Notre Dame Journal of Formal Logic
Volume 35, Number 4, Fall 1994

Reply to Robert Koons

ANIL GUPTA and NUEL BELNAP

We are grateful to Professor Robert Koons for his excellent, and generous, review (henceforth *KR*) of our book *The Revision Theory of Truth* (henceforth *RTT*). Koons provides in *KR* a welcome guide to our *RTT*, and he puts forward objections that deserve serious consideration. In this note we shall respond only to his principal objection.[1] This objection, which is developed on pp. 625–628 of *KR*, calls into question our main thesis. As we argue below, however, the objection is not successful. We should forewarn the reader that this note is not self-contained. It presupposes familiarity with *RTT* (primarily, Chapter 4) and with *KR*.

The main thesis of *RTT* is that truth is a circular concept. We argued that the Tarski biconditionals, read as partial definitions, constitute an intensionally adequate definition of truth. In other words, if T is a predicate defined by the Tarski-style infinitistic definition (1),

$$(1) \quad x \text{ is } T \ =_{\mathrm{Df}} \ (x = \text{“}p\text{”} \text{ and } p) \text{ or } (x = \text{“}q\text{”} \text{ and q}) \text{ or } \ldots,$$

then truth and T have the same signification in all possible worlds.[2] Since T is obviously circular, we concluded that truth is circular also. Koons finds fault with this last step. He writes:

> We can use the Tarski biconditionals to define a new notion, Tarski-wahrheit, which is certainly a circular concept, since the Tarski biconditionals are circular. Gupta and Belnap argue that truth and Tarski-wahrheit are intensionally equivalent. However, is that enough to enable us to conclude that truth itself is a circular concept? [*KR*, pp. 625–626]

Koons goes on to argue that it is not enough. He puts forward a condition, stated in (2) below, that he thinks ought to be satisfied before we can legitimately claim that truth is circular. He believes that this condition cannot be satisfied and concludes that our argument is flawed and our main thesis false.

Putting in a little more detail, Koons's objection is as follows. Koons observes that the intensional equivalence of a definiendum (e.g., "x is a bachelor") to a circular definiens (e.g., "x is either a bachelor or an unmarried male adult") is insufficient to establish any circularity in the definiendum. Furthermore, Koons thinks, this kind

Received May 10, 1995

of intensional equivalence does not establish that the definition is, in his terminology, "philosophically adequate." To establish the circularity of truth, Koons maintains, we need to show something stronger than the intensional equivalence of truth and *T*. We need to show, he thinks, that a circular definition of truth, such as (1), is philosophically adequate.[3] He spells out this demand as follows:

(2) A definition is philosophically adequate just in case, were someone lacking the concept of the definiendum to acquire a new concept through acceptance of the definition as a stipulation, the new concept so acquired would be intensionally equivalent (coextensive in all possible worlds) to the definiendum [*KR*, p. 626].

Koons thinks that a person lacking the concept of truth is unable to understand any indicative sentence and, in particular, any stipulative definition. He concludes, "truth is essentially indefinable. Hence, it is not definable in a circular fashion and it is not …a circular concept [p. 627]." The context makes plain the basis of this conclusion: No philosophically adequate definition of truth is possible.

We wish to make several points in response. First, the objection rests on some strong theses—theses that are either doubtful or unacceptable. For example, it is doubtful whether an understanding of truth is required for an understanding of an arbitrary stipulative definition, as *KR* claims. Also, (2) is unacceptable as a universal account of philosophical adequacy on definitions. For, surely, there is no *one* absolute condition of adequacy on all definitions. Adequacy conditions vary with the philosophical purpose that a definition is meant to serve, as we emphasize in the fourth point below.[4]

Second, we urge that the relevance of conditions such as (2) to the project of *RTT* is highly doubtful. That project was to try to understand the puzzling logical behavior of the concept of truth and to provide some explanation for it. We argued that this behavior can be explained if truth is a circular concept, one whose signification is given by a rule of revision. We urge that the success of *RTT* is to be judged, not through conditions such as (2), but by examining whether the proposed semantics actually illuminates and explains the behavior of truth.

The objection helps to bring out an important point here: the question whether a concept is circular is largely unrelated to questions about the actual or hypothetical origin of the concept. How persons actually acquire concepts—and how they *might* acquire them—is not at present well understood. But this does not hinder useful semantical inquiries into concepts. Suppose we know that a concept **C** is psychologically primitive (i.e., one not acquired through the acceptance of a definition), and even that it is necessarily so. This tells us little about the signification or the intension of **C**. For all we know **C** is two-valued, but it can equally well be indexical, or vague, or circular. Nothing in our present understanding of concept-acquisition entitles us to draw any conclusions about the semantics of **C** from the fact that it is psychologically primitive. It seems to us a better methodology to use prior semantical investigations to evaluate generalizations linking concept-acquisition and semantics than to evaluate semantical theories on the basis of such (presumably, synthetic) generalizations, especially when these generalizations are arrived at *a priori*. If it were to be discovered that truth is necessarily psychologically primitive, then it would be proper to conclude that some psychologically primitive concepts are circular. The discovery would not, by itself, constitute any argument against the semantical views of *RTT*.

Third, with circular definitions certain distinctions that are hardly in view elsewhere become vitally important. One such is the distinction between *intensional equivalence* and *intensional adequacy* of a definiens to a definiendum. With noncircular definitions, the intensional equivalence between a definiendum, "x is G," and a definiens, "$A(x)$," proves that the definition,

$$x \text{ is } G \ =_{\mathrm{Df}} \ A(x),$$

is adequate from the intensional point of view. Not so with circular definitions. For example, the definiendum "x is a bachelor" is intensionally equivalent to itself. But this does not show that the circular definition,

(3) x is a bachelor $=_{\mathrm{Df}}$ x is a bachelor,

is intensionally adequate to "x is a bachelor." For, according to the definition, the statement "Benazir Bhutto is a bachelor" is Truth-Teller-like, whereas actually it is categorically false. Thus, the circular definiens "x is a bachelor" is intensionally equivalent to the definiendum but is not intensionally adequate to it. Another way of putting the point is this. The signification of the circular concept defined by (3) is not the same as that of "bachelor." Hence, the concept defined by (3) is not intensionally equivalent to the concept of bachelor.

Two claims about circular definitions should be clearly distinguished. Let G be an ordinary concept and let (4) be proposed as a definition for it.

(4) x is G $=_{\mathrm{Df}}$ $A(x, G)$.

The two claims are:

(5) If the definiens "$A(x, G)$" contains "G" essentially and is intensionally equivalent to the definiendum "x is G" then G is a circular concept.

(6) If the definiens "$A(x, G)$" contains "G" essentially and is intensionally adequate to the definiendum "x is G" then G is a circular concept.[5]

It is correctly observed in *KR* that (5) is false; the "bachelor" example above establishes this. But this shows nothing about (6), for (5) and (6) do not say the same thing. (5) is false but (6), we think, is true. And it is (6) that we need for our argument. Observe that (6) is equivalent to (7).

(7) If a concept G is intensionally equivalent to a concept H, where H is given an essentially circular stipulative definition, then G is a circular concept.

Since truth and T are agreed to be intensionally equivalent, and since (1) is an essentially circular stipulative definition of T, the desired conclusion follows: truth is a circular concept.

Fourth, *RTT* argues that circular concepts have a unique kind of signification. Circular concepts behave in a distinctive way. They exhibit a distinctive pattern of pathological and nonpathological behavior. This pattern ought to be reflected in, and explained by, their signification. The source of the pattern, according to *RTT*, is the hypothetical character that circular concepts impart to their signification. Since the hypothetical character is associated only with the signification of circular concepts, we think that (7) is plausible. The following analogy might be helpful here. Intensional equivalence with an (essentially) vague concept proves vagueness; intensional

equivalence with an (essentially) partial concept proves partiality; similarly, we think, intensional equivalence with an (essentially) circular concept proves circularity.

Let's note a methodological consequence of (7): one does not need a finegrained, or substantive, analysis of a concept to establish its circularity. All one needs is an analysis that fixes the intension; nothing more. Consequently, an analysis that suffices to prove circularity may well not—and typically will not—explain many important features of the concept. This is precisely what we think happens with the Tarski biconditionals (understood definitionally). The biconditionals fix the intension of truth, but they do not explain many important features of the concept. (See the discussion of the Intension and Implication Theses in *RTT*; see also [1].) Towards the end of *KR*, it is objected that definitions such as (1) do not explain a certain asymmetry of truth and falsehood. We willingly grant this point, but it leaves undisturbed the argument and the main thesis of *RTT*. Certain philosophical purposes require a more substantive account of truth than that found in (1). But to establish the circularity of truth, (1) suffices.

We have discussed in this note only some of the objections in *KR*. There are other important objections in *KR*—objections that deserve serious consideration in any assessment of the revision theory of truth.

NOTES

1. Our responses to some other objections are in *RTT* itself and in a letter we wrote to Koons. Koons has, very generously, included extracts from this letter in his review.

2. See *RTT*, Section IV of Chapter 4 and Sections 6A–6C of Chapter 6. For an explanation of our use of "signification," see pp. 30–31 of *RTT*.

3. At an earlier point in his review, Koons formulates a different demand. He requires that we show that "the concept of truth could be introduced for the first time to a cognitive agent by means of the deployment of [a circular] definition [p. 626]." Our view of this demand is stated in the second of our four points.

4. Furthermore, (2) implies, contrary to the message of *KR*, that every intensionally adequate definition is philosophically adequate. Note that the righthand side in (2) is of the form, "Were someone to do φ then p." That is, it is a generalized counterfactual conditional. The consequent of this conditional is bound to hold for an intensionally adequate definition. Hence, by the standard stated in (2), such a definition is going to be philosophically adequate. Condition (2) rules, therefore, that (1) is a philosophically adequate definition of truth. The arguments of *KR* show at most that, with truth, the generalized conditional contained in (2) will have a necessarily false antecedent. But this does not show that the conditional itself is false.

5. Note that the qualification "essentially" is needed in these claims to rule out trivial counterexamples such as the following:

   (i) x is H $=_{\mathrm{Df}}$ x is an even number & x is prime & (x is either H or non-H).
   (ii) x is J $=_{\mathrm{Df}}$ either [$x = a$ & (a or b is J)] or [$x = b$ & neither a nor b is J].

   H and J are defined circularly in (i) and (ii), but this does not establish that H and J express circular concepts. The reason is that circularity, though present, is eliminable and inessential in (i) and (ii). (See *RTT*, 5A.11, for a discussion of Example (ii).)

REFERENCES

[1] Gupta, A., "A critique of deflationism," *Philosophical Topics*, vol. 21 (1993), pp. 57–81.

[2] Gupta, A., and N. Belnap, *The Revision Theory of Truth*, MIT Press, Cambridge, 1993.

[3] Koons, R., Book Review, *Notre Dame Journal of Formal Logic*, vol. 35 (1994), pp. 606–631.

Anil Gupta
Department of Philosophy
Indiana University
Bloomington, IN 47405
email: agupta@indiana.edu

Nuel Belnap
Department of Philosophy
University of Pittsburgh
Pittsburgh, PA 15260
email: belnap@pitt.edu