

# The Paradox of Knowability from an Intuitionistic Standpoint

Gabriele Usberti

**Abstract** An intuitionistic solution to the Paradox of Knowability is given. It consists (i) in accepting  $\alpha \rightarrow K\alpha$ , the ordinary formalization of the principle of Radical Anti-Realism (RAR) that “Every truth is known”, since, intuitionistically understood, it means that proofs are epistemically transparent; and (ii) in accepting (RAR) itself, on the basis of the fact that knowledge is an intuitionistic internal truth notion. Some neo-verificationist approaches are criticized. Finally the problem of how to frame a rational discussion between Classicism and Intuitionism is briefly discussed.

**Keywords** Intuitionism · Knowability paradox · Anti-realistic theory of meaning · Truth notions · BHK-interpretation

## 1 Introduction

The Paradox of Knowability<sup>1</sup> is an argument that from the principle of Knowability (K) Every truth is knowable, accepted by anti-realists of all sorts, derives the principle of Radical Anti-Realism (RAR) Every truth is known, which, on the contrary, virtually no anti-realist would be prepared to accept. Contraposing, if one does not want to accept (RAR), one must reject (K) as well.<sup>2</sup>

---

<sup>1</sup>The paradox is usually ascribed to F. Fitch, but is due in fact to A. Church. For its history see [18].

<sup>2</sup>Of course, an intuitionist might also accept the negation of (K) without accepting that there are unknowable truths.

---

G. Usberti (✉)

Università degli Studi di Siena, DISPOC, Via Roma 56, 53100 Siena, Italy  
e-mail: gabriele.usberti@unisi.it

In a nutshell, the argument has the following structure. I shall use the symbols  $\&$ ,  $+$ ,  $\supset$ ,  $=$ ,  $-$ ,  $\Pi$ ,  $\Sigma$  for the classical logical constants; and  $\wedge$ ,  $\vee$ ,  $\rightarrow$ ,  $\leftrightarrow$ ,  $\neg$ ,  $\forall$ ,  $\exists$  for the intuitionistic ones. First, (K) and (RAR) are formalized by the two following schemas, respectively:<sup>3</sup>

- (1)  $\alpha \supset \Diamond K \alpha$
- (2)  $\alpha \supset K \alpha.$

Then replace  $\alpha$  in (1) with the proposition “ $q \& -K q$ ”; you obtain the following instance:

$$(3) \quad (q \& -K q) \supset \Diamond K(q \& -K q),$$

from which it is not difficult to derive, by means of intuitively acceptable principles, the unacceptable (2). The principles are the following:

- (a)  $\Box(K(\alpha \& \beta) \supset (K \alpha \& K \beta))$
- (b)  $\Box(K \alpha \supset \alpha)$
- (c)  $\Box((\alpha \& -\alpha) \supset \perp)$
- (d)  $-(\alpha \& -\beta) \supset (\alpha \supset \beta),$

and this is the derivation:<sup>4</sup>

$$\frac{\frac{\frac{[q \& -K q]^1}{\Diamond K(q \& -K q)} \text{ by (3)}}{\Diamond(K q \& K -K q)} \text{ by (a)}}{\Diamond(K q \& -K q)} \text{ by (b)}}{\Diamond(\perp)} \text{ by (c)} \quad \frac{-\Diamond(\perp)}{\perp} \supset E$$

$$\frac{\perp}{-(q \& -K q)} \text{ -I, 1} \quad \frac{-(q \& -K q)}{q \supset K q} \text{ by (d)}$$

The paradox is usually viewed as an argument against anti-realism, when this is conceived, according to a famous proposal by Michael Dummett, as a doctrine concerning meaning rather than ontology. Dummett writes:

Realism I characterise as the belief that statements of [a certain] class possess an objective truth-value, independently of our means of knowing it: they are true or false in virtue of a reality existing independently of us. The anti-realist opposes to this the view that the statements of [that] class are to be understood only by reference to the sort of thing which we count as evidence for a statement of that class.<sup>5</sup>

<sup>3</sup>“K” should be read as “It is known that”.

<sup>4</sup>Observe that this last step, and principle (d), are not intuitionistically valid.

<sup>5</sup>Reference [2], p. 146.

The fundamental opposition between realism and anti-realism concerns therefore, according to Dummett, the key notion of the theory of meaning, i.e. the notion in terms of which the meaning of the statements of the given class is to be explained: truth according to the realist, evidence according to the anti-realist. However, since Dummett holds that meaning is to be explained in any case in terms of truth-conditions, and that for the anti-realist truth can consist only in the existence of evidence, the realism/anti-realism opposition, in the final version he offers, concerns the notion of truth to be used in explaining meaning: the bivalent notion according to the realist, some non-bivalent notion according to the anti-realist.<sup>6</sup> The criterion of realism is therefore, in Dummett's opinion, the acceptance/refusal of the bivalence principle concerning truth.

Notice that this raises immediately a question: if we are ready to oppose the realistic, bivalent, notion of truth with other, non-bivalent, notions, we must ask at which conditions a notion can be considered as a notion of *truth*. I shall return to this question later. For the time being let me observe that the story of the semantic characterization of the realism/anti-realism debate is not finished. After Dummett it has been observed that all the anti-realistic notions of truth on the market (truth as assertibility, truth as existence of a verification, and so on) share a general characteristic: that truth is an epistemic notion, and therefore is essentially *knowable*. Knowability has therefore been identified as the essential property of truth, and anti-realism has been characterized as the view that every truth is knowable. It is precisely this feature of anti-realistic truth that the paradox is intended to hit.

In this paper I shall first argue that an intuitionistic solution to the paradox is available; then I shall examine the position of neo-verificationism concerning the paradox; finally I shall briefly consider the general question of how a rational discussion of alternative logics is possible at all.

## 2 An Intuitionistic Solution

The first step towards an intuitionistic solution is the remark that the formalizations of (K) and (RAR) by (1) and (2), respectively, acquire a meaning very different from the intuitive one if the logical constants occurring in them are understood according to the BHK-explanation, i.e. the explanation of their intuitionistic meaning given by Brouwer, Heyting and Kolmogorov, and that the intuitionistic formulas corresponding to (1) and (2), namely

$$(1') \quad \alpha \rightarrow \diamond K \alpha$$

$$(2') \quad \alpha \rightarrow K \alpha,$$

---

<sup>6</sup>As it will become evident below, the argument stated in this paper in no way relies on Dummett's opinion (for which see for instance [4], p. XXII) that also an anti-realist should explain meaning in terms of truth-conditions.

are valid, independently from there being a Church-Fitch argument (whose last step is not intuitionistically valid).

## 2.1 (2') is intuitionistically valid

Intuitionists do not agree with Dummett and other neo-verificationists that meaning is to be explained in any case in terms of truth-conditions. According to them, «The notion of truth makes no sense [...] in intuitionistic mathematics»<sup>7</sup>; the key notion of the theory of meaning is the notion of *proof*, and understanding  $\alpha$  (knowing its meaning) is to be explained as being capable to recognize the proofs of  $\alpha$ . The content of a mathematical statement  $\alpha$  (what Frege would have called the thought expressed by  $\alpha$ ) is characterized by Heyting as the expectation of a proof of  $\alpha$ . What a proof of  $\alpha$  is, is explained by recursion on the logical complexity of  $\alpha$ , under the assumption that we have an intuitive understanding of what is a proof of an atomic statement. This is the BHK-explanation. I think a revision of this explanation is necessary concerning disjunction and existential quantification.<sup>8</sup> According to Heyting, a proof of  $\alpha \vee \beta$  is either a proof of  $\alpha$  or a proof of  $\beta$ ; however, even the intuitionists consider, for instance, “Prime( $n$ )  $\vee$   $\neg$  Prime( $n$ )”, where  $n$  is some very large number, as assertible even if neither “Prime( $n$ )” nor “ $\neg$  Prime( $n$ )” is; I propose therefore to revise the BHK-explanation in the following way: A proof of  $\alpha \vee \beta$  is a procedure such that its execution yields,<sup>9</sup> after a finite time, either a proof of  $\alpha$  or a proof of  $\beta$ .<sup>10</sup> An analogous modification of the clause for  $\exists x\alpha$  can be similarly motivated.

Summing up, the revised version of the BHK-explanation I will make reference to is the following:

---

<sup>7</sup>Reference [12], p. 279. It should be stressed that Heyting is speaking of the realist, platonic, notion of truth. His assertion cannot therefore be understood as excluding that, within an intuitionistic framework, it is possible to define some notion that, on the one hand, can plausibly be proposed as a notion of truth, and, on the other hand, is reducible to others already present within that framework. I shall come back to this point in Sect. 2.2.

<sup>8</sup>For a more detailed motivation of this revision see [22], p. 42. The possibility of such a revision is explicitly envisaged, but discarded, in [3], p. 20.

<sup>9</sup>“Yields” is to be understood as equivalent to “is known to yield”.

<sup>10</sup>For instance, a proof of “Prime( $n$ )  $\vee$   $\neg$  Prime( $n$ )” is a primality test for  $n$ , i.e. an algorithm for determining whether  $n$  is prime. Such a test should not be confused with a general method consisting in applying to every number  $x$  a test for the primality of  $x$  (this general method is a proof of “ $\forall x(\text{Prime}(x) \vee \neg \text{Prime}(x))$ ”).

**Definition 1**

<i>A proof of</i>	<i>is</i>
$\alpha \wedge \beta$	a pair $\langle \pi_1, \pi_2 \rangle$ , where $\pi_1$ is a proof of $\alpha$ and $\pi_2$ is a proof of $\beta$
$\alpha \vee \beta$	a procedure $p$ whose execution yields, after a finite time, either a proof of $\alpha$ or a proof of $\beta$
$\alpha \rightarrow \beta$	a constructive function <sup>a</sup> $f$ such that, for every proof $\pi$ of $\alpha$ , $f(\pi)$ is a proof of $\beta$
$\neg\alpha$	a constructive function $f$ such that, for every proof $\pi$ of $\alpha$ , $f(\pi)$ is contradiction
$\forall x\alpha^\ddagger$	a constructive function $f$ such that, for every $c \in D$ , $f(c)$ is a proof of $\alpha(c)^\ddagger$
$\exists x\alpha^\ddagger$	a procedure $p$ whose execution yields, after a finite time, a pair $\langle c, \pi \rangle$ , where $c \in D$ and $\pi$ is a proof of $\alpha(c)^\ddagger$

<sup>†</sup> where  $x$  varies on  $D$

<sup>‡</sup>  $c$  is a name of  $c$

<sup>a</sup>Heyting speaks of “general method”; I shall use throughout “constructive function” (or briefly “function”) with the same meaning.

In order to arrive at defining the meaning of (2'), we must ask how to define a proof of  $K\alpha$ . There is no official intuitionistic answer, and there are several possibilities, according to the intended intuitive reading of  $K$ : “ $\alpha$  is presently known by someone”, “ $\alpha$  is known by someone at some time”, “ $\alpha$  is presently known by the one who is considering  $\alpha$ ”, and so on. I shall choose the last reading because, on the one hand, it seems to be the most congenial to intuitionistic ideas and, on the other hand, it is equivalent to the other readings for my present purposes. Here is my proposal:

**Definition 2** Whenever one is presented with a proof of  $\alpha$ , a proof of  $K\alpha$  is the observation that what one is presented with is a proof of  $\alpha$ .

In the light of Definitions 1 and 2, (2') expresses the expectation, for any proposition  $\alpha$ , of a function  $f$  associating to every proof  $\pi$  of  $\alpha$  the observation that  $\pi$  is a proof of  $\alpha$ .

Now, a fundamental characteristic of proofs, as the intuitionists conceive them, is that «for them, *esse est concipi*», to quote Dummett’s illuminating formulation.<sup>11</sup> In other words, a proof of  $\alpha$  is essentially what is recognized as such by an idealized knowing subject: there is no point of view from which something can be judged *to be* a proof of  $\alpha$  in spite of the fact that an idealized subject who is presented with it does not judge it as a proof of  $\alpha$ , or from which something can be judged not to be a proof of  $\alpha$  in spite of the fact that an idealized subject who is presented with it does judge it as a proof of  $\alpha$ . I shall call this characteristic of intuitionistic proofs their *epistemic transparency*; it can be expressed in the following way:

- (4) A proof of  $\alpha$  is *epistemically transparent* if and only if a subject who is presented with it is in a position to know that it is a proof of  $\alpha$ .

---

<sup>11</sup>«[M]athematical objects [...] are mental constructions [...] in the sense that, for them, *esse est concipi*.» [3, p. 7].

A little<sup>12</sup> reflection shows that, if proofs are transparent, the function  $f$  whose expectation is expressed by (2') does exist: by (4), for any proof  $\pi$  of  $\alpha$ , if one is presented with  $\pi$ , one is in a position to know that  $\pi$  is a proof of  $\alpha$ ; so one can associate to  $\pi$  the observation that  $\pi$  is a proof of  $\alpha$ ; and this, by Definition 2, is the proof of  $K\alpha$  required by the definition of  $f$ . Conversely, if a subject  $s$  knows, i.e. can compute, a function  $f$  associating to every proof  $\pi$  of  $\alpha$  the observation that  $\pi$  is a proof of  $\alpha$ , then  $s$ , when presented with a proof of  $\alpha$ , is in a position to know that it is a proof of  $\alpha$ . In conclusion (2'), far from saying that every intuitive truth is known, says that proofs are epistemically transparent, and is therefore obviously true; moreover, it remains true if  $K$  is read as "is known by someone at some time", since, if  $\alpha$  is presently known by me, then  $\alpha$  is known by someone at some time.

Williamson [25], pp. 430–1, raises the following objection to the validity of (2'). He first argues that a proof of  $\alpha \rightarrow \beta$  should be conceived by intuitionists as a function  $f$  from proof-tokens to proof-tokens «that is *unitype* in the sense that if  $p$  and  $q$  are proof-tokens of the same type then so are  $f(p)$  and  $f(q)$ .» Then, under the assumption that

a proof of  $\alpha \rightarrow K\alpha$  is a unitype function that evidently takes any proof-token of  $\alpha$  to a proof-token, for some time  $t$ , of the proposition that  $\alpha$  is proved at  $t$ ,

he shows that, if  $\alpha$  has not yet been decided, the function  $f$  that associates to every proof-token of  $\alpha$  a proof-token of the proposition that  $\alpha$  is proved at  $t$  is not unitype: if the proof-token  $p$  is carried out at  $t_1$  and the proof-token  $q$  is carried out at  $t_2$ , where  $t_1 \neq t_2$ , then  $f(p) \neq f(q)$ , since the proposition that  $\alpha$  is proved at  $t_1$  is different from the proposition that  $\alpha$  is proved at  $t_2$ . However, the quoted assumption is by no means conceptually necessary, nor is it a consequence of the general conception of proofs of conditionals as unitype functions. If we assume that a proof of  $\alpha \rightarrow K\alpha$  is a unitype function that takes any proof-token of  $\alpha$  to a proof-token of the proposition that  $\alpha$  is proved (with no mention of the time at which it is proved),  $f$  is unitype.

Let us now consider (1'). It is more difficult to suggest an intuitionistic reading for it, since it is not easy to devise a clear intuitionistic sense for the possibility operator. Here is one plausible candidate:

**Definition 3** A proof of  $\diamond\alpha$  is a procedure such that its execution yields,<sup>13</sup> after a finite time, a proof of  $\alpha$ .

According to this explanation, (1') expresses the expectation, for any proposition  $\alpha$ , of a function  $g$  associating, to every proof  $\pi$  of  $\alpha$ , a procedure  $p$  whose execution yields, after a finite time, the observation that what one is presented with is a proof of  $\alpha$ . Now, if we remember that the function  $f$  whose expectation is expressed by (2') does exist, and that  $f$  associates, to every proof  $\pi$  of  $\alpha$ , directly the empirical observation that  $\pi$  is a proof of  $\alpha$ , we see that also the function  $g$  exists: the procedure

---

<sup>12</sup>«To be in a position to know  $p$ , it is neither necessary to know  $p$  nor sufficient to be physically and psychologically capable of knowing  $p$ . No obstacle must block one's path to knowing  $p$ . If one is in a position to know  $p$ , and one has done what one is in a position to do to decide whether  $p$  is true, then one does know  $p$ . [...] Thus being in a position to know [...] is factive.» [27, p. 95].

<sup>13</sup>“Yields” is to be understood as equivalent to “is known to yield”.

consists precisely in effecting the empirical observation. In conclusion, also between the content of (I') and the intuitive content of (K) there is a substantial difference.

## 2.2 Truth Notions

The second step towards a solution consists in looking for a plausible intuitive sense of (K) and (RAR), according to which not only (K), but also (RAR), becomes acceptable. Of course, there is a sense in which (RAR) is *not* acceptable; my question is whether there is a sense in which it is. A first component of such a sense has already been made explicit: it consists in giving the logical constants their (revised) BHK-sense. The second component is of course the concept of truth, which (K) and (RAR) explicitly refer to. It is at this point that a problem mentioned above becomes relevant: at which conditions is a notion a notion *of truth*?

First, let me explain why, exactly, the question is crucial. If we read a formula of the language of classical propositional logic (CPL), it is natural and correct to read an occurrence in it of a propositional letter, say  $p$ , as “ $p$  is true”; for example, the intuitive reading of an instance of the schema  $\alpha + \neg\alpha$  would be, “Either  $p$  is true or  $\neg p$  is true” (which, given the definition of “ $p$  is false” as “ $\neg p$  is true”, is equivalent to “Either  $p$  is true or  $p$  is false”). This is correct because the key notion of the realistic explanation of the meaning of the logical constants is the realistic (i.e. bivalent) notion of truth; but it is no longer legitimate when we consider a formula of the language of IPL, since the key notion of the BHK-explanation is not the (bivalent) notion of truth. As a consequence, the simple occurrence of  $p$  will not be sufficient to make reference to the truth of  $p$ : in order to make reference to the truth of  $p$  it will be necessary to use a truth-predicate, or a truth-operator. Notoriously, the choice between expressing truth with a predicate or an operator has an impact on many other things, in particular on the possibility of expressing semantic paradoxes; since the questions discussed in this paper are independent of such a possibility, I shall choose the simpler alternative of expressing truth with an operator. The question arises at this point: what makes an operator a *truth* operator?

A plausible answer to this question is offered by Tarski's Convention T, in the case truth is expressed by a predicate. Tarski has proposed to consider a definition of truth as materially adequate if it entails every sentence of the form

$$(5) \quad N \text{ is true if and only if } t,$$

where  $N$  is the name of a sentence of the object language, and  $t$  is a translation of that sentence into the metalanguage. Since “materially adequate” means faithful to our intuitions about the notion of truth, we can take the validity of (5) as a criterion for a formally defined predicate to be a truth-predicate, i.e. a predicate defining a notion we are intuitively prepared to consider a notion of truth.<sup>14</sup> From this we may

<sup>14</sup>If I understand it correctly, [15], p. 148, makes essentially the same point.

easily extract an analogous condition for an operator: an operator  $O$  can be seen as a truth-operator if it is defined in such a way that it entails every sentence of the form

$$(6) \quad O\alpha \text{ if and only if } t,$$

where  $\alpha$  is a sentence of the object language and  $t$  is a translation of that sentence into the metalanguage. Finally, if we make the further simplifying assumption that the metalanguage is an extension of the object language, (6) is equivalent to

$$(7) \quad O\alpha \text{ if and only if } \alpha,$$

which is the usual version of what I shall call “The (T) Schema”.

So, my proposal is that an operator is to be considered as a truth operator if its meaning is defined in such a way as to satisfy the (T) Schema. Before going on, let me examine an objection to this proposal raised by Dummett. In *The Logical Basis of Metaphysics* he writes:

It is sometimes alleged that what makes a given notion a notion of truth is that it satisfies all instances of the (T) schema. This is wrong [...]. If a constructivist proposes that the only intelligible notion of truth we can have for mathematical statements is that under which they are true just in case we presently possess a proof of them, he is offering a characterisation of truth for which the (T) schema fails, since truth, so understood, does not commute with negation.<sup>15</sup>

Let me try to make the argument explicit. Dummett is envisaging the case of a constructivist who equates the truth of a (mathematical) statement  $\alpha$  with the actual possession of a proof of  $\alpha$ . The intuitionist may be seen as a case in point, and in a moment I myself shall explicitly endorse this view. At this point Dummett, assuming that a consequence of the (T) schema is that the following principle is valid:

$$(8) \quad T\neg\alpha \text{ if and only if } \neg T\alpha,$$

remarks that (8) is invalid when truth is equated to the actual possession of a proof (since from the fact that one does not possess a proof of  $T\alpha$  it does not follow that one possesses a proof of  $T\neg\alpha$ ), and concludes, by contraposing, that the (T) schema is not valid. Here is the derivation of (8) from the (T) schema:

$$\begin{array}{lll} (9) & (i) & T\neg\alpha \text{ iff } \neg\alpha \quad [\text{from (7), replacing } \alpha \text{ with } \neg\alpha] \\ & (ii) & \neg\alpha \text{ iff } \neg T\alpha \quad [\text{from (7), by contraposition}] \\ & (iii) & T\neg\alpha \text{ iff } \neg T\alpha \quad [\text{from (i) and (ii), by transitivity}]. \end{array}$$

It seems to me that Dummett’s remark that (8) is invalid is not correct:  $\neg T\alpha$  does not mean that one does not possess a proof of  $T\alpha$ , but that one possesses a method to transform every proof of  $T\alpha$  into a contradiction. When one possesses such a

---

<sup>15</sup>Reference [5], p. 166.



method, one has a proof of  $\neg\alpha$ ; owing to the epistemic transparency of intuitionistic proofs, one can effect the (empirical) observation that what one has is a proof of  $\neg\alpha$ , and this observation is a proof of  $K\neg\alpha$ , i.e. of  $T\neg\alpha$  under the present identification of truth with the actual possession of a proof. In conclusion, when the truth of a (mathematical) statement  $\alpha$  is equated with the actual possession of a proof of  $\alpha$ , truth does commute with intuitionistic negation.<sup>16</sup>

### 2.3 Internal and Intuitive Truth

The next question to consider is whether the validity of the (T) Schema picks out a unique notion of truth. Tarski seems to hold that it does. In [21] he expresses the conviction that the material adequacy condition imposed onto the definition of truth, is capable to select the classical Aristotelian notion of truth as correspondence. The conviction is not explicitly stated, but it can be inferred from the following facts:

(i) In section I.3 Tarski expresses an intention:

We should like our definition to do justice to the intuitions which adhere to the *classical Aristotelian conception of truth* [...] we could perhaps express this conception by means of the familiar formula:

*The truth of a sentence is its agreement with (or correspondence to) reality.*<sup>17</sup>

(ii) In section I.4 the same intention is made precise by requiring that the definition satisfies the material adequacy condition. Hence, the material adequacy condition ‘does justice’ to the intuitive notion of truth as correspondence. The question whether the intuitive notion of truth as correspondence is bivalent is not explicitly addressed by Tarski; an affirmative answer from him is suggested by the fact that in [20] he derives the principle of bivalence from the (materially adequate) definition of

---

<sup>16</sup>Notice that truth commutes with intuitionistic disjunction as well, since the principle

$$(*) \quad K(\alpha \vee \beta) \rightarrow (K\alpha \vee K\beta)$$

is valid. A proof of (\*) is a function  $f$  transforming every proof of  $K(\alpha \vee \beta)$  into a proof of  $K\alpha \vee K\beta$ ; a proof  $\pi$  of  $K(\alpha \vee \beta)$  is the observation that what one is presented with is a procedure  $p_1$  such that its execution yields, after a finite time, either a proof of  $\alpha$  or a proof of  $\beta$ . Define the following procedure  $p_2$ :

- apply  $p_1$ ;
- if the outcome is a proof of  $\alpha$ , perform the observation that what one is presented with is a proof of  $\alpha$ , getting a proof of  $K\alpha$ ;
- if the outcome is a proof of  $\beta$ , perform the observation that what one is presented with is a proof of  $\beta$ , getting a proof of  $K\beta$ .

Of course  $p_2$  is a procedure such that its execution yields, after a finite time, either a proof of  $K\alpha$  or a proof of  $K\beta$ ; we can therefore take  $p_2$  as  $f(\pi)$ .

<sup>17</sup>Reference [21], pp. 342–3.

truth.<sup>18</sup> However, the derivation crucially uses Excluded Middle, as is made clear by the following steps:

- (10)           (i)  $\alpha + \neg\alpha$            [the law of Excluded Middle]  
                  (ii)  $T\alpha \equiv \alpha$        [the (T) Schema]  
                  (iii)  $T\alpha + T\neg\alpha$    [from (i) and (ii) by Replacement].

So, Tarski's conviction is correct only under the premiss that the metalanguage is associated to a metatheory whose semantics validates Excluded Middle. If this principle is not valid in the metatheory, it is possible to exhibit counterexamples to Tarski's conviction, namely it is possible to define a non-bivalent notion of truth satisfying the (T) Schema. I shall now show how.<sup>19</sup>

Let us adopt a metatheory in which the logical constants are read according to the (revised) BHK-explanation (which, of course, does not validate the Excluded Middle). We are looking for a materially adequate definition, i.e. such that all the equivalences

$$(11) \qquad T\alpha \leftrightarrow \alpha$$

are logical consequences of it (where T is the intended truth operator). The definition I suggest is the following<sup>20</sup>:

**Definition 4**  $T\alpha =_{\text{def.}} K\alpha$ ,

where the meaning of K is defined by Definition 2. Definition 4 is materially adequate: (2') is valid for the reasons explained above; and the converse

$$(12) \qquad K\alpha \rightarrow \alpha$$

is valid as well: it expresses the expectation of a function  $h$  associating, to every observation  $o$  that what one is presented with is a proof of  $\alpha$ , a proof  $h(o)$  of  $\alpha$ , and  $h$  is warranted to exist by the factivity of proof observation.<sup>21</sup> In conclusion, the knowledge operator K is a truth operator, and of course this operator does not satisfy the principle of bivalence.

Concluding, it is true both (i) that the validity of the (T) schema expresses our essential intuition about the notion of truth, and (ii) that our most common intuitive notion of truth is realistic; but the reason why (ii) holds is bivalence, not the (T) schema: the validity of the (T) schema is neutral among different intuitive notions of truth. It is therefore possible, and necessary, to introduce a clear distinction between

<sup>18</sup>Reference [20], pp. 197–8. The principle of bivalence is called by Tarski “The principle of excluded middle”.

<sup>19</sup>Another example is the notion of truth defined in [14].

<sup>20</sup>Remember that the definition is intended to apply to mathematical statements.

<sup>21</sup>The factivity is warranted by the assumption that proof observation can plausibly be conceived as a computational process.

the condition at which an operator is a truth operator and the condition at which an operator reflects our realistic intuitions about the notion of truth; the former consists in the validity of the (T) schema,<sup>22</sup> the latter may be epitomized into the slogan of truth as correspondence and consequently into the validity of the law of bivalence. A notion satisfying the former condition is capable to play (at least some of) *the roles* of the notion of truth; truth as correspondence constitutes our predominant common-sense notion of truth. Between these two extremes there is a variety of truth notions, of which knowability and existence of a verification are two instances. I shall call “internal” these theoretical notions of truth, to stress the fact that each of them is capable to play the, or at least some of the, conceptual roles of the notion of truth within the framework of the related theory of meaning and of the formal semantics that adopts it. In this terminology we can say that bivalent truth is both the predominant intuitive notion of truth and the internal notion of classical logic; and that, besides it, there are several other internal notions of truth.

At this point it should be clear that an intuitive sense of the principle (RAR), according to which it becomes acceptable, does exist: for, if the logical constants are understood according to the BHK-explanation, and truth is understood according to Definition 4, then (RAR) is a tautology, saying that every known statement is known. The intuitionistic solution of the paradox consists therefore in accepting (RAR) as obvious when the logical constants are understood intuitionistically and truth as internal.

Is the idea of equating truth to knowledge consistent? There is an argument—called by [16] “The Standard Argument”—that purports to show that it is not.<sup>23</sup> It consists in the following derivation of  $\exists\alpha(\alpha \wedge \neg K\alpha)$  from the assumptions  $p \vee \neg p$  and  $\neg K p \wedge \neg K \neg p$ :

$$(13) \quad \frac{\mathcal{D} \quad \frac{(p \wedge \neg K p) \vee (\neg p \wedge \neg K \neg p)}{\exists\alpha(\alpha \wedge \neg K \alpha)} \quad \frac{[p \wedge \neg K p]^3}{\exists\alpha(\alpha \wedge \neg K \alpha)} \quad \frac{[\neg p \wedge \neg K \neg p]^4}{\exists\alpha(\alpha \wedge \neg K \alpha)}}{\exists\alpha(\alpha \wedge \neg K \alpha)} 3, 4$$

where  $\mathcal{D}$  is:

$$\frac{p \vee \neg p \quad \frac{[p]^1 \quad \frac{\neg K p \wedge \neg K \neg p}{\neg K p}}{p \wedge \neg K p} \quad \frac{[\neg p]^2 \quad \frac{\neg K p \wedge \neg K \neg p}{\neg K \neg p}}{\neg p \wedge \neg K \neg p}}{\frac{(p \wedge \neg K p) \vee (\neg p \wedge \neg K \neg p)}{(p \wedge \neg K p) \vee (\neg p \wedge \neg K \neg p)}} 1, 2$$

<sup>22</sup>The claim that an operator O is a truth operator iff it satisfies the schema (7) should not be confounded with the minimalist claim that (7) is a good definition of the meaning of O. The former claim is perfectly compatible with the idea, embraced above, that the validity of (7) is not the definition, but the material adequacy condition of the definition, of O.

<sup>23</sup>See [16], p. 275.

Now, if we observe that there are statements  $p$  that the intuitionist acknowledges as being decidable (i.e. such that  $p \vee \neg p$  is assertible), and that, as a matter of fact, are unknown (i.e., such that  $\neg K p \wedge \neg K \neg p$  is true),<sup>24</sup> we obtain that  $\exists \alpha (\alpha \wedge \neg K \alpha)$  is assertible.

My answer consists in observing that the argument is valid but unsound, since  $\neg K p \wedge \neg K \neg p$  is intuitionistically inconsistent. Assume that  $\neg K p \wedge \neg K \neg p$  is assertible, and reason in the following way:

$$(14) \quad \frac{\frac{\frac{[\neg K p \wedge \neg K \neg p]^1}{\neg K p} \quad \frac{[p]^2}{K p}}{\perp} 2 \quad \frac{\frac{[\neg K p \wedge \neg K \neg p]^1}{\neg K \neg p} \quad \frac{[\neg p]^3}{K \neg p}}{\perp} 3}{\perp} 1}{\neg(\neg K p \wedge \neg K \neg p)}$$

The formula  $\exists \alpha (\alpha \wedge \neg K \alpha)$  may therefore be false. In order to show that it is actually false, let us wonder whether there could be a proof of it, i.e. a procedure  $p$  whose execution yields, after a finite time, a pair  $\langle c, \pi \rangle$ , where  $c$  is a proposition and  $\pi$  is a proof of  $c \wedge \neg K c$ . A proof of  $c \wedge \neg K c$  is a pair  $\langle \pi_1, \pi_2 \rangle$ , where  $\pi_1$  is a proof of  $c$  and  $\pi_2$  is a proof of  $\neg K c$ ; such a pair cannot exist, on pain of contradiction: being presented with  $\pi_1$ , one can effect the observation that what one is presented with is a proof of  $c$ , thereby obtaining a proof  $\pi_3$  of  $K c$ ; coupling  $\pi_3$  with  $\pi_2$  we obtain a proof of  $K c \wedge \neg K c$ : a contradiction;  $p$  cannot therefore exist. In conclusion, the intuitionist cannot assert  $\exists \alpha (\alpha \wedge \neg K \alpha)$ , and the idea of statements that, being unknown, are not yet true nor false is not inconsistent.

The intuitionistic inconsistency of  $\neg K p \wedge \neg K \neg p$  may sound unacceptable from the intuitive standpoint, since it seems to conflict with the idea, which also an intuitionist should accept, that there are undecided, hence unknown, statements. Here it is important, again, to pay attention to the intuitionistic meaning of the logical constants, in particular of negation. The assertibility of  $\neg(\neg K \alpha \wedge \neg K \neg \alpha)$  means that a method is known to transform every proof of  $\neg K \alpha \wedge \neg K \neg \alpha$  into a contradiction, hence that a logical obstacle is known to the possibility that there is a *proof* of  $\neg K \alpha \wedge \neg K \neg \alpha$ ; it does not exclude the *fact* that neither  $\alpha$  nor  $\neg \alpha$  are known. We will see in a moment whether the existence of such a fact can be acknowledged within the intuitionistic conceptual framework. Before, I want to comment upon the existence of a logical obstacle to the possibility that there is a *proof* of  $\neg K \alpha \wedge \neg K \neg \alpha$ . This is neither unacceptable nor unexpected if we keep present that the operator  $K$  is, in intuitionistic logic, a truth-operator; for it is a principle valid in general, i.e. for *every* internal notion of truth, that the formula expressing the proposition “ $p$  is neither true nor false” is inconsistent. Take for instance the formula  $\neg T \alpha \& \neg T \neg \alpha$ , expressing the same proposition within classical logic, and reason exactly in the same way as in (14), simply replacing  $\neg$  with  $\neg$ , and  $\wedge$  with  $\&$ . The crucial step is the

<sup>24</sup>An example is “Prime( $n$ )”, where  $n$  is some very large number.

inference of  $T\alpha$  from  $\alpha$ ; in other terms, the inconsistency of the formula expressing the proposition “ $\alpha$  is neither true nor false” depends on the validity of the principle  $\alpha \rightarrow T\alpha$  (together with propositional laws that are common to classical and intuitionistic logic). We have seen that the reason why that principle is intuitionistically valid is the assumption that proofs are epistemically transparent; of course this very assumption may be questioned,<sup>25</sup> but the issue of its truth or falsity is utterly different from the question whether there are intuitive truths that, as a matter of fact, are unknown.

## 2.4 Unknown Statements

I have said that the assertibility of  $\neg(\neg K\alpha \wedge \neg K\neg\alpha)$  does not exclude the existence of the fact that neither  $\alpha$  nor  $\neg\alpha$  are known. Can the intuitionist *assert* the existence of such a fact? I think not, and in this section I shall try to motivate this opinion.

Let me observe first that “ $K\alpha$ ”, in all its possible readings, clearly is an *empirical* statement, not a mathematical one. I have argued elsewhere that the negation of many empirical statements, and in particular of  $K$ -statements, cannot be plausibly equated to intuitionistic negation  $\neg$ , and I have proposed that it be equated to Nelson’s strong negation  $\sim$ .<sup>26</sup> So, if we add  $\sim$  to the language  $\mathcal{L}_{\text{IPL}K}$  of Intuitionistic Propositional Logic plus the operator  $K$ , and we assume for simplicity that all the empirical sentences of  $\mathcal{L}_{\text{IPL}\sim K}$  have proofs,<sup>27</sup> we must add, to the Definition 2 of the notion of proof of  $K\alpha$ , a definition of the notion of proof of  $\sim K\alpha$ . Here is my proposal:

**Definition 5** Whenever one is presented with something that is not a proof of  $\alpha$ , a proof of  $\sim K\alpha$  is the observation that what one is presented with is not a proof of  $\alpha$ .

It should be noticed that Dummett’s remark—that intuitionistic truth, when it is equated with the actual possession of a proof, does not commute with negation—is certainly correct when it is understood as referring to strong negation. For example, the observation that what one is presented with is not a proof that it is raining is not the same thing as the observation that what one is presented with is a proof that it is not raining. As a consequence, Dummett’s objection to the validity of the (T) schema as a criterion for being a truth operator seems to cause trouble in this case. However, in this case the argument (9) is no longer valid: the second step is an application of contraposition, but contraposition is not valid for strong negation. As a consequence, the fact that strong negation does not commute with truth does not entail the invalidity of the (T) schema. We can therefore conclude that, even when we add to intuitionism strong negation, the knowledge operator  $K$  is a truth operator.

---

<sup>25</sup>A discussion of this assumption is beyond the limits of this paper.

<sup>26</sup>Reference [24].

<sup>27</sup>In general empirical sentences have (non-conclusive) justifications. A definition of the notion of justification for the sentences of  $\mathcal{L}_{\text{IPL}\sim K}$  presupposes a solution of Gettier problems. I have suggested such a definition in [23].

It seems to me that, if the existence of undecided statements can be expressed at all in an intuitionistic language, it should be expressible in  $\mathcal{L}_{\text{IPL}\sim, \text{K}}$ . Take for instance  $g$ , Goldbach's Conjecture: for the statement

(15) Goldbach's conjecture is undecided (unknown)

the following formula seems to be a plausible formalization in  $\mathcal{L}_{\text{IPL}\sim, \text{K}}$ :

(16)  $\sim \text{K } g \wedge \sim \text{K } \neg g.$

Williamson argues against this formalization:

if  $\sim$  is to count intuitionistically as any sort of negation at all,  $\sim A$  should at least be inconsistent with  $A$  in the ordinary intuitionistic sense.<sup>28</sup>

In other words, the schema

(17)  $\sim \alpha \rightarrow \neg \alpha;$

should be valid; then, from (16) one could derive  $\neg \text{K } g \wedge \neg \text{K } \neg g$ , which, by (2) and (13), is equivalent to  $\neg g \wedge \neg \neg g$ : a contradiction. However, the assumption that (17) is valid for all  $\alpha$  of  $\mathcal{L}_{\text{IPL}\sim, \text{K}}$  seems to be a sort of *petitio principii*, since, on the one hand, it almost amounts to assuming what one wants to conclude, i.e. that (16) is inconsistent, and, on the other hand, the motivation for it seems insufficient. Notice that (17) is valid for all  $\alpha$  belonging to  $\mathcal{L}_{\text{IPL}\sim}$ ,<sup>29</sup> so, according to Williamson's criterion,  $\sim$  does count intuitionistically as a sort of negation; the possible invalidity of (17) when  $\alpha$  contains occurrences of  $\text{K}$  can therefore be imputed to the interplay between the meanings of  $\text{K}$  and  $\sim$ . On the other hand, (17) is clearly invalid when  $\alpha$  contains occurrences of  $\text{K}$ . Consider the instance

(18)  $\sim \text{K } \alpha \rightarrow \neg \text{K } \alpha:$

it asserts the existence of a function  $f$  associating to every proof of the antecedent a proof of the consequent. A proof of the antecedent is the observation that what one is presented with is not a proof of  $\alpha$ ; this observation is true in two cases: when one is presented with a proof of  $\sim \alpha$ , and when one is presented with some  $x$  that is neither a proof of  $\sim \alpha$  nor a proof of  $\alpha$ . In the second case  $f$  should associate to  $x$  a function  $f'$  associating to every proof of  $\text{K } p$  a contradiction; but  $f'$  cannot exist: as  $x$  is not a proof of  $\sim \alpha$ , the existence of a  $y$  that is a proof of  $\alpha$  cannot be ruled out, and if one observes that  $y$  is a proof of  $\alpha$ , that observation is a proof of  $\text{K } \alpha$ .

However, if we look at the interplay between intuitionistic logical constants, strong negation and  $\text{K}$  from the standpoint of Kripke semantics, the assertibility of (18) seems to be out of the question. A *Kripke model* for  $\mathcal{L}_{\text{IPL}\sim}$  is a quadruple

<sup>28</sup>Reference [26], p. 139.

<sup>29</sup>Reference [9].

$\mathcal{M} = \langle W, \geq, D, V \rangle$ , where  $W$  is a non-empty set (of nodes),  $\geq$  is a reflexive partial order on  $W$ , and  $V$  is a partial function from atomic formulas and nodes to  $\{0, 1\}$  satisfying the following conditions:

- If  $V(p, w) = 0$  and  $wRw'$ , then  $\text{Val}(p, w') = 0$ ;  
if  $V(p, w) = 1$  and  $wRw'$ , then  $\text{Val}(p, w') = 1$  (stability).
- For every  $w \in W$ ,  $V(\perp, w) = 0$ .  
For every  $w \in W$ ,  $V(\sim\perp, w) = 1$ .

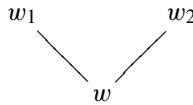
The notion  $\models_w \alpha$  ( $\alpha$  is true at  $w$ ) is defined by induction on  $\alpha$  as follows:

$$\begin{aligned} \models_w p &\text{ iff } V(p, w) = 1 \\ \models_w \sim p &\text{ iff } V(p, w) = 0 \\ \models_w \sim\perp & \\ \models_w \alpha \wedge \beta &\text{ iff } \models_w \alpha \text{ and } \models_w \beta \\ \models_w \sim(\alpha \wedge \beta) &\text{ iff } \models_w \sim\alpha \text{ or } \models_w \sim\beta \\ \models_w \alpha \vee \beta &\text{ iff } \models_w \alpha \text{ or } \models_w \beta \\ \models_w \sim(\alpha \vee \beta) &\text{ iff } \models_w \sim\alpha \text{ and } \models_w \sim\beta \\ \models_w \alpha \rightarrow \beta &\text{ iff, for every } w' \geq w, \text{ if } \models_{w'} \alpha \text{ then } \models_{w'} \beta \\ \models_w \sim(\alpha \rightarrow \beta) &\text{ iff } \models_w \alpha \text{ and } \models_w \sim\beta. \end{aligned}$$

Now, if we add the operator  $K$  to  $\mathcal{L}_{\text{IPL}\sim}$ , the only definition I can see that is faithful to Definition 5 is the following:

$$(19) \quad \begin{aligned} \models_w K\alpha &\text{ iff } \models_w \alpha \\ \models_w \sim K\alpha &\text{ iff, for some } w' \geq w, \models_{w'} \sim\alpha. \end{aligned}$$

Call any Kripke model for  $\mathcal{L}_{\text{IPL}\sim}$  in which these clauses hold a model for  $\mathcal{L}_{\text{IPL}\sim, K}$ . Consider now the following model  $\mathcal{M}$  of  $\mathcal{L}_{\text{IPL}\sim, K}$ , where  $V(g, w)$  is undefined,  $V(g, w_1) = 0$  and  $V(g, w_2) = 1$ :



$\mathcal{M}$  falsifies at  $w$  all the following formulas:

$$\begin{aligned} \sim K g &\rightarrow \neg K g \\ \sim K g &\rightarrow \sim g \\ \sim K g &\rightarrow \neg g \\ \sim K \neg g &\rightarrow \neg\neg g. \end{aligned}$$

On the other hand,  $\models_w \sim K g \wedge \sim K \neg g$ ; but the constraint of monotonicity is not met: not  $\models_{w_1} \sim K g$  and not  $\models_{w_2} \sim K \neg g$ .

### 3 Neo-Verificationist Approaches

Does the paradox of knowability threaten the neo-verificationist, who normally equates truth with knowability rather than with knowledge?

Let us observe, first of all, that, even within the intuitionistic conceptual framework, it would be possible to suggest a definition of truth different from the one given above:

**Definition 6**  $\text{TR } \alpha =_{\text{def.}} \exists \sigma (\text{proves}(\sigma, \alpha))$ .

Of course, if this definition is proposed within the intuitionistic conceptual framework, the metalinguistic existential quantifier is to be understood intuitionistically: a proof of  $\exists \sigma (\text{proves}(\sigma, \alpha))$  is a procedure  $p$  whose execution yields, after a finite time, a pair  $\langle \sigma, \pi \rangle$ , where  $\sigma$  is a construction<sup>30</sup> and  $\pi$  is a proof of “ $\sigma$  proves  $\alpha$ ”.

It is easy to see that Definition 6 is materially adequate. Define the following function  $f$ : if  $\sigma$  is a proof of  $\alpha$ ,  $f(\sigma)$  is the following procedure  $p$ : (i) take  $\sigma$ ; (ii) effect the observation  $\pi$  that  $\sigma$  proves  $\alpha$ ; (iii) construct the pair  $\langle \sigma, \pi \rangle$ . Since proofs are epistemically transparent, the observation  $\pi$  terminates after a finite time, and the pair  $\langle \sigma, \pi \rangle$  is therefore a proof of  $\exists \sigma (\text{proves}(\sigma, \alpha))$ , i.e. of  $\text{TR } \alpha$ . Conversely, define the following function  $g$ : if  $p$  is a procedure whose execution yields, after a finite time, a pair  $\langle \sigma, \pi \rangle$ , where  $\sigma$  is a construction and  $\pi$  is a proof of “ $\sigma$  proves  $\alpha$ ”, then  $g(p)$  is  $\sigma$ ; since proof observation is factive,  $\sigma$  is a proof of  $\alpha$ .

Definitions 4 and 6 are not extensionally equivalent. Consider the sentence “ $\text{Prime}(n)$ ”, where  $n$  is some very large natural number, and suppose that the primality test has never been applied to  $n$ . Then one of the two statements “ $\text{Prime}(n)$ ” and “ $\neg \text{Prime}(n)$ ” is true, according to Definition 6, since (i) we know the primality test, which is a procedure with the required properties, and (ii) we know that the primality test, if it were applied, would answer either that  $n$  is prime or that  $n$  is not prime. When truth is defined according to Definition 6, the truth of  $\alpha$  is not a cognitive state, but empirical accessibility to a cognitive state which is a proof of  $\alpha$ . On the other hand, neither “ $\text{Prime}(n)$ ” nor “ $\neg \text{Prime}(n)$ ” is true, according to Definition 4, since we have a proof of neither statement, owing to the fact that the primality test has not been applied to  $n$ . Hence, according to Definition 6 there are statements that are true although they are not known now, and possibly not even in the future; while according to Definition 4 there are no statements of this kind: there are only unknown statements waiting to be *made* true (i.e. known) by our activity of proving mathematical statements or coming to know empirical statements.

The essential point to notice in this connection is that the following formula is intuitionistically valid:

$$(20) \quad \exists \sigma (\text{proves}(\sigma, \alpha)) \leftrightarrow \text{K } \alpha,$$

<sup>30</sup>Throughout the present paper I adhere to the intuitionistic idea that proofs (of mathematical statements) belong to a domain of mental constructions. As a matter of fact, I find it much more appropriate to conceive proofs as belonging to the category of cognitive states; on this point see [24], §2.



since both subformulas are equivalent to the same formula  $\alpha$ . How is this possible? The validity of (20) puts dramatically into evidence a peculiarity of intuitionistic logic which deserves being stressed. Suppose that the procedure  $p$  described above, were it applied, would give as a result that  $n$  is prime. According to Definition 6, what determines the truth of “Prime( $n$ )” before the execution of  $p$  is a mere fact (if it is a fact): the fact that the execution of  $p$  will give as a result a proof of “Prime( $n$ )”. Now, the essential characteristic of intuitionistic logic, as Heyting conceives it, is its being a *logique du savoir*, opposed to classical logic as a *logique de l'être*<sup>31</sup>; this entails that the intuitionistic meaning of the logical constants, implication in particular, must be explained in terms of cognitive states instead of facts and relations between facts.<sup>32</sup> Hence, the mere fact that “Prime( $n$ )” is true before the execution of  $p$  plays no role in determining the assertibility or the non-assertibility of any formula of intuitionistic logic; in particular, it does not conflict with the validity of (20).

As a consequence, if one wanted to define a notion of intuitionistic truth by means of Definition 6 instead of 4, one would face a dilemma: either to accept (20), whose validity follows from the fact that the biconditional is read intuitionistically, giving up the possibility of expressing the fact that there are true but unknown statements; or to insist that there are intuitionistically true but unknown statements, giving up the intuitionistic reading of the logical constants occurring in the semantic metalanguage.

The moral drawn from this dilemma by the realist is clear: there are statements that are intuitionistically true but unknown; hence, as shown by the paradox, there are also statements that are intuitionistically true but unknowable; therefore (K) must be rejected. Equally clear is the moral drawn by the intuitionist: both linguistic and metalinguistic logical constants must be read intuitionistically, hence (20) is valid, and the notion of truth defined by Definition 6 either is to be rejected, or has intuitive consequences that cannot be expressed in an intuitionistic language. There is, however, a third answer that can be, and has been, proposed—an answer I should call “hybrid”: it consists in defining truth by Definition 6, in insisting that there are intuitionistically true but unknown statements, in giving up the intuitionistic reading of the metalinguistic logical constants, adopting for them a classical reading, and in accepting (K). This position is instantiated by whoever accepts (K) rejecting at the same time (RAR); for the reason why (RAR) is judged unacceptable can be only that it is understood as expressing the thought that every  $\alpha$  is either false or known, i.e. is understood on the basis of the classical reading of the implication occurring in its formalization.

---

<sup>31</sup> «Heyting [10] has opposed intuitionistic logic as the logic of knowledge (*logique du savoir*) to classical logic of existence (*logique de l'être*).» ([11], p. 107).

<sup>32</sup>This is the content of what Heyting calls *principle of positivity*: «Every mathematical or logical theorem must express the result of a mathematical construction» ([11], p. 108. See also [10], p. 231). In the case of implication, in particular, Heyting holds that within classical logic «il n'y a pas de place pour une implication proprement dite, car chaque proposition est vraie ou fausse, et on ne conçoit pas comment sa vérité pourrait dépendre de celle d'autres propositions.» [10, p. 226] On the contrary, «il est tout naturel que la démonstration d'une proposition dépende de la démonstration d'une autre proposition.» (p. 233).

Among the supporters of the hybrid position there are many neo-verificationists, in my opinion. Be it as it may, it seems to me that this position incurs a paradox strictly analogous to the paradox of knowability. Assume

$$(21) \quad q \wedge \neg \mathbf{K} q;$$

then, by Definition 6, there is a proof of  $q \wedge \neg \mathbf{K} q$ ; let's call such proof  $\sigma$ ; then

$$(22) \quad \text{proves}(\sigma, (q \wedge \neg \mathbf{K} q));$$

by the definition of proof of a conjunction,<sup>33</sup>

$$(23) \quad \sigma = \langle \sigma_1, \sigma_2 \rangle, \text{ where } (\text{proves}(\sigma_1, q) \ \& \ \text{proves}(\sigma_2, \neg \mathbf{K} q)).$$

Since  $\sigma_1$  proves  $q$ , and proofs are epistemically transparent, it is possible to perform the observation  $\sigma_3$  that  $\sigma_1$  proves  $q$ , and this observation is a proof of  $\mathbf{K} p$ ; then

$$(24) \quad \text{proves}(\sigma_3, \mathbf{K} q);$$

if we now construct the pair  $\sigma' = \langle \sigma_3, \sigma_2 \rangle$ , we have that

$$(25) \quad \text{proves}(\sigma', (\mathbf{K} q \wedge \neg \mathbf{K} q)),$$

hence

$$(26) \quad \Sigma \sigma (\text{proves}(\sigma, \perp));$$

on the other hand, the meaning of  $\perp$  is characterized by saying that there is no proof of  $\perp$ , hence the formula

$$(27) \quad \neg \Sigma \sigma (\text{proves}(\sigma, \perp))$$

is assertible: a contradiction. Therefore,

$$(28) \quad \neg (q \wedge \neg \mathbf{K} q),$$

from which

$$(29) \quad q \supset \mathbf{K} q.$$

---

<sup>33</sup>Cesare Cozzo argues, in [1], p. 76, that the existence of  $\sigma_1$  and  $\sigma_2$  is not a contradiction because the existence of  $\sigma_2$  does not imply that there is a not proof of  $q$ . This may be conceded, but it does not solve the paradox, as the next steps show.

A possible way out consists in giving up the intuitionistic idea that proofs are epistemically transparent; in this way the step from (23) to (24) is blocked. But the price to pay is very high: as proof, or more generally verification, is the key-notion of a neo-verificationist theory of meaning, the non-transparency of proofs/verifications would create the same difficulties the neo-verificationists impute to the realist theory of meaning because of the non-transparency of truth-conditions (essentially, the non-satisfiability of the manifestability requirement imposed onto knowledge of meaning).

Another way out has been proposed by Dummett. As a matter of fact, Dummett has tackled the paradox in two papers,<sup>34</sup> suggesting two different answers; since the former has been explicitly withdrawn by him,<sup>35</sup> I will consider only the latter. Dummett's solution consists in accepting

$$(30) \quad \alpha \rightarrow \neg\neg K \alpha,$$

rejecting at the same time (2). This is legitimated, firstly, by the fact that only (30), not (2), follows intuitionistically from (1); secondly, by the fact that, if one reads negation intuitionistically,

‘ $\neg\neg K \alpha$ ’ means ‘There is an obstacle in principle to our being able to deny that  $\alpha$  will ever be known’, in other words ‘The possibility that  $\alpha$  will come to be known always remains open’<sup>36</sup>

—which is precisely what the verificationist believes to hold good for every true  $\alpha$ . Dummett does not explain why (2) should be rejected; he only remarks that what (2) says is «contrary to our strong intuition» (p. 51). As I remarked at the beginning, what (2) says is not contrary to our intuition if (2) is read intuitionistically; on the contrary, it certainly *is* contrary to our intuition if what it says is that either the fact that  $\alpha$  does not obtain, or the fact that  $\alpha$  is (or will ever be) known obtains; but this is precisely the classical reading of the implication occurring in (2). Hence, Dummett is reading classically the implication in (30), intuitionistically the double negation. Such a hybrid reading is not justified; as a consequence, Dummett's solution seems quite *ad hoc*.

## 4 How Is a Rational Discussion Possible?

One essential ingredient of the solution I have proposed is the remark that, when the logical constants are understood intuitionistically, the formalization (2') of (RAR) becomes perfectly acceptable. On the other hand, when the logical constants are

---

<sup>34</sup>References [6, 8].

<sup>35</sup>«I do not stand by the resolution of this paradox I proposed in “Victor's Error”, a piece I wrote in a mood of irritation with the paradox of knowability.» [7, p. 348].

<sup>36</sup>Reference [8], p. 52.

understood classically, (2) is utterly unacceptable. This situation is far from surprising; on the contrary, it illustrates a general truth reminded above: the classical meaning of the logical constants is deeply different from their intuitionistic meaning. Consider for instance the schema “ $\alpha$  or not  $\alpha$ ”: classically understood (i.e., formalized as  $\alpha + \neg\alpha$ ) it expresses the intuitively true principle that every proposition is either true or false, (*intuitively* true because our common-sense or pre-theoretic intuitions about the world are predominantly realistic) whereas intuitionistically understood (i.e. formalized as  $\alpha \vee \neg\alpha$ ) it expresses the intuitively false principle that every proposition is decidable in the sense that there is either a proof or a refutation of it.

However, this situation generates a serious problem: the problem whether a rational discussion between a supporter of classical logic and a supporter of intuitionistic logic is possible at all. How is it possible that there is real disagreement or real agreement between them, given that both disagreement and agreement about a principle presuppose that the same meaning is assigned to it by both parties, while, as we have just seen, the meaning of one and the same formula drastically changes across classical and intuitionistic readings?

It seems to me that there are at least two alternative strategies to tackle the problem. The first consists in placing the discussion between the two parties *before* the formalization of the intuitive notions (as the logical constants, the notion of truth, and so on) into a formal language. The discussion, in this case, concerns questions like the following:

- (i) Which intuitive notions should be formalized? For instance: inclusive or exclusive disjunction? Which notion of implication? Which notion of truth?
- (ii) Which intuitive notion should be chosen as the key-notion of the theory of meaning, i.e. as the notion in terms of which the meaning of the expressions of the formal language (in particular of the logical constants) is to be characterized? For instance: (bivalent) truth (as the realist claims), or knowability/existence of a proof (as the neo-verificationist claims), or knowledge/actual proof (as the intuitionist claims)?

In this case the problem can be solved, *provided that* each party accepts the *intelligibility* of the key-notion adopted by the other party; for only in this case a rational discussion is possible: the same intuitive notions are accessible to both parties, and the disagreement concerns the legitimacy, the adequacy, the fruitfulness, etc. of adopting one notion or another as the key-notion. From this standpoint, Brouwer’s idea that such classical notions as bivalent truth or actual infinity are unintelligible is to be abandoned, in favor of a slightly different claim: that those classical notions, precisely because they are intelligible, turn out to be incapable to play the foundational role the classicist gives them. Of course, such a claim should be motivated by a rational argument; which means that a rational discussion would be possible.

The second strategy consists in placing the discussion between the two parties *after* the formalization of the intuitive notions. In this case the problem of course arises, owing to the fact that the choice of different key-notions for the theory of meaning induces differences in the meaning of the logical constants. However, there may be tactics to solve it.

I hold the first alternative is better, but I have not an a priori argument; I will argue for my thesis by considering what seems a very plausible tactics and explaining why, in my opinion, it is not viable.

The tactics is based on the idea of translating one logic into the other, analogously to the case of the translation of a language into another. As a matter of fact, there are several so-called ‘translations’, both of classical logic/mathematics into intuitionistic logic/mathematics—the so-called negative translations (by Kolmogorov, Gödel, Gentzen, Kuroda and others); and of intuitionistic logic/mathematics into extensions of classical logic/mathematics (Shapiro, Horsten, Artemov). I shall not enter here into a detailed discussion of this tactics. I want only to stress an obvious fact: that the so-called ‘translations’ are not translations at all. A translation, in general, must be correct, and it is correct if it is meaning-preserving, i.e. if, for every expression  $E$  of  $\mathcal{L}$  (the language to be translated), its translation  $\text{Tr}(E)$  into  $\mathcal{L}'$  has the same meaning as  $E$  (whatever meaning is). But there is no reason to believe that the ‘translations’ mentioned above are meaning-preserving. Consider for instance the BHK clause for implication; Shapiro himself admits that the notion of “transformations of proofs” cannot be captured in the language of Epistemic Arithmetic, and Smoryński has observed that the ‘translation’ of intuitionistic logic into epistemic logic «does not capture the full flavor of talk about methods» (p. 1497).<sup>37</sup> To make another example, Kuroda’s negative translation is based on a simple idea: that intuitionistic double negation is a sort of ‘equivalent’ of classical truth; this is surely true if one aims at a faithful ‘immersion’ of classical logic into intuitionistic logic (i.e. at a representation preserving theoremhood), but not if one aims at a genuine translation, for the classical truth of  $\alpha$ , expressed by its occurrence within any formula, is something very different from the existence of an obstacle in principle to our being able to deny that  $\alpha$ , expressed by  $\neg\neg\alpha$ . Moreover, there seems to be a conceptual reason for the impossibility of a genuine translation of one logic into another: on the one hand, a translation is correct only if it is meaning-preserving; on the other hand, classical logic explains the meaning of the logical constants in terms of a notion (bivalent truth) the intuitionist considers unintelligible or illegitimate, and also the converse is true (the classicist finds mysterious the intuitionistic notion of general method or effective function): so it seems unlikely that one of them finds in his own language an expression with the same meaning of an expression of the other’s language.

## 5 Conclusion

The Paradox of Knowability is a paradox if the logical constants occurring in its formalization are understood according to the realist explanation of their meaning; but in a discussion between realists and anti-realists one cannot assume that anti-realists understand them in this way, for the paradox is intended to be an argument by which the former try to convince the latter to abandon their views on the meaning

---

<sup>37</sup>Reference [13], p. 9.

of the logical constants, and such an argument cannot be convincing if, in order to be formulated, it requires anti-realists to give up preventively their views. Vice versa, the paradox completely vanishes when the logical constants occurring in its formalization are understood according to the BHK explanation, since it is now necessary to distinguish two notions of truth: internal intuitionistic truth, which coincides with knowledge, and intuitive truth, essentially consisting in correspondence to external reality; in the former sense it is obvious that every truth is known, in the latter it is equally obvious—also for the anti-realist—that not every truth is known, and also that not every truth is knowable. From this point of view, the view of the paradox as an argument against anti-realism is the result of a wrong way of conceiving the rules of a rational discussion between classicist/realist and intuitionist/anti-realist.

In conclusion, the Paradox of Knowability leaves the debate between realists and anti-realists at the same point it was before its discovery. The crucial point of the debate is which notion between truth and evidence should be adopted as the key notion of the theory of meaning, or—if we accept the (in my opinion misleading) idea that meaning is to be explained in any case in terms of truth-conditions—which notion of truth, between bivalent and non-bivalent truth, the theory of meaning should be built on; in this case, the criterion for distinguishing realism from anti-realism cannot be the acceptance or refusal of the intuitive principle (K), but the acceptance or refusal of the principle of bivalence, according to Dummett's original suggestion.

**Acknowledgments** I am indebted to Julien Murzi and Luca Tranchini for helpful comments on earlier versions of this paper. The work reported here was supported by the MIUR fund No. 20107738C5\_002.

**Open Access** This chapter is distributed under the terms of the Creative Commons Attribution Noncommercial License, which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

1. Cozzo, C.: What can we learn from the paradox of knowability? *Topoi* **13**(2), 71–78 (1994)
2. Dummett, M.: Realism. In: [4], pp. 145–165. Paper read at the Oxford Philosophical Society (1963)
3. Dummett, M.: *Elements of Intuitionism*. Clarendon Press, Oxford (1977)
4. Dummett, M.: *Truth and Other Enigmas*. Duckworth, London (1978)
5. Dummett, M.: *The Logical Basis of Metaphysics*. Duckworth, London (1991)
6. Dummett, M.: Victor's error. *Analysis* **61**, 1–2 (2001)
7. Dummett, M.: Reply to Wolfgang Künne. In: Auxier, R.E., Hahn, L.E. (eds.) *The Philosophy of Michael Dummett*. The Library of Living Philosophers, pp. 345–350. Open Court, Chicago (2007)
8. Dummett, M.: Fitch's paradox of knowability. In: [18], pp. 51–52 (2009)
9. Gurevich, Y.: Intuitionistic logic with strong negation. *Studia Logica* **36**(1/2), 49–59 (1977)
10. Heyting, A.: La conception intuitionniste de la logique. *Les études philosophiques* **2**, 226–233 (1956)
11. Heyting, A.: Intuitionism in mathematics. In: Klibansky, R., (ed.) *Philosophy in the Mid-century*. A Survey, pp. 101–115. La Nuova Italia, Firenze (1958)

12. Heyting, A.: On truth in mathematics. Verslag van de plechtige viering van het honderdvijftigjarig bestaan der Koninklijke Nederlandse Akademie van Wetenschappen, pp. 277–279. North Holland, Amsterdam (1958)
13. Horsten, L.: In defense of epistemic arithmetic. *Synthese* **116**, 1–25 (1998)
14. Kripke, S.: Outline of a theory of truth. *J. Philos.* **72**, 690–716 (1975)
15. Milne, P.: Tarski, truth and model theory. In: *Proceedings of the Aristotelian Society*, vol. XCIX, pp. 141–167 (1999)
16. Murzi, J.: Knowability and bivalence: intuitionistic solutions to the paradox of knowability. *Philos. Stud.* **149**(2), 269–281 (2010)
17. Salerno, J.: Knowability Noir: 1945–1963. In: [18], pp. 29–48 (2009)
18. Salerno, J. (ed.): *New Essays on the Knowability Paradox*. Oxford University Press, Oxford (2009)
19. Smoryński, C.A.: Review of *Intensional Mathematics* by Shapiro. *J. Symb. Log.* **56**, 1496–1499 (1991)
20. Tarski, A.: *The Concept of Truth in Formalized Languages*. Logic, Semantics, Metamathematics, pp. 152–278. Clarendon Press, Oxford (1936)
21. Tarski, A.: The semantic conception of truth and the foundations of semantics. *Philos. Phenomenol. Res.* **4**(3), 341–376 (1944)
22. Usberti, G.: Anti-realist truth and truth-recognition. *Topoi* **31**(1), 37–45 (2012)
23. Usberti, G.: Gettier problems, C-justifications, and C-truth-grounds. In: Moriconi, E., Tesconi, L. (eds.) *Second Pisa Colloquium in Logic, Language and Epistemology*, pp. 325–361. ETS, Pisa (2014)
24. Usberti, G.: A notion of C-justification for empirical statements. In: Wansing, H. (ed.) *Dag Prawitz on Proofs and Meaning*, pp. 415–450. Springer, Cham (2015)
25. Williamson, T.: Knowability and constructivism. *Philos. Q.* **38**(153), 422–432 (1988)
26. Williamson, T.: Never say never. *Topoi* **13**(2), 135–145 (1994)
27. Williamson, T.: *Knowledge and Its Limits*. Oxford University Press, Oxford (2000)