusion is only required if we wish QLPE to satisfy the Constructive Necessitation Theorem stated as Proposition 3.2 below.[26]

It should also be noted that none of the QLPE axioms explicitly mention the predicate $E(x)$. This reflects our intention to interpret $E(t)$ as expressing that the verification denoted by t has been constructed. As we discussed in the previous section, the fact that a given verification has been constructed appears to be a contingent matter reflecting not the properties of proofs themselves, but merely the exigencies of our deductive

[26] For more on the status of this principle, see Dean and Kurokawa (2008a).

practice. From this it follows that it would be inappropriate to adopt as an axiom a statement such as

EApp $(\forall x)(\forall y)[(E(x) \wedge E(y)) \rightarrow E(x \cdot y)]$

For if we were to adjoin EApp to QLPE, it would then follow that the extension of E was closed under application. But note that it seems possible that we might have constructed a verification, denoted by t, of $\varphi \rightarrow \psi$ and a verification, denoted by s, of φ, but failed to have constructed the verification denoted by $t \cdot s$ which (per axiom LP2) denotes a proof ψ.

Note also that the rule RLP2 provides for a significant degree of flexibility about how proof constants are treated as naming verifications of axioms. This rule allows, for instance, the case in which $\mathcal{C}$ assigns a distinct constant to each instance of every axiom scheme. Accordingly, we also do not adopt as axioms instances of the schema

EAx $E(a)$ for a an axiom constant

To do otherwise would entail that the extension of E would be infinite were we to adopt such a constant specification. We will show below, however, that our proposed resolution to the Paradox does not depend on adopting such a "minimalist" attitude about which verifications have been constructed.

With the axioms and rules of QLPE in place, we can define the derivability relation $\Gamma \vdash_{\mathsf{QLPE}} \varphi$ in the normal manner. On this basis, we can now state two basic results about QLPE whose proofs carry over directly from QLP:

Proposition 3.1 (Deduction Theorem) *If* $\Gamma, \varphi \vdash_{\mathsf{QLPE}} \psi$, *then* $\Gamma \vdash_{\mathsf{QLPE}} \varphi \rightarrow \psi$.

Proposition 3.2 (Constructive Necessitation Theorem) *If* $\vdash_{\mathsf{QLPE}} \varphi$, *then* $\vdash_{\mathsf{QLPE}} t : \varphi$ *for some proof term t not containing free variables.*

The Constructive Necessitation Theorem expresses that QLPE is capable of internalizing its own proofs in the sense that if φ is derivable in this system, a term t which encodes the structure of its derivation can be constructed so that the formula $t : \varphi$ is itself provable. This result shows that the axiom necessitation rule RQLP2 extends to arbitrary provable formulas, meaning that QLPE satisfies an explicit form of the traditional necessitation rule of S4. Results of this kind play a significant role in the original motivation of LP and QLP.[27]

As we argued in the previous section, there should be no obstacle to the neo-verificationists accepting our proposed $\mathcal{L}_E$ formalizations of (KP3) as (KP4) and (NonOm2) as (NonOm3). We reproduce these principles here for reference:

(KP4) For all $\varphi \in \text{Sent}_E$, $\varphi \rightarrow (\exists x)x : \varphi$.
(NonOm3) For some $\psi \in \text{Sent}_E$, $\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]$.

[27] In particular, Proposition 3.2 is used in order to prove the so-called Realization Theorem for LP given in Artemov (2001). A simplified statement of this theorem is as follows: If a modal formula φ is provable in S4, then a statement φ^r is provable in LP where $(\cdot)^r$ represents some mapping between the language of S4 and LP which replaces each occurrence of $\Box$ in φ with a proof term t in φ^r. This result has been shown to have a number of useful applications, among them giving an arithmetic provability interpretation for IPC.

We now wish to see what happens when we adjoin these statements to QLPE. To facilitate this, let $\vdash_{\mathsf{QLPE+KP4}}$ denote derivability in the system consisting of QLPE together with all instances of (KP4) and let NonOm3$_\chi$ denote the single statement formed by taking $\psi \equiv \chi$ in (NonOm3).[28]

Proposition 3.3 *For any* $\psi \in Sent_E$, *NonOm3*$_\psi \vdash_{\mathsf{QLPE}+KP4} (\exists x)\neg E(x)$.

Proof Writing $\vdash$ for $\vdash_{\mathsf{QLPE+KP4}}$ we may reason as follows:

(1)	NonOm3$_\psi \vdash \psi$	propositional logic
(2)	NonOm3$_\psi \vdash (\forall x)[E(x) \to \neg x : \psi]$	propositional logic
(3)	NonOm3$_\psi \vdash \psi \to (\exists x)x : \psi$	KP4
(4)	NonOm3$_\psi \vdash (\exists x)x : \psi$	(1), (3)
(5)	NonOm3$_\psi \vdash (\forall x)[E(x) \to \neg x : \psi] \to$	
	$((\exists x)x : \psi \to (\exists x)\neg E(x))$	QLP1, QLP2, QLP4
(6)	NonOm3$_\psi \vdash (\exists x)x : \psi \to (\exists x)\neg E(x)$	(2), (5)
(7)	NonOm3$_\psi \vdash (\exists x)\neg E(x)$	(4), (6)

□

This result demonstrates that when formalized in $\mathcal{L}_E$, the premises of the Knowability Paradox lead directly to the conclusion that there are non-constructed proofs. This is in line with the conclusion we reached in the previous section on the basis of our discussion of decidable predicates and has nothing directly to do with the reasoning of the Paradox.[29] But one is also naturally drawn to ask what happens when we attempt to reconstruct the original Fitch derivation on the basis of these premises. As we will now attempt to demonstrate, such an attempt does not lead to a reiteration of the Paradox. It does, however, provide another way of demonstrating how (NomOm3) entails the existence of non-constructed proofs.

In order to reconstruct the reasoning of the Fitch derivation, we will work in the system QLPE + KP4 and this time treat NonOm3$_\psi$ as a reductio assumption. In parallel to the original derivation, we endeavor to derive

(13) (Om3) For all $\varphi \in \mathcal{L}_E, \varphi \to (\exists x)[E(x) \wedge x : \varphi]$.

by taking NonOm3$_\psi$ as the substitution instance for φ in (KP4). This gives us

[28] For purposes of definiteness, we do not treat (KP4) as an axiom and hence RPL2 does not apply to it. Nothing in the sequel will depend on this.

[29] It may also be noted that $(\exists x)\neg E(x)$ is also derivable in QLPE + KP4 on the basis of somewhat weaker premises than those employed in Proposition 3.3. In particular, suppose on the basis of the discussion in footnote 22 about the assertablity of a particular instance of (NonOm3) that a neo-verificationist is only willing to accept

(NonOm3d) $(\psi \vee \neg\psi) \wedge (\forall x)[E(x) \to (\neg x : \psi \wedge \neg x : \neg\psi)]$

It follows by essentially the same reasoning employed in Proposition 3.3, plus an intuitionistically valid proof by cases, that for any ψ, NonOm3d$_\psi \vdash_{\mathsf{QLPE+KP4}} (\exists x)\neg E(x)$. We take this to demonstrate that a commitment to the existence of non-constructed proofs is already implicit in (KP4) even when taken in conjunction with very weak formulations of non-omniscience.

$(1')$ $\text{NonOm3}_\psi \vdash \psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]$
$(2')$ $\text{NonOm3}_\psi \vdash (\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]) \rightarrow$
$(\exists y)y : [\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]]$ KP4
$(3')$ $\text{NonOm3}_\psi \vdash (\exists y)y : [\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]]$ $(1'), (2')$

In the original derivation at line (4), the statement $K(\psi \wedge \neg K\psi)$ is assumed as another reductio premise. If we continue to use the analysis of knowledge given by (2b), then this statement should go over into $\mathcal{L}_E$ as

(14) $(\exists z)[E(z) \wedge z : (\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi])]$

In order to mimic the effect of assuming this statement as an open assumption in a natural deduction proof, it is most convenient to introduce a free-variable instance of this statement $E(z) \wedge z : (\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]) \equiv \Theta$ as an assumption. On this basis, we may continue as follows:

$(4')$	$\Theta \vdash z : [\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]]$	propositional logic
$(5')$	$\Theta \vdash E(z)$	propositional logic
$(6')$	$\Theta \vdash a_1 : ([\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]] \rightarrow \psi)$	RQLP2
$(7')$	$\Theta \vdash a_2 : ([\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]] \rightarrow$	
	$(\forall x)[E(x) \rightarrow \neg x : \psi]])$	RQLP2[30]
$(8')$	$\Theta \vdash a_1 \cdot z : \psi$	LP1 $(5'), (6')$
$(9')$	$\Theta \vdash a_2 \cdot z : (\forall x)[E(x) \rightarrow \neg x : \psi]$	LP1 $(5'), (7')$
$(10')$	$\Theta \vdash a_2 \cdot z : (\forall x)[E(x) \rightarrow \neg x : \psi] \rightarrow$	
	$(\forall x)[E(x) \rightarrow \neg x : \psi]$	LP3
$(11')$	$\Theta \vdash (\forall x)[E(x) \rightarrow \neg x : \psi]]$	$(9'), (10')$
$(12')$	$\Theta \vdash E(a_1 \cdot z) \rightarrow \neg a_1 \cdot z : \psi$	QLP $(11')$
$(13')$	$\Theta \vdash \neg E(a_1 \cdot z)$	propositional logic

Steps $(4')$–$(13')$ can be taken to parallel steps (4)–(9) in the original derivation—i.e. the hypothetical subproof with premise $K(\psi \wedge \neg K\psi)$. At the end of this subproof we were originally able to derive $\neg K(\psi \wedge \neg K\psi)$ based on the conflict between step (7) (i.e. $K\psi$) and step (9) (i.e. $\neg K\psi$). In our reconstructed derivation, however, we derive only $E(z)$ and $\neg E(a_1 \cdot z)$. These statements respectively express that the verification denoted by z has been constructed and that the verification denoted by $a_1 \cdot z$ has not. Note that this is not a formal contradiction relative to the axioms we have presented for QLPE. For pursuant to our discussion above, it is at least consistent to think that we may have constructed the verification denoted by z but have not performed the additional step of applying a_1 to it so as to construct the verification denoted by $a_1 \cdot z$.

One might take this to suggest that the original derivation of the Paradox breaks down in QLPE only because we have not elected to impose closure axioms on the

[30] Steps $(5')$ and $(6')$ respectively follow by applying RLP2 to the Hilbert axioms $(P \wedge Q) \rightarrow P$ and $(P \wedge Q) \rightarrow Q$ with $P \equiv \psi$ and $Q \equiv (\forall x)[E(x) \rightarrow \neg x : \psi]$. Here we take a_1, a_2 such that $(P \wedge Q) \rightarrow P \in \mathcal{C}(a_1)$ and $(P \wedge Q) \rightarrow Q \in \mathcal{C}(a_2)$.

extension of $E(x)$ such as EApp and EAx. Once we decide to represent knowledge in terms of proof existence in the sense of (2a), it can readily be seen that closure conditions of this sort correspond to the traditional assumption that knowledge is closed under deductive consequence. As we noted above, if we adopt such conditions, then we are immediately pushed down a slope which eventuates in the conclusion that the extension of $E(x)$ must be infinite. This is out of keeping with our intention to interpret $E(x)$ as holding only of proof terms which denote verifications which have been constructed.[31]

But even taking this into account, the degree of closure which is required to bring steps (5′) and (13′) into conflict seems minor indeed. In particular, it may seem as if the foregoing derivation allows us to "block" an analog to the Knowability Paradox only by refusing to acknowledge that knowledge of a conjunction might not entail knowledge of its conjuncts. However, we believe this not to be the case. For as we will now attempt to make clear, (Om3) is not derivable from (NonOm3) and (KP4), even if we assume the closure conditions needed to derive a contradiction from (5′) and (13′). For suppose we are now reasoning in the system QLPE + KP4 + EAx + EApp. Then we may continue as follows:

(15′)	$\Theta \vdash E(a_1)$	EAx
(16′)	$\Theta \vdash (E(a_1) \wedge E(z)) \rightarrow E(a_1 \cdot z)$	RQLP1, EApp
(17′)	$\Theta \vdash E(a_1 \cdot z)$	(5′), (15′), (16′)
(18′)	$\Theta \vdash \perp$	(13′), (17′)
(19′)	$\vdash \neg(E(z) \wedge z : (\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]))$	(18′), deduction theorem
(20′)	$\vdash \neg(\exists z)(E(z) \wedge z : (\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]))$	(19′), RQLP1, QLP4

The derivation (4′)–(20′) can be taken to structurally parallel steps (4)–(10) in the Fitch derivation. Step (10) in the original derivation can be taken to express that it cannot be known that $\psi \wedge \neg K\psi$. But since this conclusion is provable from no premises,

[31] It is at this point where our approach to the Knowability Paradox comes into closest contact with the traditional problem of logical omniscience—cf., e.g., Parikh (1987) and Fagin et al. (1995). For present purposes, a logic including a propositional knowledge operator K (such as that employed in the Fitch derivation) can be taken to be "logically omniscient" just in case the class of formulas φ of which $K\varphi$ can be proven to hold is deductively closed. This is often described as a negative feature of such systems, as it collapses the distinction between those statements which we normally describe as *known* (in the sense of being explicitly present in our ken) and those we would describe as being merely *knowable* (in the sense of being potentially unrealized consequences of statements we know explicitly). Note, however, that this distinction between knowledge and knowability does *not* coincide with the one which is relevant to the Knowability Paradox. As we have seen, in this setting "φ is knowable" has traditionally been rendered not as $K\varphi$ but rather as $\Diamond K\varphi$. As the foregoing discussion of closure conditions brings out, our means of analyzing knowledge and knowability via (2a, b) is more closely connected with the traditional distinction than it is with the analysis of knowability in terms of possibility. We note in passing that there is at least a conceptual affinity between our use of E for formalizing reasoning about constructed proofs and the explicit knowledge modalities employed by Fagin and Halpern (1988) to blunt the problem of logical omniscience. However, we take Corollary 3.1 below to show that our resolution to the Paradox itself does not depend on any particular means of analyzing (or resolving) the logical omniscience problem relative to QLPE.

it follows by the modal necessitation rule that (11) $\Box\neg K(\psi \wedge \neg K\psi)$, which in turn can be taken to express that $\psi \wedge \neg K\psi$ is an *unknowable* truth. A parallel claim in the current setting would be that $\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]$ is an *unprovable* truth, contra (3′). But note that (20′) expresses merely that no proof of this statement can fall under the extension of E. Relative to the interpretation of $\mathcal{L}_E$ introduced above, this entails not the claim that this statement is unprovable, but merely that no proof of this statement has already been constructed.

Although this conclusion does not stand in formal conflict with (3′), one might at first think there was still a sort of conceptual tension between (20′) and (3′). For while on the one hand (3′) reports that there is a proof z of $\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]$, the interpretation we have just provided for (20′) appears to suggest that this proof can never be constructed. One might thus wonder whether z is a sort of "ideal" proof which exists in principle but could never be actualized in practice. But recall that the fundamental motivation behind the neo-verificationist's adoption of a principle such as (KP) is to rule out the existence of unverifiable truths. As such, the conclusion that there are true statements which are verifiable only by non-constructible proofs might be taken to show that the formalization of (KP) by (KP4) does not adequately capture the intentions of these theorists in advocating an epistemic constraint on truth.

In order to see that this tension is merely apparent, some additional consideration must be paid to the role played by the predicate E within QLPE. As we have stated above, this predicate is intended to mark a distinction between actual and potential proofs which we have argued is both acknowledged in the writings of the neo-verificationists and required in virtue of standard assumptions about the status of decidable predicates within intuitionism. The question remains as to how this distinction can be made precise with respect to a description of our actual deductive activities. We will defer a detailed examination of this topic to Sect. 4. For now, however, it will suffice to note that it is at least consistent with the writing of the neo-verificationists to think of the quantifiers of QLPE as ranging over all verifications which are constructible in the sense of corresponding to the denotations of proof terms $t \in \text{Term}_E$. Note, however, that this class is inductively generated and hence infinite. Thus at any given stage in our deductive activity, we will only have constructed finitely many of its members. The predicate E is intended to hold of precisely those proof terms denoting verifications we have constructed at or prior to the current stage.[32]

[32] The relevant contrast between constructible and constructed proofs can also be characterized by likening the quantifier $(\exists x)$ to a *possibilist* quantifier over proofs—i.e. one which ranges not only over proofs which are actual (in the sense of having been constructed) but also over ones which are merely possible (in the sense of being constructible in principle). This helps to explain why $\mathcal{L}_E$ sentences of the form $(\exists x)x : \varphi$ can be taken to express the *knowability* of φ as opposed to merely the fact that a verification of this statement has been found. Note that if we were to adopt an *actualist* interpretation of proof quantification—i.e. one in which proof quantifiers ranged only proofs which had been constructed at some particular stage—we would be forced to introduce traditional modal operators to express statements about provability in principle. (KP3) and (NonOm2) would then be respectively formalized as (KP_m) $\varphi \rightarrow \Diamond(\exists x)x : \varphi$ and (NonOm_m) $\psi \wedge \neg(\exists x)x : \psi$. However, it can readily be verified that these principles lead to a contradiction via reasoning which parallels the Fitch derivation. As such, a possibilist interpretation of the QLPE quantifiers appears to be preferred both because it conforms better with the use of proof quantification in the writings of the neo-verificationists and also because it allows us to avoid becoming trapped in the original reasoning of the Paradox.

If we apply this interpretation to (20′), this statement no longer has the appearance of expressing that no proof of $\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]$ is constructible in principle. This sense could be expressed in QLPE by a statement with an unrestricted proof quantifier of the form

(15) $\neg(\exists y) y : (\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi])$

This statement expresses that there is no proof (and therefore no constructible proof) of $\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]$ and as such obviously stands in contradiction to (3′). However, (20′) differs from (15) in that it begins with a quantifier restricted by E. It consequently expresses not that there is no proof of $\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]$ (or that it is impossible to construct such a proof), but merely that we cannot currently have constructed a proof of a statement of this form.

The provability of (20′) in QLPE thus does not reflect that statements of the form $\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]$ possess some mysterious epistemic property. Rather, it illustrates the more mundane observation that were we to have constructed a proof of such a statement at the present stage in our deductive activity, we would thereby have refuted its second conjunct. For suppose that z verifies $\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]$ and also that $E(z)$. Then it follows both that $a_1 \cdot z$ verifies ψ and also that $(\forall x)[E(x) \rightarrow \neg x : \psi]$ must be true. But this latter statement expresses that no verification of ψ has been constructed, meaning in particular that $\neg E(a_1 \cdot z)$. If we assume the minimal closure conditions on constructed verifications discussed above, it follows that z itself cannot fall under E on pain of contradicting this conclusion.

Hand (2003, 2008) has called attention to essentially the same phenomenon in the context of the traditional presentation of the Knowability Paradox. He suggests that the situation just described can be taken to reflect a genuine limitation of our ability to construct proofs of statements which describe the current extent of our knowledge. In particular, we cannot, as a matter of practical fact, verify a conjunction like $\psi \wedge \neg K\psi$ one of whose constituents expresses that the other has not been verified. For in the course of so doing we would make this conjunct false. Hand maintains on behalf of the neo-verificationists that such statements might still be verifiable in the sense that there exists a constructive procedure by which they could be shown to be true. But for the operational reason just described, such procedures cannot be carried out in practice.[33]

In summary, the sort of resolution to the Paradox we are proposing relies on a related distinction between constructed and constructible proofs. For although (20′)

[33] In his earlier paper, Hand (2003) suggests that these considerations demonstrate that the fundamental semantic principle endorsed by the neo-verificationists must not be formulated in a manner which entails that every true statement is verifiable by an *executable* verification procedure. However, this might appear to entail that it is consistent to maintain (KP) only for statements which do not give rise to paradoxical conclusions in the manner of the so-called "restrictionist" strategy of Tennant (1997). In his more recent paper, Hand (2008) modifies his position somewhat by proposing that from the semantic perspective, it may be maintained that all true statements may be verified by constructive procedures. However, what he refers to as "pragmatic" factors intervene to prevent the procedure associated with $\psi \wedge \neg K\psi$ from being carried out. While we are largely in sympathy with Hand's deployment of a distinction of this sort, we contend that a better way to aid the neo-verificationists is to take their writings at face value and formalize (KP) using proof quantifiers rather than propositional modal operators. This allows us to formulate their views in a manner which avoids the necessity of either limiting the scope of (KP) or explaining away potentially problematic instances of this principle on non-semantic grounds.

reports that no verification of $\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]$) can be constructed now, this is not in formal conflict with (3′) which states that such a proof exists. What (20′) does entail, however, is that if NonOm3_{ψ} is true for any ψ, then (KP4) entails that there must exist a non-constructed proof. Although this is essentially the same result as Proposition 3.3, (4′)–(13′) show how the result can be reached in a manner which resembles the original reasoning of Fitch. This same result can be presented more generally by subsuming NonOm3_{ψ} into the antecedent of a conditional to yield

Proposition 3.4 *For any* $\psi \in Sent_E$,

$$\vdash_{\mathsf{QLPE}} (\forall z)(z : [\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]] \rightarrow \neg E(a_1 \cdot z)).$$

On the basis of the foregoing discussion, it should now be clear that there is no sense in which the verification promised by (3′) of $\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]$ must be viewed as a *non-constructible* proof. For note that not only is (20′) premised on our acceptance of NonOm3_{ψ}, there is nothing about (20′) or Proposition 3.4 which precludes that a verification of $\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]$ might come to be constructed at some later point in our deductive practice than that which E describes. Put in terms of the original distinction drawn by the neo-verificationists, this is to say that while the proof claimed to exist by (3′) must of necessity be merely potential, there is no in-principle reason why it cannot be made actual by further deductive effort.[34]

The foregoing considerations also appear to suggest that there is no straightforward way of continuing the above derivation in QLPE + (KP4) which eventuates in (Om3) via the reasoning of the Fitch derivation.[35] However, in order to demonstrate this formally, it is not sufficient to merely note that there is no contradiction between (3′) and (20′). Rather, we must somehow show that for at least certain choices of ψ, NonOm3_{ψ} is consistent with all instances of (KP4). We now turn to this task by presenting a semantics of QLPE. This will allow us to both prove such a consistency result and also to frame more precisely the distinction between actual and potential proofs which we have been discussing.

3.2 The consistency of (KP4) and (NonOm3)

The foregoing discussion suggests that the consistency of (KP4) and (NonOm3) rests on the admissibility of a situation in which an instance of (NomOm3) is provable,

[34] This means of resolving the issue does, of course, depend on the assumption that E does not vary in extension over time but always holds of just those verifications which have been constructed at or prior to the present moment. This might be achieved, for instance, by determining its extension indexically by referring to the current moment through the use of a word like ‘now’ in a manner which also allowed us to make subsequent reference to this time. We will return to question of whether such indexical devices are ultimately required in order to fix the extension of E in Sect. 4.1.

[35] Another reason for this is that QLPE has no precise analogue to the modal necessitation rule by which the step from (10) to (11) is mediated in the Fitch derivation. It is possible to mimic this step indirectly by applying the Constructive Necessitation Theorem 3.2 to produce a proof term t such that $\vdash t : \neg(\exists z)(E(z) \wedge z : (\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]))$ and then existentially generalizing to obtain $\vdash (\exists u)u : \neg(\exists z)(E(z) \wedge z : (\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]))$. If provability is likened to a kind of necessity, then this statement can be taken to express the necessity of the non-existence of a *constructed* proof of $\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi]$. However, this does not rule out the possibility that there might exist a non-constructed proof of this statement and therefore does not stand in conflict with (3′).

but its proof has not been constructed. Turning this observation into a formal proof of consistency which will ensure that no contradiction can be derived from (KP4) and (NonOm3) requires either a more sophisticated proof theoretic argument or a formal semantics which is at least sound for QLPE. On the basis of the historical roots of explicit modal logics in Provability Logic, one might initially think that such a semantics should take the form of an arithmetical interpretation for $\mathcal{L}_E$ analogous to that described for GL or LP above. However, there are serious technical obstacles to providing such a semantics for systems containing QLP.[36] But as the origin of these difficulties is largely unrelated to the sort of use to which we are attempting to put QLPE here, we will make use of a form of so-called *evidential semantics* whose roots can be traced to Mkrtychev's (1997) term-based semantics for LP and Fitting's later adaption of this idea to provide Kripke-like semantics for LP (Fitting 2005) and QLP (Fitting 2004, 2006). For the moment, we do not wish to endow the semantics which we are about to present with any particular intuitive significance with respect to the intended interpretation of $\mathcal{L}_E$. Rather, we wish only to show the formal consistency of the set

$$\Gamma_\psi = \{\psi \wedge \forall x\,[E(x) \rightarrow \neg x : \psi]\} \cup \{\varphi \rightarrow (\exists x)x : \varphi \mid \varphi \in \text{Sent}_E\}$$

consisting of an appropriately chosen instance (NonOm3_ψ) of non-omniscience together with all instances of (KP4) by constructing a model of QLPE in which Γ is satisfied. We now embark on a sequence of definitions in order to fulfill this goal.

A QLPE model $\mathcal{M}$ will consist of two domains $\mathcal{D}_1$ and $\mathcal{D}_2$ such that $\mathcal{D}_1 \subseteq \mathcal{D}_2$. The members of the domains will correspond to the denotations of proof terms—i.e. what we have been calling *verifications*. For heuristic purposes, $\mathcal{D}_2$ may be thought of as the class of all verifications which exist in a particular model and $\mathcal{D}_1$ as the class of verifications which have been constructed in that model. We start out by defining the notion of an interpretation $\mathcal{I}$ for $\mathcal{M}$:

Definition 3.5 A QLPE *interpretation* for $\mathcal{M}$ is a function $\mathcal{I}$ which interprets the proof operations $\cdot$, !, + and $(\cdot\forall\cdot)$ as follows:

(i) $\mathcal{I}$ assigns to each constant symbol a a member $a^{\mathcal{I}}$ of $\mathcal{D}_2$
(ii) $\mathcal{I}$ assigns to ! a mapping $!^{\mathcal{I}} : \mathcal{D}_2 \rightarrow \mathcal{D}_2$
(iii) $\mathcal{I}$ assigns to $\cdot$ a binary operation $\cdot^{\mathcal{I}} : \mathcal{D}_2 \times \mathcal{D}_2 \rightarrow \mathcal{D}_2$
(iv) $\mathcal{I}$ assigns to $(\cdot\forall\cdot)$ a mapping $\forall^{\mathcal{I}}$ from the function space of $\mathcal{D}_2$ to $\mathcal{D}_2$ itself, i.e. $(\cdot\forall\cdot)^{\mathcal{I}} : (\mathcal{D}_2 \rightarrow \mathcal{D}_2) \rightarrow \mathcal{D}_2$

Suppose we have a structure of the form $\langle \mathcal{D}_1, \mathcal{D}_2, \mathcal{I}\rangle$ where $\mathcal{D}_1 \subseteq \mathcal{D}_2$ and $\mathcal{I}$ meets the foregoing conditions. An *assignment* v is a mapping from proof variables to members of $\mathcal{D}_2$. It is not required that $v(x)$ be in $\mathcal{D}_1$. An assignment w is an *x-variant* of an

[36] These difficulties center on the fact that since $\neg(\exists x)x : \bot$ is a theorem of QLP, there is no straightforward way of interpreting $x : \varphi$ as $Proof_T(x^*, \ulcorner\varphi^*\urcorner)$ for any recursively axiomatizable $T \supseteq PA$. This follows for essentially the reasons which stand in the way of defining such an interpretation for S4 as discussed in footnote 6. We can also start with a fixed definition of an arithmetic interpretation such as this and then ask which statements in the language of QLP have arithmetical interpretations which are provable in T for all choices of $(\cdot)^*$ (note that this class does not include $\neg(\exists x)x : \bot$). The class of such statements has been shown not to be recursively axiomatizable by Yavorsky (2001).

assignment v if v and w agree on all variables except possibly for x. We write $v(x/r)$ for the x-variant of v that maps x to r. For an arbitrary assignment v, every proof term t is mapped to a member of $\mathcal{D}_2$, denoted by t^v, by the following recursive rules:

(i) $c^v = c^{\mathcal{I}}$ for a constant symbol c
(ii) $x^v = v(x)$ for x a variable
(iii) $(t \cdot u)^v = (t^v \cdot^{\mathcal{I}} u^v)$
(iv) $(t + u)^v = (t^v +^{\mathcal{I}} u^v)$
(v) $(!t)^v = !^{\mathcal{I}}(t^v)$
(vi) Suppose t^w has been defined for all w. For each x, define a mapping, $\langle \lambda x.t \rangle^v : \mathcal{D}_2 \to \mathcal{D}_2$ as follows: for each $r \in \mathcal{D}_2$, $\langle \lambda x.t \rangle^v(r) := t^{v(x/r)}$. We now define $(t\forall x)^v := \forall^{\mathcal{I}}(\langle \lambda x.t \rangle^v)$.

An *evidence function* $\mathcal{E}$ is a mapping that assigns to each $r \in \mathcal{D}_2$ and to each valuation v a set $\mathcal{E}(r, v)$ of formulas of $\mathcal{L}_E$. The range of this function $\mathcal{E}$ is a set of formulas. An evidence function is stipulated to satisfy the following closure conditions:

(i) Let φ be an axiom of QLPE and suppose $\varphi \in \mathcal{C}(c)$. Then $\varphi \in \mathcal{E}(c^{\mathcal{I}}, v)$.
(ii) If $(\varphi \to \psi) \in \mathcal{E}(r, v)$ and $\psi \in \mathcal{E}(s, v)$ then, $\psi \in \mathcal{E}((r \cdot^{\mathcal{I}} s), v)$.
(iii) If $\varphi \in \mathcal{E}(r, v)$ and t is any proof term such that $t^v = r$, then $t : \varphi \in \mathcal{E}(!^{\mathcal{I}}(r), v)$.
(iv) $\mathcal{E}(r, v) \cup \mathcal{E}(s, v) \subseteq \mathcal{E}((r +^{\mathcal{I}} s), v)$.
(v) Let t be a proof term and suppose that for all $r \in \mathcal{D}_2$, $\varphi \in \mathcal{E}(\langle \lambda x.t \rangle^v(r), v(x/r))$. Then $(\forall x)\varphi \in \mathcal{E}((t\forall x)^v, v)$.
(vi) If v and w agree on the free variables of φ, then $\varphi \in \mathcal{E}(r, v)$ iff $\varphi \in \mathcal{E}(r, w)$.

We now give the full definition of a QLPE model and what it means for a statement to be true in a model with respect to an assignment:

Definition 3.6 A QLPE model is a structure $\mathcal{M} = \langle \mathcal{D}_1, \mathcal{D}_2, \mathcal{I}, \mathcal{E}, \mathcal{V} \rangle$ such that $\langle \mathcal{D}_1, \mathcal{D}_2, \mathcal{I} \rangle$ is a pre-model, $\mathcal{E}$ is an evidence function and $\mathcal{V}$ is a propositional valuation. If $\mathcal{M}$ is a model and v is an assignment, we define truth in $\mathcal{M}$ relative to v as follows:

(i) $\mathcal{M} \vDash_v P_i$ iff $\mathcal{V}(P_i) = \mathrm{T}$
(ii) $\mathcal{M} \vDash_v E(t)$ iff $t^v \in \mathcal{D}_1$
(iii) $\mathcal{M} \nvDash_v \bot$
(iv) $\mathcal{M} \vDash_v \varphi \to \psi$ iff $\mathcal{M} \nvDash_v \varphi$ or $\mathcal{M} \vDash_v \psi$
(v) $\mathcal{M} \vDash_v (\forall x)\varphi$ iff $\mathcal{M} \vDash_w \varphi$ for every w where $w = v(x/r)$ and $r \in \mathcal{D}_2$
(vi) $\mathcal{M} \vDash_v (\exists x)\varphi$ iff $\mathcal{M} \vDash_w \varphi$ for some w where $w = v(x/r)$ and $r \in \mathcal{D}_2$
(vii) $\mathcal{M} \vDash_v t : \varphi$ iff $\varphi \in \mathcal{E}(t^v, v)$ and $\mathcal{M} \vDash_v \varphi$

Based on this definition of truth, we can now show that QLPE is sound with respect to the semantics we have just given.

Theorem 3.1 (Soundness) *For all QLPE models $\mathcal{M}$ and assignments v, if $\vdash_{\mathsf{QLPE}} \varphi$ then $\mathcal{M} \vDash_v \varphi$.*

Proof A straightforward adaptation of that given for QLP in Fitting (2004).

The Soundness Theorem is already sufficient to prove the consistency of Γ_ψ with respect to certain choices of ψ. Note that a caveat to this effect is necessary because

Γ_ψ will be inconsistent whenever ψ directly implies the negation of an instance of (KP4) (i.e. $\vdash_{\mathsf{QLPE}} \psi \rightarrow (\varphi \wedge \neg(\exists x)x : \varphi)$ for some $\varphi \in \mathrm{Sent}_E$). However, we do get the following:

Proposition 3.5 *Let Q be an arbitrary propositional letter. Then $\Gamma_Q \not\vdash \bot$.*

Proof By the Soundness Theorem, it suffices to construct $\mathcal{M}^*$, v such that $\mathcal{M}^* \models_v \Gamma_Q$. Let $\mathcal{M} = \langle \mathcal{D}_1, \mathcal{D}_2, \mathcal{I}, \mathcal{E}, \mathcal{V}\rangle$ be any QLPE model such that $\mathcal{V}(Q) = \mathrm{T}$ and $\mathcal{D}_1 = \emptyset$. Clearly $\mathcal{M} \models_v Q \wedge \forall x[E(x) \rightarrow \neg x : Q]$ for any v. Let r^+ be a new verification not in $\mathcal{D}_2$ and let $\mathcal{D}_2^* = \mathcal{D}_2 \cup \{r^+\}$. Define $\mathcal{E}^+$ as follows:

(i) $\mathcal{E}^+(r, v) = \mathcal{E}(r, v)$ for all v and $r \neq r^+$
(ii) $\mathcal{E}^+(r^+, v) = \{\varphi \mid \varphi \in \mathrm{Sent}_E\}$

Let $\mathcal{E}^*$ be the result of closing $\mathcal{E}^+$ under conditions (i)–(v) on evidence functions. Then $\mathcal{M}^* = \langle \mathcal{D}_1, \mathcal{D}_2^*, \mathcal{I}, \mathcal{E}^*, \mathcal{V}\rangle$ is also QLPE model. We now claim $\mathcal{M}^* \models_v \Gamma_Q$ for all v. Note that since we have changed neither $\mathcal{V}$ nor $\mathcal{D}_1$, $\mathcal{M}^* \models_v Q \wedge \forall x[E(x) \rightarrow \neg x : Q]$ for any v. Now suppose $\mathcal{M}^* \models_v \varphi$ and $\varphi \in \mathrm{Sent}_E$. Then $\mathcal{M}^* \models_{v(x/r^+)} \varphi$ (since $\varphi \in \mathrm{Sent}_E$ and thus has no free variables). Note that $\varphi \in \mathcal{E}^*(r^+, v)$ and hence also $\mathcal{M}^* \models_{v(x/r^+)} x : \varphi$. Thus $\mathcal{M}^* \models_v (\exists x)x : \varphi$ and hence $\mathcal{M}^* \models_v \varphi \rightarrow (\exists x)x : \varphi$. Since φ was arbitrary, we thus have $\mathcal{M}^* \models_v \Gamma_Q$. □

It may be noted with respect to the foregoing proof that the model $\mathcal{M}^*$ constructed in order to show the consistency of (KP4) and (NonOm3) is such that $\mathcal{M}^* \models_v \neg(\exists x)E(x)$ (since we set $\mathcal{D}_1 = \emptyset$ for both $\mathcal{M}$ and $\mathcal{M}^*$). It is thus natural to ask whether the consistency of Γ_ψ depends on our decision to not include closure principles for $E(x)$ among the axioms of QLPE itself. Some delicacy is required in order to answer this question, for as we have seen above it may not be clear exactly what sort of statements should be counted as reasonable closure principles for constructed proofs. Suppose, however, that we define $\mathfrak{C}$ to be the class of all *closed* proof terms—i.e. those containing no free proof variables. $\mathfrak{C}$ corresponds to the minimum class of terms whose denotations must be assumed to be constructed if, in addition to EAx and EApp, we assume analogous closure axioms for the operations $+$, ! and $(\cdot\forall\cdot)$. Note that if we assume that $E(t)$ for all $t \in \mathfrak{C}$, it follows from Proposition 3.2 that for all *provable* φ, we have $\varphi \rightarrow (\exists x)[E(x) \wedge x : \varphi]$. And thus the following schema

(CP) For all $t \in \mathfrak{C}$, $E(t)$.

expresses a relatively strong closure principle relative to the desiderata discussed above.

In order to demonstrate that (CP) does not bring (NonOm3) and (KP4) into conflict, we may note the following:

Corollary 3.1 *Let Q be an arbitrary propositional letter. Then $\Gamma_Q \cup \{E(t)|t\in\mathfrak{C}\} \not\vdash \bot$.*

Proof By the Soundness Theorem, it suffices to construct $\mathcal{M}^*$, v such that $\mathcal{M}^* \models_v \Gamma_Q \cup \{E(t) \mid t \in \mathfrak{C}\}$. Such an $\mathcal{M}^*$ can be constructed by starting with $\mathcal{M} = \langle \mathcal{D}_1, \mathcal{D}_2, \mathcal{I}, \mathcal{E}, \mathcal{V}\rangle$ such that $\mathcal{V}(Q) = \mathrm{T}$ and $\mathcal{D}_1 = \mathfrak{C}$. Note that since $\not\vdash_{\mathsf{QLPE}} Q$, it follows by induction on terms that for all $t \in \mathfrak{C}$, and all v, $\mathcal{M} \not\models_v t : Q$. From this it

follows that for all v, $\mathcal{M} \models_v Q \wedge \forall x[E(x) \rightarrow \neg x : Q]$. $\mathcal{M}^*$ can now be constructed from $\mathcal{M}$ by a straightforward adaptation of the construction given in Proposition 3.5.
□

The foregoing results show that regardless of whether we assume that all verifications denoted by closed terms are constructed, any statement of the form $Q \wedge (\forall x)[E(x) \rightarrow \neg x : Q]$ can consistently be conjoined with (KP4). This case appears to cover the sorts of (potential) examples of non-omniscience which are typically discussed in relation to the Knowability Paradox. For in attempting to illustrate the intuitive plausibility of the fact that there are true but unknown propositions, commentators have focused on contingent empirical statements which might express truths which no one ever bothers to verify—e.g.: "The comet contains pre-biotic molecules" (Edgington 1985), "The number of tennis balls in Timothy Williamson's garden on 4 July 1990 was even" (Williamson 1990), "The number of hairs on Wolfgang Künne's head on 1 January 2007 was odd" (Künne 2007). Note that such statements neither possess logical structure which can be discerned in propositional languages like $\mathcal{L}_{LK}$ or $\mathcal{L}_E$, nor do they themselves contain epistemic operators. As such they fall naturally under the scope of Proposition 3.5, which demonstrates that when knowledge and knowability are formalized according to (2a, b), no inconsistency results from the assumption that an unknown truth is expressed by an atomic statement even if all truths are assumed to be knowable.[37]

We take the results of this section to show in two different ways that there is no strictly logical way in which (KP4) and (NomOm3) conflict. For on the one hand, even when we reason from these principles using what amounts to classical first-order logic with quantifiers over proofs, no contradiction results. And on the other, Proposition 3.5 shows that since we can construct a QLPE model in which an instance of (NonOm3) is true together with all instances of (KP4), our inability to derive a contradiction from (KP4) and (NonOm3) by imitating the original reasoning of Fitch was not an accident. Note, however, that it follows from Proposition 3.3 that any model $\mathcal{M}$ satisfying (KP4) and (NonOm3) must be such that $\mathcal{M} \models (\exists x)\neg E(x)$. And from this it follows that for any such model there must exist verifications $r \in \mathcal{D}_2 \backslash \mathcal{D}_1$. Relative to the heuristic interpretation we assigned our semantics above, such r must constitute non-constructed or "merely potential" proofs. Even though a commitment to such entities seems to be already entailed by Propositions 3.3 and 3.4 themselves, it remains to be seen whether the neo-verificationists should assign these

[37] For the reasons just cited we take Proposition 3.5 to be sufficiently strong to settle the rhetorical issues about the consistency of verificationism which have animated much of the debate about the Knowability Paradox. However we have already noted that there will obviously be cases in which Γ_ψ will be inconsistent due to a direct conflict between NonOm3_ψ and an instance of (KP4). It is thus also natural to ask whether Proposition 3.5 can be extended to show that Γ_ψ is consistent for broader classes of formulas. The strongest result we know of this kind is the following: Let $\psi \in \text{Sent}_E$ not contain $E(x)$ and be such that all instances of $(\exists x)$ are in positive position and all instances of $(\forall x)$ are in negative position. Then $\Gamma_\psi \not\vdash_{\mathsf{QLPE}} \bot$ if and only if $\not\vdash_{\mathsf{QLPE}} \neg\psi$. The proof of this statement appears to require the completeness of QLPE with respect to the semantics we have presented. We have established this fact, but its proof is sufficiently involved that we have elected to omit it here. We note in passing, however, that a stronger result of this form can be used to demonstrate the consistency of Γ_ψ where ψ is of the form $(\exists x)x : Q$. This would be relevant in the case where we took Q to already express a provable (but currently unproven) proposition—e.g. (possibly) $Q \equiv$ "$2^{41,854,457} - 1$ is prime."

results anything more than instrumental significance. We will return to this issue in Sect. 4.2.

4 Taking stock

4.1 Comparisons and contrasts

A useful taxonomy of strategies for resolving the Knowability Paradox is provided by Brogaard and Salerno (2002), who distinguish between what they refer to as *logical revisionism*, *syntactic restrictionism* and *semantic revisionism*.[38] Logical revisionists seek to resolve the Paradox by changing the underlying logic which is employed to reason about these principles in a manner which prevents the Fitch derivation from going through. Syntactic restrictionists similarly seek to block the derivation by limiting the class of allowable substitution instances in (KP1). Semantic revisionists attempt to resolve the Paradox in a more general manner by replacing (KP1) with some other principle which they suggest plays a theoretical role similar to that of (KP), but which can be shown to be consistent with either (NonOm1) or an appropriate reformulation thereof.

Relative to this taxonomy, the resolution we have pursued here is a form of semantic revisionism. We hasten to add, however, that since the term "revisionism" is assessed relative to the formalization of (KP) and (NomOm) as (KP1) and (NonOm1), our acceptance of this categorization should not be taken to suggest that our approach is premised on the suggestion that neo-verificationism can be salvaged only by revising (KP1). Rather, our strategy has been to examine the views of theorists who are the presumptive targets of the Knowability Paradox and to argue that (KP1) represents, at best, a simplistic means of formalizing their views. From our perspective it is thus Hart's original use of (KP1) to formulate verificationism which constitutes an instance of revision.

We have already commented on the logical revisionist strategy in the course of addressing Williamson's proposal that verificationism can be salvaged by arguing that intuitionistic logic should be employed when reasoning from (KP1) and (NonOm1).[39] As we have portrayed the neo-verificationists as staunch supporters of intuitionistic logic, we are sympathetic to his desire to ground a resolution to the Paradox in considerations which have a basis in a detailed formulation of their views. Nonetheless, we have already noted that there are several reasons to think that the mere adoption of an

[38] Brogaard and Salerno actually employ the term "semantic restrictionism" to describe the proposals to which we wish to compare our own. However, since these proposals are put forth by theorists who advocate reformulating (KP) based on considerations about in the intended interpretation of such a principle, the term revisionism seems equally apt.

[39] A related proposal is that of Beall (2000) and Wansing (2002) who suggest that we use paraconsistent logic to reason about (KP1) and (NonOm1). Beall motivates his proposal on the basis of his claim that Montague and Kaplan's (1960) *Paradox of the Knower* provides independent motivation for accepting that the class of known statements is inconsistent. But not only do we fail to share this conviction (cf., Dean and Kurokawa 2008b), we take the proposal described herein to demonstrate that it is possible to resolve the Knowability Paradox in a manner which is both less radical and better motivated with respect to its intellectual origins.

intuitionistic background logic may be insufficient to adequately resolve the Paradox. And thus since we believe that the results we have presented in the previous section show that the Paradox can be resolved even if we reason classically, our approach provides no motivation for changing the background logic. And since our resolution also does not require restricting substitution into (KP), we feel the same way about the restrictionist strategy.[40]

The issue to which we now turn is that of comparing our proposed reformulation of (KP) to those of other semantic revisionists. The best known proposal in this class owes to Edgington (1985).[41] Edgington's proposal is grounded in the fact that we may distinguish between knowing that φ *is true* in a world w and knowing that φ *would be true* in another (potentially non-actual) world. On the basis of this observation, she proposes that (KP) be understood thusly:

(16) For all statements φ and worlds w_1, if φ is true in w_1, then there is a world w_2 in which it is known that φ is true in w_1.[42]

In order to reason about the consequences of this principle, Edgington proposes that we employ an actuality operator A with the intended interpretation $A\varphi$ if and only if φ is true in the actual world. Her proposed reformulation of (KP) thus takes the form

$$(\mathrm{KP_e}) \quad \text{For all } \varphi,\ A\varphi \rightarrow \Diamond KA\varphi.$$

Suppose we additionally assume that there is a statement which is both true and unknown in the actual world—i.e.

$$(\mathrm{NomOm_e}) \quad \text{For some } \psi,\ A(\psi \wedge \neg K\psi).$$

From this it follows that

(17) $\Diamond KA(\psi \wedge \neg K\psi)$

which can be taken to express that there is some world w where $KA(\psi \wedge \neg K\psi)$ holds.

In order to assess the consequences of (17), we must additionally note that Edgington's intended interpretation of $K\varphi$ is "φ is known by some agent i at some

[40] Dummett's (2001) original resolution to the Paradox fits into this class. He proposed to limit substitution into (KP) to logically unstructured (or "basic") sentences. This obviously blocks the derivation of the Paradox since no instance of (NonOm) satisfies this criterion. However, the cost of this step is high, as it significantly limits the class of statements to which an epistemic constraint on truth can be asserted to apply. We will discuss a more recent proposal of Dummett in Sect. 4.2.

[41] Another proposal in this class is that of Kvanvig (1995, 2006). His resolution is based on the observation that the occurrence of the possibility operator in (KP) creates a context in which the substitution of what he refers to as "modally non-rigid statements" (i.e. ones which may take on different truth values in different worlds) is not legitimate. Kvanvig's diagnosis of the locus of the Paradox is thus different from Edgington's. However, the revised version of (KP) which he formulates has problematic consequences similar to those incurred by Edgington's proposal which we will presently discuss.

[42] Edgington elects to explicate modal notions in terms of situations rather than worlds. She takes a situation to differ from a possible world in that the latter but not the former must be descriptively complete in settling the truth values of all propositions. Williamson (1987), however, has argued that if the modal notions involved in the traditional exposition of the Paradox are understood in this way, then ($\mathrm{KP_e}$) becomes trivial. We have thus modified our presentation of Edgington's interpretation of A and $\Diamond$ to take this into account.

time t." If we now let α denote the actual world, it follows from (17) by a variant of the Fitch derivation that $w \neq \alpha$. Thus from (17) it follows that there must be non-actual agents who at some non-actual time have knowledge not just about the actual world α, but also about the epistemic situation of the agents who reside in it.

Williamson has argued that since this consequence of (17) rivals the implausibility of (Om) itself, we must reject (KP_e) as a viable means of expressing an epistemic constraint on truth. In order to appreciate Williamson's concern, consider what would be required in order for an agent i in a world $w \neq \alpha$ to entertain a thought which is uniquely about α. i cannot refer to α by using an indexical device akin to Edgington's A operator, as when uttered in w, statements of the form $A\varphi$ express that φ is true not in the actual world α, but rather in his own world w. And in virtue of common assumptions about the nature of modality, i is also barred from referring to α either by ostension or through any form of causal interaction. Thus despite whatever intuitions might be taken to motivate (KP_e), this principle appears to lead to the untenable consequence that there must be "non-actual knowledge about the actual world."

We are sympathetic to Williamson's criticisms. And thus as fellow semantic revisionists, we must also ask whether the problem he raises applies to our own proposal for resolving the Paradox. On the face of things, it might not seem as if there is cause for concern as we have not explicitly employed modal or indexical devices which could result in the sort of consequence just described. However, recall the QLPE statement (3′) which arises in parallel to (17) by substituting an instance of (NonOm3) for φ in (KP4):

(3′) $(\exists y)y : (\psi \wedge (\forall x)[E(x) \rightarrow \neg x : \psi])$

As we have discussed above, this statement reports that there is a proof which simultaneously serves to verify ψ and also the statement $(\forall x)[E(x) \rightarrow \neg x : \psi]$ which we have taken to formalize the proposition that ψ is not known. As there is no guarantee that such a proof will be the denotation of a closed term, let us temporarily introduce the term p to denote a verification with this property. It would then follow from (3′) that $a_2 \cdot p : (\forall x)[E(x) \rightarrow \neg x : \psi]$. Relative to the interpretation of explicit modalities given by (5), this formula asserts that $a_2 \cdot p$ is a proof of $(\forall x)[E(x) \rightarrow \neg x : \psi]$. And relative to our analysis of knowledge in (2a), it then follows that if an agent were to have constructed $a_2 \cdot p$, he would thereby know $(\forall x)[E(x) \rightarrow \neg x : \psi]$. We must now ask ourselves whether this situation leads to any implausible consequences analogous to those entailed by (17).

Note first that Proposition 3.4 implies that $\neg E(a_1 \cdot p)$—i.e. that the proof denoted by $a_1 \cdot p$ has not been constructed. It is not possible to derive an analogous conclusion about $a_2 \cdot p$ unless we also assume that $E(x)$ satisfies some additional closure properties. But in this case, the conditions needed to derive $\neg E(a_2 \cdot p)$ are again quite plausible. For if we (i) assume EAx, then it follows that (ii) $E(a_1)$, from which (iii) $\neg E(p)$ follows, and then since (iv) $E(a_2)$ also follows from EAx, it follows that $\neg E(a_2 \cdot p)$ as long as we assume the "downward" closure condition that a proof cannot be constructed unless its constituents are. For purposes of argument, it thus seems reasonable to assume that the proof denoted by $a_2 \cdot p$ has not been constructed and, as such, is a "merely potential" proof. Note, however, that this proof itself serves to verify a statement which can be taken to express that no proof of ψ has actually been constructed. If we also assume that all proofs in a model of QLPE are potentially constructible (and hence potentially graspable by a deductive agent), the existence of $a_2 \cdot p$ might be taken to entail something akin to the conclusion that there must be non-actual agents who have knowledge of the actual world.

Despite the similarity in form between (17) and (3′), we do not believe that the latter should be taken to incur the same sort of problems relative to the interpretation we have argued that the neo-verificationists should assign to the language of QLPE. In order to see this, we must examine in slightly more detail the basis for including the predicate $E(x)$ in $\mathcal{L}_E$. We have thus far taken largely for granted the notion of constructedness which $E(x)$ is intended to express on the basis of the use of this notion by the neo-verificationists in framing the distinction between actual and potential proofs which we discussed in Sect. 2. However, if we wish to say more about this concept, it is difficult to resist adopting an analysis resembling

(18) proof t is constructed $\rightleftharpoons$ there exists some agent who has constructed t at or before the current time

The basis for adopting a temporal interpretation of $E(x)$ emerges most directly from Prawitz's remarks on the nature of potential proofs, which he repeatedly described as existing in an "abstract" or "tenseless" sense (cf., e.g., Prawitz 1987). Actual proofs are thus presumably understood by Prawitz in contradistinction to potential ones in virtue of the fact that they have already been concretely produced. And it thus follows that a constructed proof can be considered to be a concrete entity existing in time.

If we adopt this interpretation of $E(x)$, we must now ask if it is contradictory to assume that there can exist potential proofs which verify statements about our actual (i.e. current) epistemic situation. In order to see why the existence of such proofs does not have counterintuitive consequences, it will be useful to temporarily employ an idealized model of deductive activity which was first employed in intuitionistic mathematics. Relative to this model, mathematical activity is represented by the operation of a single idealized mathematical agent [IM], who comes to prove (or otherwise "experience the truth of") various statements at a series of discrete temporal stages which are assumed to be indexed by natural numbers 0, 1, ... Since the fact that IM has or has not proven a given statement at a particular stage seems to correspond to a contingent fact about his operation, it might first appear that the picture envisioned by this model is extraneous to the subject matter of mathematics itself. However Brouwer (1948) made explicit use of reference to the IM in order to produce weak counterexamples to principles in classical analysis. In particular, he assumed that it was allowable to refer to stages at which IM has *not* proven certain statements.[43]

The first attempt to formalize Brouwer's methods was given by Kreisel (1967), who proposed that the operation of IM could be axiomatized and then adjoined to a background system of intuitionistic analysis resulting in what came to be known as the Theory of the Creative Subject. We believe that this theory represents the closest

[43] For a reconstruction of one of Brouwer's so-called Creative Subject arguments with this property, see Troelstra and van Dalen (1988, pp. 842–843).

point of contact between an intuitionistic formal system and the neo-verificationists' distinction between actual and potential proofs.[44] The precise formulation of this theory need not detain us here as many of its details pertain to specific ways in which the operation of IM is assumed to be idealized. But there are two relevant senses in which we take the Creative Subject picture to *not* be an idealization of our actual deductive activity: (1) it allows that we may be ignorant of a statement now but come to know it later in virtue of coming to prove it; (2) it is consistent with conceiving of the construction of a proof as an activity which itself has duration.

These both seem to be undeniable features of our actual mathematical practices. For on the one hand, we conventionally speak of our mathematical knowledge as growing as we construct new proofs. And on the other, what we ultimately come to identify as a proof in practice is often the result of considerable effort on the part of a community of mathematicians whose work may proceed over a period of years. Thus it does not seem so far-fetched to think that proofs appear either with explicit dates attached (e.g. their dates of publication) or at least that the times at which they became constructed can be determined on the basis of *descriptive* (as opposed to purely indexical) features pertaining to the conditions under which they were produced.

These observations suggest that we may legitimately take the construction of proofs to be *events* which, under only mild idealization, can be assumed to have a definite beginning, middle and end. This in turn suggests that it is allowable to employ what Burgess has called *chronometry*—i.e. our normal means of dating events using calendrical terms like years, months, days, hours, etc.—as a canonical means of referring to the times at which the construction of proofs are completed (and at which they hence become constructed).[45] And for this reason, it does not appear to be necessary

[44] For a discussion of the interpretative issues which arise when we attempt to formulate the Theory of the Creative Subject axiomatically, see Dummett (1977). Williamson (1992) has also discussed this theory in the context of the Knowability Paradox. His primary conclusion is that the Paradox cannot be developed in the language which is commonly used to formulate this theory because this language allows us to quantify only over "abstract" or "mathematical" times and as such does not allow us to analyze the everyday concept of knowability (which Williamson takes to require knowledge at some "concrete" or "non-mathematical" time). Although we agree with this conclusion, we also believe the distinction which Williamson attempts to draw overlooks the fact that Brower employed reference to IM explicitly as a means of introducing reflection upon the *current* state of IM's knowledge into the practice of intuitionistic mathematics itself.

[45] This observation highlights the connection between the way we have suggested the neo-verificationists understand (KP) and its temporal analog—i.e. if φ is true, then φ *will* be known to be true—which Burgess refers to as the *Discovery Principle* (DP). One problem which Burgess suggests stands in the way of finding a plausible formulation of this principle is that of the applicability of chronometry to arbitrary events—i.e. not just ones involving the construction of proofs, but potentially durationless empirical phenomena like the collision of elementary particles. One difference between the intended interpretation we have assigned to the language of QLPE and that of the tense logics discussed by Burgess is that while the former only allows us to formulate tensed statements about the construction of proofs, the latter allows for the application of tense operators to arbitrary statements. We take the considerations mentioned in the text to suggest that the applicability of chronometry is relatively unproblematic in the former case. We note in passing one other salient difference between (DP) and the interpretation which the neo-verificationists may wish to assign to (KP3). We have argued that their notion of an *actual proof* is most naturally understood as one which has been constructed *now*. However, nothing about this interpretation requires that we also adopt a purely temporal understanding of the notion of *potential proof*. Thus although (KP3) requires that all truths are potentially proven, for all we have said here, it is consistent with this principle to deny that there is an *actual* time at which their proofs will be constructed.

to assume that the mechanism by which $E(x)$ succeeds in picking out just those proofs which have been constructed at the current time must be indexical. Rather, it could function in a descriptive manner which is analogous to the function achieved by reference to stages in the Theory of the Creative Subject. Admittedly, the means by which reference to our current state of knowledge is achieved in practice is considerably more complex than the use of natural numbers to index discrete times. It will, for instance, be informed by knowledge of a great many contingent facts about the course which our actual mathematical practice has taken in the past. But the fact that we are in possession of such a mechanism seems to be no more contentious than the claim that we can come to know that the sentence

(19) Every map is four colorable and it was not known prior to 1976 that every map is four colorable.

expresses a true proposition by surveying the recent literature on graph theory and topology.

4.2 Dummett and Prawitz on bivalence and potential proof

We have now seen by several different routes that if we represent the premises of the Knowability Paradox in terms of proof existence as (KP4) and (NonOm3), we reach the conclusion that there must exist non-constructed (or merely potential) proofs. This was the upshot not only of our informal argument based on decidable predicates in Sect. 2, but also of our attempt to reconstruct the Fitch derivation considered in Sect. 3.1 and its model theoretic analog considered in Sect. 3.2. We have thus far presented this outcome not only as a natural consequence of the attempt to define truth in terms of proof existence, but also as a viable means by which the neo-verificationists can respond to their critics in the face of the Paradox.[46]

The root of the controversy about the notion of potential proofs can be traced to a long-standing debate between Dummett and Prawitz about the proper characterization of truth within an anti-realist theory of meaning. As we have seen, Prawitz is quite explicit in his insistence that if the truth of a statement is to be identified with the existence of a proof which verifies it, we must allow that there exist proofs that have not been constructed. In particular, he has argued that such an understanding of proof existence is necessary if we wish to be able to account for the fact that mathematical

[46] In what, to the best of our knowledge, are his only remarks directed at the knowability Paradox itself, Prawitz appears to reach a similar conclusion: “The situation envisaged ... may be described by saying that there are *potential* verifications of the conjuncts but that there can never be an *actual* verification of the conjunction.” (Prawitz 1998b, emphasis in the original) This conclusion might be taken to recapitulate Hand’s observation that while statements of the form $\psi \wedge \neg K\psi$ can be verified in principle, they cannot be verified in practice on pain of defeating their own content. If we adopt this view, however, we must also accept that the connection between verifiability and knowledge is severed in the sense that although ψ is asserted to be verifiable, it cannot be known by any agent. Although Prawitz also accepts this conclusion, it follows that he must then accept the existence of potential verifications which are not even in principle constructible by an agent. For this reason, we propose that a better way for him to reply to the Paradox would be to reject the claim that $\psi \wedge \neg K\psi$ is an adequate formalization of (NonOm) in favor of the view that it should formalized in a manner more closely resembling (NonOm2).

statements like the Fermat Conjecture did not *become* true only when an appropriate proof was constructed, but rather have always been true.

Dummett (e.g. Dummett 1998) has argued that adopting such a view will inevitably drive us to accept that truth is *bivalent*—i.e. that every statement is either true or false—a position which is incompatible with intuitionistic logic and which he associates with realism. His argument appears to be based on the assumption that in accepting the notion of potential proof, Prawitz is also committed to regarding all proofs as inhabiting some determinate domain. In this case, Prawitz's view implies that a statement is true if there exists a proof which verifies it within this domain and false if there exists a proof which verifies its negation within this domain. Such a picture does not immediately collapse into bivalence because there might exist statements which are neither provable nor refutable. However, Dummett has argued that such a "trichotomous" situation—i.e. one in which all statements are either provable, refutable or neither provable nor refutable—implies the Law of the Excluded Middle [LEM] and hence bivalence.[47]

Based on what he takes to be the inexorable link between the recognition of potential proofs and bivalence, Dummett has claimed that the former notion cannot be consistently employed in the context of a semantic theory which is consonant with the principles of anti-realism. On the basis of their disagreement about the admissibility of potential proofs, one might reasonably conclude that while Prawitz can embrace the resolution to the Knowability Paradox which we have presented here, Dummett has no choice but to reject it. In the remainder of this section we will try to use some of the machinery of QLPE to examine whether this is actually the case.

The first observation we must make relates to the fact that QLPE is based on classical logic. This means that every instance of LEM is derivable in this system. And so, a fortiori, we have

(20) For all φ, $\vdash_{\mathsf{QLPE}} (\exists x)x : \varphi \vee \neg(\exists x)x : \varphi$.

If we identify the truth of φ with the existence of a proof which verifies it, and its falsity with the non-existence of such a proof (and additionally assume that such quantification is adequately expressed by the quantifiers of QLPE), then $(\exists x)x : \varphi \vee \neg(\exists x)x : \varphi$ can be taken to be a schematic representation of bivalence. Since this statement is provable in QLPE, this system cannot reasonably be taken to represent the theoretical commitments of the neo-verificationists even in the presence of (KP4).[48]

The rhetorical significance of this observation is not as clear cut as it might first seem. For as we suggested at the end of Sect. 1, the mere fact that QLPE as a whole

[47] In brief, Dummett's argument is as follows: (1) suppose $\neg\neg\varphi$; (2) this is equivalent to the claim that it is provable that φ cannot be refuted; (3) on the trichotomous view, this is in turn equivalent to the claim that it can be proven that either (i) φ is provable or (ii) that φ is neither provable nor refutable; (4) case (ii) is ruled out since it can be shown to be inconsistent for any statement to be provably neither provable nor refutable; (4) hence (i) must hold, which is equivalent to φ itself. This argument demonstrates that the trichotomous view licenses the inference from $\neg\neg\varphi$ to φ. But now note that since $\neg\neg(\varphi \vee \neg\varphi)$ is intuitionistically provable, so is $\varphi \vee \neg\varphi$.

[48] In fact, the adjunction of (KP4) to such a system might even be considered to be harmful. For note that if we decided to work in some system which satisfied (20) but not LEM itself, then it is easy to see that LEM would follow if we also adopted (KP4).

validates classical principles which the neo-verificationists do not accept need not be taken as a pretext for rejecting this system for purposes of resolving the Knowability Paradox. For it is at least open to these theorists to use the language of QLPE to formalize the premises of the Paradox, and then to simply reason from them according to whatever fragment of this system they accept. If they are willing to do this, there seems to be no reason why they should not endorse Propositions 3.3 and 3.4, as they are derivable intuitionistically. And as we argued in Sect. 3, these results can also be taken to point to a means of resolving the Paradox by acknowledging the existence of non-constructed proofs.

But inasmuch as the neo-verificationists might take the validity of (20) to stand in the way of employing QLPE in this manner, it is also reasonable to ask whether it is possible to formulate an intuitionistic version of this system which does not validate this principle. Such a system can be obtained by replacing by LP1 with the axioms of IPC and modifying the definition of constant specification accordingly. We will call this system iQLPE.[49] In assessing the status which the neo-verificationist ought to assign to iQLPE, we face several interpretative challenges. For on the one hand, iQLPE is strong enough to derive Propositions 3.3 and 3.4 which imply the existence of non-constructed proofs given the truth of any instance of (NonOm3). But on the other hand not only is $(\exists x)x : \varphi \vee \neg(\exists x)x : \varphi$ not a theorem of iQLPE, it may also be shown that this remains true when we adjoin (KP4)—i.e.

Proposition 4.1 *There exist* φ *such that* $\nvdash_{\mathsf{iQLPE+KP4}} (\exists x)x : \varphi \vee \neg(\exists x)x : \varphi$.[50]

Although together with (NonOm3), (KP4) does imply the existence of non-constructed proofs, the foregoing result suggests that this principle does not on its own imply bivalence. As such, we take Proposition 4.1 to bear against Dummett's argument that the acceptance of non-constructed or merely potential proofs inevitably leads to the acceptance of bivalence.

We may also observe, however, that (KP4) by itself is not sufficient to derive the existence of non-constructed proofs in either the classical or intuitionistic forms of QLPE. For although Propositions 3.3 and 3.4 are derivable in the intuitionistic system, to obtain the conclusion $(\exists x)\neg E(x)$, we obviously must also assume an instance of

[49] One technical complication which arises in the formulation of such a system relates to the decidability of the relation expressed by "t is a proof of φ." On the basis of the assumption that it is implicit in the notion of proof that we can recognize a proof of a statement when one is presented to us (a view which has its roots in Kreisel's (1962) reading of Wittgenstein (1967)), it has been assumed by most theorists that this relation must be decidable. This would appear to suggest that we ought to adopt as axioms of iQLPE all instances of $t : \varphi \vee \neg t : \varphi$. Such a schema is adopted by Artemov and Iemhoff (2007) in order to prove the arithmetical completeness of an intuitionistic form of LP relative to HA. We have elected not to adopt axioms of this form here both because they raise technical questions which are unrelated to the use of iQLPE in the present setting and also because of the independent arguments of Beeson (1985).

[50] This appears to require a semantic construction as developed in Dean and Kurokawa (2008a). In fact, once it is shown that $\nvdash_{\mathsf{iQLPE+KP4}} \neg\neg\varphi \rightarrow \varphi$, another proof of (4.1) can be given by reconstructing Dummett's trichotomy argument inside iQLPE. We also note in passing that result of adding a principle resembling (KP4) to an intuitionistic background theory has been discussed by a number of theorists, among them Beeson (1985), Pagin (1998) and McCarty (2006). The formalisms presented by these theorists are based on logical systems which extend the elementary propositional language of iQLPE to first-order logic or beyond. As such, we view iQLPE as a minimal intuitionist base for investigating the consistency of principles like (KP4) which attempt to define truth in terms of proof existence.

(NonOm3). If Dummett wishes to retain his critique of Prawitz's notion of potential proof, it thus appears that he has the option of denying (NomOm3). However, at this point the rhetorical landscape becomes complicated due to both the connection of this principle to (NonOm) (and thus to the standard setting of the Knowability Paradox) and also to the need for Dummett to find some means of formalizing his own Principle (K) in order to respond to the Paradox.

In a recent paper, Dummett (2007) takes up this issue directly in the context of revising the response to the Paradox he gave in Dummett (2001). His new view is that Principle (K) should be formalized not as (KP1) but rather as

(KP_d) For all φ, $\varphi \rightarrow \neg\neg K\varphi$.

Dummett's rationale for adopting this new principle is that in the context of intuitionistic logic $\neg\neg\varphi$ can be taken to express that φ is possible in the sense that it can be shown that the assumption that it is not the case leads to a contradiction. From (KP_d) it can be shown that

(Om_d) For all φ, $\neg(\varphi \wedge \neg K\varphi)$.

Dummett refers to (Om_d) as the "supposedly absurd" consequence of the Paradox. However, he elects to embrace this principle, again on the basis of an intuitionistic interpretation of negation. For understood in this manner, (Om_d) expresses not that all truths are known but rather that, for each φ, we will never be able to prove that φ is true but unknown (or more precisely, that the assumption that there is such a φ leads to a contradiction).

While this is an internally coherent means of responding to the Paradox, we also believe that it is by no means the best response that Dummett could give. For note that both (KP_d) and (Om_d) are intended to formalize principles which would be stated in natural language by using the term "known." In particular, (Om_d) presumably formalizes

(21) For all φ, it is not the case that φ and φ is not known.

Although such a statement is not intuitionistically equivalent to the Omniscience Principle (Om) itself, (Om_d) is at the very least counterintuitive as it prevents us from expressing the truism that we are not omniscient via (NonOm). For this reason alone, it seems that Dummett would be better off if he were not committed to accepting (21).

We believe that a resolution to the Paradox with this feature is near at hand. The key observation is the recognition that there is no intrinsic reason why Dummett must accept that informally stated claims like (21) or his own Principle (K) should be formalized using propositional modal operators as in (KP1). In fact we have seen in Sect. 2 that, like the other neo-verificationists, his acceptance of such principles appears to be premised on a background analysis of knowledge resembling (2a, b). Once this is acknowledged, (NonOm) is naturally understood as abbreviating a proposition whose full statement would more closely resemble (NonOm2). And as we pointed out above (cf., footnote 23), there are in fact occasions on which Dummett appears to *endorse* such a principle, despite the fact that it potentially stands in conflict with (Om_d). Thus if Dummett were to embrace (NonOm2) as a more felicitous analysis of (NonOm) than (NonOm1), he would have a rationale for rejecting statements like (Om_d) which appear to be radically out of keeping with our everyday intuitions.

At the same time, however, if (NonOm) is embraced, the necessity of modifying the formalization of (KP) in order to avoid the Paradox again becomes an issue. Dummett has undertaken this step by proposing to formalize (KP) via (KP_d). However, relative to the considerations adduced in Sect. 2, this treatment must at least be taken to be unmotivated. For like Prawitz and Martin-Löf, Dummett repeatedly states that truth should be analyzed in terms of proof existence in a manner which resembles the traditional interpretation of the intuitionistic connectives. Not only is the use of the knowledge operator in (KP_d) out of keeping with this approach, but this principle itself expresses a substantially weaker form of epistemic constraint on truth than Dummett appears to have endorsed in the past. There thus seem to be considerations which push him toward formalizing Principle (K) via a principle such as (KP3) which parallel those that recommend (NonOm2) over (Om_d).

What we take ourselves to have shown in this paper is that not only can the statements (NonOm2) and (KP3) be naturally formalized in a language with proof quantifiers, but that no analog to the Paradox arises once we take this step. But from these principles, it does follow by an intuitionistically valid argument that there exist non-constructed proofs. As is attested to by Proposition 4.1, however, formalizing (KP) as (KP4) does not commit us to such strong assumptions about proof existence as to entail bivalence. And for this reason, we take the considerations adduced in this section to present a tentative argument that Dummett's best response to the concerns raised by the Knowability Paradox is not to embrace (Om_d), but rather to abandon his critique of Prawitz's notion of potential proof. For by employing a system like iQLPE, it is possible to formulate (NonOm2) and (KP3) in a manner which enables us to show both that these principles are immune to the Paradox and also that they do not imply bivalence. More can be said about how Prawitz's views about meaning and verification contrast with those of Dummett by using this system. However, we leave this topic for another occasion.

References

Artemov, S. N. (2001). Explicit provability and constructive semantics. *The Bulletin of Symbolic Logic, 7*(1), 1–36.

Artemov, S., & Iemhoff, R. (2007). The basic intuitionistic logic of proofs. *Journal of Symbolic Logic, 72*(2), 439–451.

Beall, J. C. (2000). Fitch's proof, verificationism and the knower paradox. *Australasian Journal of Philosophy, 78*(2), 241–247.

Beeson, M. (1985). *Foundations of constructive mathematics: Metamathematical studies*. Berlin: Springer Verlag.

Binmore, K., & Shin, H. (1993). Algorithmic knowledge and game theory. In C. Bicchieri & M. L. D. Chiara (Eds.), *Knowledge, belief, and strategic interaction* (Chap. 9, pp. 141–154). Cambridge University Press.

Boolos, G. (1992). *The Logic of provability*. Cambridge: Cambridge University Press.

Brogaard, B., & Salerno, J. (2002). Fitch's Paradox of knowability. In E. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Stanford: Stanford University. Cited in http://plato.stanford.edu/cite.html.

Brouwer, L. (1948). Essentieel negative eigennschappen [Essentially negative properties]. *Indagationes Math, 10*, 322–323.

Brouwer, L. (1983). Consciousness, philosophy and mathemtics. In *Philosophy of mathematics: Selected papers* (2nd ed., pp. 90–96). New York: Cambridge University Press.
Burgess, J. P. (Forthcoming). Can truth out? In J. Salerno (Ed.), *New essays on the knowability paradox*. Oxford: Oxford University Press.
Cozzo, C. (1994). What can we learn from the paradox of knowability? *Topoi, 13*(2), 71–78.
Dean, W., & Kurokawa, H. (2008a). On semantics for iQLP and related systems (Manuscript).
Dean, W., & Kurokawa, H. (2008b). The knower paradox and the quantified logic of proofs. In A. Hieke (Ed.), *Proceedings of the Austrian Ludwig Wittgenstein Society* (Vol. 31).
Dummett, M. (1977). *Elements of intuitionism*. Oxford: Clarendon Press.
Dummett, M. (1978). The philosophical basis of intuitionistic logic. In *Truth and other enigmas* (pp. 215–247). New York: Oxford University Press.
Dummett, M. (1982). Realism. *Synthese, 52*(1), 55–132.
Dummett, M. (1987). Reply to Prawitz. In B. M. Taylor (Ed.), *Michael Dummett: Contribution to philosophy* (pp. 117–165). Dortrecht: Martinus Nijhoff Publishers.
Dummett, M. (1991). *The logical basis of metaphysics*. London: Duckworth.
Dummett, M. (1993). What is a theory of meaning? (II). In *The seas of language* (pp. 67–137). Oxford: Oxford University Press.
Dummett, M. (1998). Truth from the constructive standpoint. *Theoria, 64*, 123–138.
Dummett, M. (2001). Victor's error. *Analysis, 61*, 1–2.
Dummett, M. (2007). Reply to Künne. In R. Auxier & L. Hahn (Eds.), *The philosophy of Michael Dummett* (pp. 345–350). La Salle, IL: Open Court Press.
Edgington, D. (1985). The paradox of knowaiblity. *Mind, 94*, 557–568.
Fagin, R., & Halpern, J. (1988). Belief, awareness, and limited reasoning. *Artificial Intelligence, 34*, 39–76.
Fagin, R., Halpern, J. Y., Moses, Y., & Vardi, M. Y. (1995). *Reasoning about knowledge*. Cambridge: MIT Press.
Fitch, F. (1963). A logical analysis of some value concepts. *Journal of Symbolic Logic, 28*(2), 135–142.
Fitting, M. (2004) Quantified LP. Technical Report, CUNY PhD. Program in Computer Science Technical Report TR-2004019.
Fitting, M. (2005). The logic of proofs, semantically. *Annals of Pure and Applied Logic, 132*(1), 1–25.
Fitting, M. (2006). A quantified logic of evidence. *Electronic Notes in Theoretic Computer Science, 143*, 59–71.
Gödel, K. (1986). An interpertation of intuitionistic propositional calculus. In S. Feferman, et al. (Eds.), *Kurt Gödel. Collected works* (Vol. I, pp. 144–195). Oxford: Oxford University Press.
Goodman, N. D. (1970). A theory of constructions equivalent to arithmetic. In A. Kino, J. Myhill, & R. E. Vesley (Eds.), *Proceedings Summer Conference on Intuitionism and Proof Theory, Buffalo, NY, USA, Aug 1968, Studies in logic and the foundations of mathematics* (pp. 101–120). Amsterdam: North-Holland.
Hand, M. (2003). Knowablity and epistemic truth. *Australasian Journal of Philosophy, 81*(2), 216–228.
Hand, M. (2008). Performance and paradox. In *New essay on the knowability paradox*. Oxford: Oxford University Press.
Hart, W. (1979). The epistemology of abstract objects. *Aristotelian Society Supplementary Volume, 53*, 163–165.
Hart, W., & McGinn, C. (1976). Knowledge and necessity. *Journal of Philosophical Logic, 5*(2), 205–208.
Heyting, A. (1974). *Mathematische Grundlagenforschung, Intuitionismus, Beweistheorie*. Berlin: Springer.
Kahle, R., & Schroeder-Heister, P. (2006). Introduction: Proof-theoretic semantics. *Synthese, 148*(3), 503–506.
Kreisel, G. (1962). Foundations of intuitionistic logic. In E. Nagel, P. Suppes, & A. Tarski (Eds.), *LMPS* (pp. 198–210).
Kreisel, G. (1967). Informal rigour and completeness proofs. In I. Lakatos (Ed.), *Problems in philosophy of mathematics* (pp. 138–171). Amsterdam: North Holland.
Künne, W. (2007). Two Principles concerning truth. In R. Auxier & L. Hahn (Eds.), *The philosophy of Michael Dummett* (pp. 315–344). La Salle, IL: Open Court Press.
Kvanvig, J. (1995). The knowability paradox and the prospects for anti-realism. *Noûs, 29*(4), 481–500.
Kvanvig, J. (2006). *The knowability paradox*. Oxford: Clarendon Press.
Martin-Löf, P. (1984). *Intuitionistic type theory*. Napoli: Bibliopolis.
Martin-Löf, P. (1991). A path from logic to metaphysics. In *Atti del Congresso Nuovi problemi della logica e della filosopfia della scienza, Viareggio, 8–13 gennaio, 1990* (Vol. II, pp. 141–149). Bologna.
Martin-Löf, P. (1995). Verificationism then and now. In W. DePauli-Schimanovich, et al. (Eds.), *The foundational debate* (pp. 187–196). The Netherlands: Kluwer.

Martin-Löf, P. (1996). On the meanings of the logical constants and the justifications of the logical laws. *Nordic Journal of Philosophical Logic, 1*(1), 11–60. Text of lectures originally given in 1983 and distributed in 1985.
Martin-Löf, P. (1998). Truth and knowability: On the principles C and K of Michael Dummett. In G. Dales & G. Oliveri (Eds.), *Truth in mathematics* (pp. 105–114). New York: Oxford University Press.
Martino, E., & Usberti, G. (1994) Temporal and atemporal truth in intuitionistic mathematics. *Topoi, 13*, 83–92.
McCarty, D. (2006) The coherence of antirealism. *Mind, 115*, 947–956.
McKinsey, J. C. C., & Tarski, A. (1948). Some theorems about the sentential calculi of Lewis and Heyting. *Journal of Symbolic Logic, 13*, 1–15.
Mkrtychev, A. (1997). Models for the logic of proofs. In S. Adian & A. Nerode (Eds.), *Proceeding of LFCS* (pp. 266–275). Heidelberg: Springer.
Montague, R., & Kaplan, D. (1960). A paradox regained. *Notre Dame Journal of Formal Logic, 1*, 79–90.
Pagin, P. (1998). Bivalence: Meaning theory vs metaphysics. *Theoria, 64*, 37–66.
Parikh, R. (1987). Knowledge and the problem of logical omniscience. In Z. Ras & M. Zemankova (Eds.), *Methodologies for intelligent systems: Proceedings of the sixth international symposium* (pp. 432–439). Amsterdam: North Holland.
Percival, P. (1990). Fitch and intuitionistic knowability. *Analysis, 50*, 182–187.
Prawitz, D. (1977). Meanings and proofs: On the conflict between classical and intuitionistic logic. *Theoria, 43*(1), 2–40.
Prawitz, D. (1980a). Ideas and results in proof theory. In J. E. Fenstad (Ed.), *Proceedings of the second Scandinavian logic symposium* (pp. 235–307). Amsterdam: North Holland.
Prawitz, D. (1980b). Intuitionistic logic: A philosophical challenge. In G. H. von Right (Ed.), *Logic and philosophy* (pp. 1–10). Hague: Martinus Nijhoff Publishers.
Prawitz, D. (1987). Dummett on a theory of meaning and its impact on logic. In B. M. Taylor (Ed.), *Michael Dummett: Contribution to philosophy* (pp. 117–165). Dortrecht: Martinus Nijhoff Publishers.
Prawitz, D. (1998a). Comments on Göran Sundholm's paper "Proofs as acts and proofs as objects". *Theoria, 64*, 318–329.
Prawitz, D. (1998b). Comments on Lars Bergström's paper "Prawitz's Version of Verificationism". *Theoria, 64*, 293–303.
Prawitz, D. (1998c). Problems for a generalization of a verificationist theory of meaning. *Topoi: An International Review of Philosophy, 64*, 87–92.
Prawitz, D. (1998d). Truth and Objectivity from a verificationist point of view. In G. Dales & G. Oliveri (Eds.), *Truth in mathematics* (pp. 41–51). New York: Oxford University Press.
Raatikainen, P. (2004). Conceptions of truth in intuitionism. *History and Philosophy of Logic, 25*(2), 131–145.
Salerno, J. (2008). Knowablity noir. In *New essays on the knowability paradox*. Oxford: Oxford University Press.
Scott, D. (1970). Constructive validity. In M. Laudet, D. Lacombe, L. Nolin, & M. Schützenberger (Eds.), *Proceedings symposium on automatic demonstration, Versailles, France, December 1968*, Vol. 125 of *Lecture Notes in Mathematics* (pp. 237–275). Berlin: Springer-Verlag.
Smorynski, C. (1985). *Self-reference and modal logic*. Berlin: Springer-Verlag.
Sundholm, G. (1983). Constructions, proofs and the meaning of logical constants. *Journal of Philosophical Logic, 12*, 151–172.
Tennant, N. (1997). *The taming of the true*. Oxford: Oxford University Press.
Troelstra, A., & van Dalen, D. (1988). *Constructivism in mathematics* (Vols. I, II). Amsterdam: North Holland.
Usberti, G. (1995). *Significato e Conoscenza: Per una Critica del Neoverificazionismo*. Milano: Guerini Scientifica.
Wansing, H. (2002). Diamonds are a philosopher's best friends. *Journal of Philosophical Logic, 31*, 591–612.
Weinstein, S. (1983). The intended interpretation of intuitionistic logic. *Journal of Philosophical Logic, 12*, 261–270.
Williamson, T. (1982). Intuitionism disproved? *Analysis, 42*, 203–207.
Williamson, T. (1987). On the paradox of knowability. *Mind, 96*, 256–261.
Williamson, T. (1990). Two incomplete anti-realist modal epistemic logics. *The Journal of Symbolic Logic, 55*(1), 297–314.

Williamson, T. (1992). On intuitionistic modal epistemic logic. *Journal of Philosophical Logic, 21*, 63–89.
Wittgenstein, L. (1967). *Remarks on the foundations of mathematics*. In G. H. von Wright, R. Rees, & G. E. M. Anscombe (Eds.). MIT Press.
Yavorsky, R. E. (2001). Provability logics with quantifiers on proofs. *Annals of Pure and Applied Logic, 113*(1–3), 373–387.