

Introduction to knowability and beyond

Joe Salerno

Received: 7 April 2008 / Accepted: 25 August 2009 / Published online: 15 October 2009
© Springer Science+Business Media B.V. 2009

Beyond the limits of human knowledge lies nothing at all—no understandable proposition whose truth may outrun our idealized epistemic capacities. So says the semantic anti-realist, who advocates an epistemic theory of truth—one for which, as a matter of metaphysical necessity, anything true is knowable in principle. The interesting idealization remains elusive but is meant to lie somewhere between the triviality of equating truth with God’s knowledge (uninformative realism) and the naivety of equating it with what humans actually know (naive idealism). The former offers no explanatory gain, and the latter fails to appreciate the objectivity and discoverability of truth. The middle way aims for something like this: barring vagueness and ambiguity, we could in principle know the truth value of any fully understood proposition, given only finite improvements to our epistemic capacities, resources or environment. Call this “moderate anti-realism”. Many forms of non-realism are in the neighborhood of the thesis, including some versions of ethical expressivism, Michael Dummett’s mathematical intuitionism, Putnamian internal realism, Peircean pragmatism, logical positivism, Kantian transcendental realism, and Berkeleyian idealism.

One great problem for our understanding of the middle way is the knowability paradox—also known as “Fitch’s paradox” because Frederic Fitch first published the result in 1963, and the “Church–Fitch paradox” because Church was the anonymous referee who conveyed it to Fitch in 1945.¹ It consists of a proof that appears

¹ The referee reports are printed in their entirety in [Salerno \(2009\)](#).

J. Salerno (✉)
Department of Philosophy, Saint Louis University, Humanities Bldg., Rm 130, 3800 Lindell Blvd,
Saint Louis, MO 63104, USA
e-mail: knowability@gmail.com

J. Salerno
Australian National University, Canberra, ACT, Australia

to collapse the distinction between moderate anti-realism and naive idealism. That is because it shows that if each truth is knowable in principle, then indeed each truth is actually known. The Church–Fitch result is naturally expressed in the following manner:

$$(CFP) \forall p(p \rightarrow \diamond Kp) \vdash \forall p(p \rightarrow Kp),$$

where \diamond is an alethic possibility operator ‘it is possible that’, K abbreviates ‘somebody at some time knows that’, and our variables are propositional.

The premise is the *knowability principle*, and is thought best to capture the brand of anti-realism in question. The conclusion is often called the *omniscience principle*, because it avows a kind of group omniscience: each truth is at some time known by someone. Furthermore, if there is a true infinitary conjunction of all these truths, then there is someone that knows everything.² We might better take the non-infinitary conclusion as an expression *epistemic determinism*, following Proietti and Gabriel Sandu (this volume).

The proof of (CFP) is not complicated. On the epistemic front we suppose that knowing a conjunction requires knowledge of the conjuncts:

$$(A) K(p \ \& \ q) \vdash Kp \ \& \ Kq$$

and that only truths can be known:

$$(B) Kp \vdash p$$

On the modal front we suppose only that our modal epistemic theorems are necessary:

$$(C) \text{ If } \vdash p, \text{ then } \Box p$$

and that a proposition is necessary false only if it is impossible:

$$(D) \Box \neg p \vdash \neg \diamond p$$

Suppose for reductio that the Fitch conjunction, $p \ \& \ \neg Kp$, is known. Then, by (A), $Kp \ \& \ K\neg Kp$. By (B) and some trivial logic, $Kp \ \& \ \neg Kp$. The contradiction allows us to negate and discharge our only supposition, giving the theorem, $\neg K(p \ \& \ \neg Kp)$, which we may now necessitate by (C). So, by (D), our supposition is impossible. This together with the relevant instance of the knowability principle, $(p \ \& \ \neg Kp) \rightarrow \diamond K(p \ \& \ \neg Kp)$, entails $\neg(p \ \& \ \neg Kp)$. A classical equivalent is $p \rightarrow Kp$, which we may now generalize—giving $\forall p(p \rightarrow Kp)$. Heres the proof at a glance:

² Interesting discussed appears in Humerstone (1985), Bigelow (2005) and Fox (unpublished manuscript).

$$\begin{array}{c}
 \frac{}{K(p \ \& \ \neg Kp)} \text{ (1)} \\
 \frac{}{Kp \ \& \ K\neg Kp} \text{ (A)} \\
 \frac{}{Kp \ \& \ \neg Kp} \text{ (B) \ \& \ trivial \ logic} \\
 \frac{}{\neg K(p \ \& \ \neg Kp)} \text{ (1)} \\
 \frac{}{\Box\neg K(p \ \& \ \neg Kp)} \text{ (C)} \\
 \frac{}{\neg\Diamond K(p \ \& \ \neg Kp)} \text{ (D)} \\
 \hline
 \frac{\forall p(p \rightarrow \Diamond Kp)}{(p \ \& \ \neg Kp) \rightarrow \Diamond K(p \ \& \ \neg Kp)} \\
 \frac{\neg(p \ \& \ \neg Kp)}{p \rightarrow Kp} \\
 \forall p(p \rightarrow Kp)
 \end{array}$$

It should be clear that (CFP) is underwritten by modest modal and epistemic principles. The converse is trivial, since truth entails possibility. The lesson appears to be that a prima facie moderate anti-realism is inherently unstable, since it collapses into naive idealism. Consequently, if we have expressed anti-realism correctly and the thesis is true, then the seemingly modest epistemic idealization is not very modest at all since it blurs the difference between knowable truth and known truth:

$$\forall p(p \rightarrow \Diamond Kp) \vdash \forall p(p \rightarrow (\Diamond Kp \leftrightarrow Kp)).$$

The point is made more sharply if knowability is both a necessary and sufficient condition for truth. A so-called moderate anti-realist of this sort will not have the conceptual resources to distinguish possible from actual knowledge:

$$\forall p(p \leftrightarrow \Diamond Kp) \vdash \forall p(\Diamond Kp \leftrightarrow Kp).$$

So is there logical space for an interesting anti-realism that differs from the space occupied by naive idealism? One traditional strategy is to argue that an epistemic theory of truth carries with it independent reasons for rejecting classical logic in favor of intuitionistic logic. The penultimate step of the proof, which derives $p \rightarrow Kp$ from $\neg(p \ \& \ \neg Kp)$, abbreviates a piece of classical reasoning that is not intuitionistically valid. I quietly employed double-negation elimination. The intuitionist herself only gets as far as $p \rightarrow \neg\neg Kp$. Dummett (2009) happily accepts this intuitionistic consequence, and indeed thinks $\forall p(p \rightarrow \neg\neg Kp)$, rather than $\forall p(p \rightarrow \Diamond Kp)$, best expresses his epistemic theory of truth.

The first set of papers in the present issue explore options for one sympathetic to a theory of meaning and truth that falls within the tradition of Michael Dummett’s verificationism, although none of these papers is happy merely to appeal to intuitionistic logic. The second set focus on Dorothy Edgington’s situation-theoretic epistemic theory of actual truth, and part three explores interesting consequences about the essential limits of human knowledge. I shall here provide a bit of background³ and say more about the papers, none of which overtly dispute the validity of the Church–Fitch proof. The papers herein are dedicated either to the philosophical consequences of the proof, or to illuminating alternative approaches for the expression of moderate anti-realism.

³ A broader account of the history of the problem can be found in the introduction and Chap. 3 of Salerno (2009).

Tennant (1997, 2001) prefers to restrict the main quantifier in the knowability principle to Cartesian propositions. A proposition is Cartesian just in case knowing it does not entail a logico-mathematical absurdity. Since, the Fitch conjunction, $p \ \& \ \neg Kp$, is not Cartesian, it appears to pose no difficulty for the Cartesian knowability principle as an expression of anti-realism. Williamson (2000b) objects that not even all Cartesian truths are knowable in principle. The argument goes roughly like this. Suppose there is a Cartesian truth p that is not known—giving

$$(i) \ p \ \& \ \neg Kp.$$

In the intuitionistic relevant logic that Tennant favors, it follows that $p \ \& \ (Kp \rightarrow q)$, for arbitrary q . So it follows that both

$$(ii) \ p \ \& \ (Kp \rightarrow q)$$

and

$$(iii) \ p \ \& \ (Kp \rightarrow \neg q).$$

Now q must be Cartesian, if (ii) is to be Cartesian. That is because knowing (ii) entails knowing q . Analogously, $\neg q$ must be Cartesian, if (iii) is to be Cartesian. Williamson, substitutes ‘ n is even’ for ‘ q ’, where ‘ n ’ is a name, the reference of which is fixed rigidly by the description ‘the number of books actually now on my table’. This gives us,

$$(ii') \ p \ \& \ (Kp \rightarrow En)$$

and

$$(iii') \ p \ \& \ (Kp \rightarrow \neg En)$$

(ii') and (iii') will be knowable in Tennant's sense, only if ‘ En ’ and ‘ $\neg En$ ’ are both Cartesian. Williamson argues that they are. Now if (ii') is knowable, then it ultimately follows that it is possible that En ; and if (iii') is knowable, then it is possible that $\neg En$.⁴ Now, ‘ En ’ and ‘ $\neg En$ ’ are claims about a mathematical subject matter, for which possibility entails truth. Hence we may derive the contradiction, $En \ \& \ \neg En$. The contradiction forces us to reject our assumption that our Cartesian truth p is unknown. Generalizing, all Cartesian truths are known. Debate about the soundness of this argument can be found in Tennant (2001) and Williamson (2009). Variations on the argument that do not turn on the necessity of $q/\neg q$ appear in Rosenkranz (2004) and Brogaard and Salerno (2006).

The first paper in the present issue is by Neil Tennant. His paper, “Williamson's Woes”, is a self-contained response to Williamson (2009). Much of the discussion centers on whether (ii') and (iii') are both Cartesian. If they are, then both En and $\neg En$ are Cartesian. Williamson presupposes that En and $\neg En$ are propositions whose possibility entails truths. They are mathematical necessities. So it would seem that

⁴ That is because, necessarily, knowing (ii')/(iii') implies knowing $En/\neg En$. And since strict conditionals with possible antecedents have possible consequents, it follows that it is possible that $En/\neg En$.

either En or $\neg En$ is not Cartesian, and Williamson's argument is blocked. Much of this hangs on the nature of ' n ', the kind of term it is, and how precisely its reference is fixed. If ' n ' just were the description, 'the number of books on the table', then both ' n is even' and ' n is not even' would be Cartesian propositions, but we would lose the validity of the move from possibility to truth. After all, 'The number of books on the table is even' is a contingent claim. If n just were the numeral that refers to the number of books on the table, then the possibility of $En/\neg En$ would entail truth. However, it follows that either En or $\neg En$ would fail to be Cartesian since knowing it would mean knowing a mathematical absurdity. Ultimately, Williamson treats ' n is even' as an a posteriori necessity, akin to *Hesperus is Phosphorus*. Whether such a claim (and its negation) are Cartesian will depend on the consequence relation employed in the definition of a Cartesian proposition. Tennant evaluates his past debate with Williamson on these and related matters.

In the same section is a paper by Michael Hand, "Antirealism and Universal Knowability". It is a particularly helpful discussion of the options available for a Dummettian theory of truth. Dummett in places provides a modal characterization of the proof-conditional theory of mathematic truth. On that view truth is what we can prove/verify, and as such, puts its head in the Fitchian noose. Dag Prawitz provides an alternative characterization in terms of the existence of proof-types. On that view something is true just in case there exists a direct proof/verification of it (whether or not the object is or can be possessed). The connection between knowledge and proof/verification is more intimate when 'proof' is taken as 'proof-token', and less intimate when taken as 'proof-type'. Indeed, Hand argues that the existence of a proof-type doesn't entail its performability. There may be pragmatic reasons why some proof-types are unperformable. The Prawitzian proposal has the benefit of distancing anti-realism from the knowability paradox by distancing itself from the modal characterization of knowability. A question arises about the extent to which the Prawitzian ontology is a concession to realism. Hand argues that the ontology does not betray anti-realism, because not every form of recognition-transcendence is antithetical to an anti-realist meaning-theory.

Dorothy Edgington's contribution, "Possible Knowledge of Unknown Truth", develops her 1985 proposal, in which it is argued that one can know from a counterfactual perspective that, as things actually are, p and it is not known that p . The temporal analogy is helpful here: One cannot now know both that p is true and that p is not now known, but one can know at some time t both that p is now true and that it is not now known. So long as t and the present moment are not identical, no contradiction emerges. Similarly, one cannot in the actual situation know both that p is true and that p is not actually known, but in a counterfactual situation one can (without contradiction) know that, in the actual situation, p is true but unknown. Timothy Williamson's important objection (reiterated in 2000a) is that this sort of transworld knowability (or counterfactual knowledge of what is actually the case) is metaphysically impossible. The reason is that transworld knowability is knowledge *de re*; it is about the actual world uniquely. To achieve this, a merely possible agent would need to stand in the required causal relations or relations of acquaintance to the actual world. But she never does. So a merely possible agent cannot directly refer to, let alone have knowledge of, the actual world. On the other hand, if the reference to the actual world is fixed

descriptively, the suggestion might be that our possible thinker can know that the relevant proposition p is true at the world specified by the description. Williamson's response is that such knowledge would be uninterestingly trivial. The description would have to be detailed enough to uniquely specify the actual world. Knowledge that a proposition is true at such a world would be trivial logical knowledge, because it would amount to something like knowing that ' pn ' is true at the world where ' $\dots \& pn \& \dots$ ' is true.

Edgington's present paper responds to Williamson's concerns. She argues that one really can have non-trivial counterfactual knowledge of the actual situation. If we deny that possible-world talk always commits us to maximal specificity and deny a Lewisian metaphysics in which worlds/possibilities stand entirely in causally isolation from one another, then a plausible story can be had. Edgington makes use of the analogy of reference to future objects. I can refer to the future boiling pot of water, having just turned on the burner, even though the water was not boiling at the time of successful reference. And I can refer to a given future table, even if only the parts lie before me yet to be assembled. Just as the relevant shared temporal history of the thinker and future objects plays an important role in securing reference of those objects, the right kind of overlap between actual and counterfactual possibilities can help us to understand how reference and knowledge is achievable across worlds/possibilities. Edgington's view is that non-trivial counterfactual knowledge of what is actually the case is plausible, when the counterfactual thinker can identify her point of departure from the actual world and know the relevant details about how things would have developed from there.

Michael Fara's paper, "Knowability and the Capacity to Know", is less hopeful about counterfactual knowledge of actuality. Indeed, he finds Williamson's objections convincing and agrees that there is no metaphysically possible, but non-actual, world where someone knows a proposition of the form 'Actually p '. However, he argues that the relevant sense of 'knowability' does not employ a concept of metaphysical possibility. Rather it employs the notion of capacity. Fara distinguishes capacities from abilities, and goes on to argue that not all capacities are possibilities. For instance, one might have the capacity to swim across mile-wide rivers, even if for whatever deep physical reasons liquids could not form rivers a mile wide. And if time travel is impossible, then I cannot interact with people who no longer exist even though I have capacities to interact with such people. With the reinterpretation of the modal, Fara goes on to defend the Edgingtonian strategy. His favored knowability principle states that a proposition p is actually true, only if, in fact, someone has the capacity to know that p is actually true. Fara points out that this principle is unaffected by the familiar knowability paradoxes.

Carlo Proietti and Gabriel Sandu, in "Fitch's Paradox and *Ceteris Paribus* Modalities", focus on a quite different sort of worry that arises for Edgington's strategy. Within the standard semantics for basic modal logic, propositions of the form 'Actually p ' are necessary—giving $@p \leftrightarrow \Box @p$. This is problematic if we think that Edgington's knowability principle should tell us something about contingent, as well as necessary, truth. Moreover, when the knowledge operator K is treated as a universal modal operator (such that ' Kp ' is true iff ' p ' is true in every epistemically accessible possible world), we get the analogous epistemic collapse, $@p \leftrightarrow K @p$. That is, we regain the very sort of epistemic determinism that Edgington's strategy was

designed to avoid. Modeling the knowability paradox in a temporal framework and utilizing *ceteris paribus* modalities, Proietti and Sandu develop notions of necessity and knowledge that are not susceptible to the above collapses. They go on to formulate knowability principles in the spirit of Edgington’s proposal.

Jonathan Kvanvig argues in “The Incarnation and the Knowability Paradox” that the knowability paradox threatens Christian theistic metaphysics. A traditional doctrine of that metaphysics is the Incarnation, which says that Jesus was a person who had both a divine and a human nature. A divine nature includes having the usual omniproperties, including omniscience, essentially. Inconsistency threatens if a human nature includes having epistemic limitations essentially. Kvanvig lays out a standard reply to this kind of worry—viz., our epistemic limitations are accidental. Kvanvig explains that the doctrine of the Incarnation increases in plausibility if Jesus is not unique among humans in being open to omniscience; to the extent that any human could know any truth, then the omniscience of a divine human is not an ad hoc exception to the rule. However, this sets the theistic doctrine on a collision course with Fitch’s paradox [more specifically the contrapositive of (CFP)], which if valid guarantees some human unknowability in principle.

The paper by Richard Routley (later Richard Sylvan) also concerns the necessary limits of human knowledge. The paper is a reprint, first published in 1981. The source, however, is no longer in print and copies are exceedingly difficult to come by. So I thank Floor Oosting at Springer, the Synthese editors, and Nick Griffin, the executor of the Routley (Sylvan) estate, for allowing its inclusion. The paper is of historical importance for a number of reasons, not least because it includes the earliest suggestion for treating the knowability paradox with paraconsistent logic. In the context of a more general discussion of the inconsistency of Liars and Knowers (i.e., statements of the form “This very statement is not true” and “This very statement is not known”, respectively), Routley points out that the paraconsistentist has reason to embrace and contain the inconsistency that derives from assuming that the Fitch-conjunction is knowable. He does not endorse this strategy himself, but prescribes it only for the “hardened paraconsistentist”. The first defense of the strategy is found in Beall (2000), and further developments appear in Priest (2009) and Beall (2009). Routley’s paper is historically important for other reasons. It marks the earliest publication where interesting connections are drawn between liar, knower and knowability paradoxes, and it makes a case for thinking that the limits of human knowledge are necessary limits.

The contrapositive of (CFP) puts Routley’s argument in better focus. It says that any ignorance entails some unknowability in principle:

$$\exists p(p \ \& \ \neg Kp) \vdash \exists p(p \ \& \ \neg \diamond Kp)$$

More carefully, the existence of truth never known entails the existence of truth unknowable in principle. Unlike Gödel’s famous incompleteness proof showing that, for any consistent theory T (that includes arithmetic), there is a truth unprovable in T , the conclusion above promises a truth not verifiable by any means whatsoever. A question remains about the modal status of the conclusion. If our ignorance is only a contingent matter, it remains open whether the limits of human knowledge are contingent or necessary. However, if some ignorance is necessary, then the above

result (together with the modal platitude that only necessities follow from necessities) promises limits that are essential to human knowledge:

$$\Box\exists p(p \ \& \ \neg Kp) \vdash \Box\exists p(p \ \& \ \neg\Diamond Kp)$$

Routley provides reasons for thinking that our actual ignorance is not a contingent matter. The reasons are informal but are in the spirit of Gödel's thought. Just as Gödel restricts the epistemic means to theory T , Routley notices that humans are necessarily restricted by their numbers and capacities relative to the sheer size of space-time. As a result, and of necessity, some truths go forever unknown. By the above result, it follows that, of necessity, some truths are unknowable in principle.⁵ Routley's paper joins the wave of early reactions to Fitch's 1963 paper, which included Hart and McGinn (1976), Hart (1979), and Mackie (1980), all of which can be understood as employing Fitch's proof to advance criticisms against various forms of verificationism, the view that all meaningful statements (and so all truths) are knowable.

References

- Beall, Jc. (2000). Fitch's proof, verificationism, and the knower paradox. *Australasian Journal of Philosophy*, 78, 241–247.
- Beall, Jc. (2009). Knowability and possible epistemic oddities. In J. Salerno (Ed.), *New essays on the knowability paradox*. Oxford: Oxford University Press.
- Bigelow, J. (2005). Omnificence. *Analysis*, 65, 187–196.
- Brogaard, B., & Salerno, J. (2006). Knowability and a modal closure principle. *American Philosophical Quarterly*, 43, 261–270.
- Dummett, M. (2009). Fitch's paradox of knowability. In J. Salerno (Ed.), *New essays on the knowability paradox*. Oxford: Oxford University Press.
- Edgington, D. (1985). The paradox of knowability. *Mind*, 94, 557–568.
- Fitch, F. (1963). A logical analysis of some value concepts. *The Journal of Symbolic Logic*, 28, 135–142. (Reprinted in Salerno (Ed.), 2009.)
- Fox, J. *Fitch–Humberstone arguments*. Unpublished manuscript.
- Hart, W. D. (1979). The epistemology of abstract objects: Access and inference II. *Proceedings of the Aristotelian Society*, 53(Suppl.), 153–165.
- Hart, W. D., & McGinn, C. (1976). Knowledge and necessity. *Journal of Philosophical Logic*, 5, 205–208.
- Humerstone, I. (1985). The formalities of collective omniscience. *Philosophical Studies*, 48, 401–423.
- Mackie, J. (1980). Truth and knowability. *Analysis*, 40, 90–92.
- Priest, G. (2009). Beyond the limits of knowledge. In J. Salerno (Ed.), *New essays on the knowability paradox*. Oxford: Oxford University Press.
- Rescher, N. (2005). *Epistemic logic*. Pittsburgh: University of Pittsburgh Press.
- Rosenkranz, S. (2004). Fitch back in action again?. *Analysis*, 64, 67–71.
- Routley, R. (1981). Necessary limits of knowledge: Unknowable truths. In M. Edgar, N. Otto, & Z. Gerhard (Eds.), *Essays in scientific philosophy. Dedicated to Paul Weingartner/Philosophie als Wissenschaft. Paul Weingartner gewidmet* (pp. 93–115). Bad Reichenhall: Comes Verlag. (Reprinted, this volume.)
- Salerno, J. (Ed.). (2009). *New essays on the knowability paradox*. Oxford: Oxford University Press.
- Tennant, N. (1997). *The taming of the true*. Oxford: Clarendon.
- Tennant, N. (2001). Is every truth knowable? Reply to Williamson. *Ratio*, 14, 263–280.
- Williamson, T. (2000a). *Knowledge and its limits*. Oxford: Oxford University Press.
- Williamson, T. (2000b). Tennant on knowable truth. *Ratio*, 13, 99–114.
- Williamson, T. (2009). Tennant's troubles. In J. Salerno (Ed.), *New essays on the knowability paradox*. Oxford: Oxford University Press.

⁵ We find the argumentative strategy more formally in Rescher (2005).