THE JOURNAL OF PHILOSOPHY
VOLUME LXV, NO. 17, SEPTEMBER 5, 1968

LOGICAL TRUTH REVISITED

THIRTY-TWO years ago W. V. Quine proposed a definition of 'logical truth' that has been widely repeated and reprinted.[1] Quine himself seems to have recognized that this definition is wrong in detail; in section I we eliminate this fault. What has perhaps been less widely observed is that, in abandoning the model-theoretic account of logical truth in favor of a "substitutional" account, Quine's definition swells the ranks of the logical truths and makes the classification of a sentence as a logical truth dependent both on the interpretation of its extralogical words and on the extralogical vocabulary available in the language.

I

Quine's definition makes use of three preliminary notions.

> (I) An expression will be said to occur *vacuously* in a given statement if its replacement therein by any and every other grammatically admissible expression leaves the truth or falsehood of the statement unchanged (73).

'Replacement' here is to be understood in the sense of "replacement in each of its occurrences by the same expression." Thus in

(1) Paul is plump v $\sim$ Paul is plump

'Paul is plump' occurs vacuously.

> (II) . . . for any statement containing some expressions vacuously there is a class of statements, describable as *vacuous variants* of the

[1] In "Truth by Convention" first published in O. H. Lee, ed., *Philosophical Essays for A. N. Whitehead* (New York: Longmans, 1936), and reprinted in H. Feigl and W. Sellars, eds., *Readings in Philosophical Analysis* (New York: Appleton, 1949), also in P. Benacerraf and H. Putnam, eds., *Philosophy of Mathematics* (Englewood Cliffs, N. J.: Prentice-Hall, 1964), and in Quine's *The Ways of Paradox* (New York: Random House, 1966). In the present paper all references to Quine's articles will be to this last-mentioned collection.

given statement, which are like it in point of truth or falsehood, like it also in point of a certain skeleton of symbolic make-up, but diverse in exhibiting all grammatically possible variations upon the vacuous constituents of the given statement (73).

Finally,

(III) An expression will be said to occur *essentially* in a statement if it occurs in all the vacuous variants of the statement, i.e., if it forms part of the aforementioned skeleton (73).

'Logical truth' is defined relative to a previously specified class of logical words. "The logical truths . . . are those true sentences which involve only logical words essentially" (103).[2]

As it stands, the definition fails. For, as Quine has noted, provision must be made in the definition of 'vacuous occurrence' for varying two or more expressions at a time.[3] Failing this,

(2) If some men are angels, some animals are angels

becomes a logical truth, since 'men', 'angels', and 'animals' occur vacuously and only logical words occur essentially. Indeed, we need not resort to quantificational logic to demonstrate the inadequacy of Quine's formulation. Examples in sentential logic are abundant. Consider

(3) Snow is white v grass is green

No grammatically acceptable substitute for 'snow is white' will turn (3) into a falsehood. Thus, 'snow is white' occurs vacuously. Similarly, 'grass is green' occurs vacuously. Common to all vacuous variants is v alone.[4] Thus, on Quine's account, (3) is a logical truth, as is any disjunction both of whose disjuncts are true.[5]

There is a variety of steps we could take to bring Quine's definition in line with his intentions. Perhaps the simplest uses the notion of an atomic expression. This is definable with the apparatus we already have on hand. An expression E is an *atomic expression* if and only if

[2] This is from Quine's "Carnap and Logical Truth" reprinted in his *The Ways of Paradox*. The same definition is clearly intended, though not so concisely expressed, in "Truth by Convention."

[3] "Carnap and Logical Truth," p. 103. Quine attributes the example below to John Myhill and Benson Mates.

[4] We will follow the convention of using quote names for English expressions and of using abbreviations and symbolic expressions autonymously.

[5] For this example we are indebted to John R. Immerwahr.

for all expressions E', if E' occurs within E then $E' = E$.[6] Let us say that an n-tuple of distinct atomic expressions $(e_1, e_2, \ldots, e_n)$ *occurs in* a statement S if each member of the n-tuple occurs in S. To replace $(e_1, e_2, \ldots, e_n)$ in S by another n-tuple of (possibly non-atomic) expressions $(E_1, E_2, \ldots, E_n)$ is to replace each occurrence of each e_i by the corresponding E_i. Then we say that an n-tuple of distinct atomic expressions $(e_1, e_2, \ldots, e_n)$ *occurs vacuously* in a statement S if and only if it occurs in S and, for all n-tuples of expressions $(E_1, E_2, \ldots, E_n)$, if the result S' of replacing $(e_1, e_2, \ldots, e_n)$ by $(E_1, E_2, \ldots, E_n)$ is a grammatically admissible statement, then S' has the same truth value as S.

Now let $d_1, d_2, \ldots, d_m$ be a complete nonrepeating list of the non-logical atomic expressions in a statement S in (say) the order of their first occurrence. Then we may say that S is a *logical truth* if an only if

(a) S is true, and

(b) $(d_1, d_2, \ldots, d_m)$ occurs vacuously in S.

This revised definition seems to capture Quine's intention. It classifies as logical truths those truths whose nonlogical words "can be varied at will without engendering falsity," where "the phrase 'can be varied at will' . . . is understood to provide for varying the words not only singly but also two or more at a time."[7]

Thus (2) is excluded from the logical truths since ('men', 'angels', 'animals') occurs nonvacuously within it. And (3) is excluded since ('snow is white', 'grass is green') occurs nonvacuously. On the other hand,

(4) Snow is white $\supset$ ($\sim$ snow is white $\supset$ grass is green)

falls among the logical truths, since ('snow is white', 'grass is green') occurs vacuously. All this is as it should be.

II

Yet our definition as revised is still inadequate. It does not coincide with the more familiar model-theoretic definition of 'logical truth'. Rather it inflates the class of logical truths in ways curiously counter-intuitive. Further, it abandons the principle that classification of a sentence as a logical truth should depend only on the interpretation and arrangement of its logical components. For our revised definition

[6] For example, in a language with the form of the sentential calculus, the atomic expressions include the atomic sentences and the logical connectives. In a quantificational language the quantifiers, variables, and constants for individuals and relations are also atomic.

[7] "Carnap and Logical Truth," p. 103.

is sensitive both to the interpretation of a sentence's extralogical components and to the wealth of a language's extralogical apparatus.

Let us first see that Quine's account, as modified, diverges from the model-theoretic account according to which logical truths are sentences valid in every nonempty domain. It is easily seen that in any language[8] a logically valid sentence satisfies our definition. For to be valid a sentence must be true on all reinterpretations of its nonlogical parts in any nonempty domain. Thus, in particular, it will be true on those reinterpretations which assign to atomic predicates (and names, if any) the extensions of predicates formulable in the language. But the valid sentences do not exhaust the sentences that satisfy our definition. For consider a language L_1 containing all the apparatus of quantificational logic with identity and containing just one nonlogical atomic predicate, the one-place predicate P. Suppose further that at least two things are P and that at least two things are not P. (For concreteness we may think of P as an abbreviation of 'is red'.) Then

$$(5)\quad (\exists x)(\exists y)(Px \cdot Py \cdot x \neq y) \vee (x)Px \vee (x) \sim Px$$

is true. The only nonlogical atomic expression in (5) is P. Does P occur vacuously in (5)? A moment's reflection will reveal that it does. For each expression that may grammatically replace P must have as its extension one of four classes: the class of things that are P, the class of things that are not P, the null class, and the universal class. And replacing P with a predicate having any one of these four extensions leaves the truth of (5) unaltered. Thus, on our revised account, (5) is to be classified as a logical truth, although it is clearly not valid in every nonempty domain.

Nor is the problem restricted to languages as impoverished as L_1. Suppose the atomic predicates of a language L_2, identical in logical structure with L_1, consist of finitely many one-place predicates P_1, P_2, . . ., P_n. Then, regardless of what their extensions are, at most 2^{2^n} distinct subsets of the domain are extensions of formulas of L_2. Therefore, for some integer m, none of these extensions has exactly m elements. Hence, the sentence in L_2 that expresses the fact that the number of objects that are P_1 is not exactly m, i.e.,

$$(6)\quad \sim(\exists x_1) \cdots (\exists x_m)[P_1x_1 \cdot \cdots \cdot P_1x_m \cdot (x_1 \neq x_2) \cdot \cdots \cdot (x_{m-1} \neq x_m) \cdot (x_{m+1})[P_1x_{m+1} \supset (x_{m+1} = x_1 \vee \cdots \vee x_{m+1} = x_m)]]$$

would turn out on Quine's account to be logically true!

[8] We use the term 'language' for a grammar together with a scheme of interpretation. A sentence *occurs* in a language if it is well-formed according to the rules of the grammar of that language.

Note that if the language is sufficiently strong—in particular, strong enough to develop first-order arithmetic—then our revised definition does pick out exactly the valid sentences. Briefly, this is so because any sentence S that is not universally valid can be falsified under an arithmetically definable interpretation. Thus the predicates involved in this interpretation can be defined in the language, and substituting these defining expressions for the corresponding atomic predicates of S yields a false sentence. So S fails to satisfy our revised definition.

At this point let us digress for a moment to warn against an untenable objection. It might be held that the notion of logical word Quine has in mind is to include logical operators and connectives such as $\exists$ and $\supset$, constants such as $=$, and associated punctuation marks such as (,], etc., but is not to include variables. On this interpretation the nonlogical atomic expressions of (5) would be x, y, and P, and our argument would collapse, since (x, y, P) does not occur vacuously in (5)—the substitution of (x, x, P) changes the truth value. However, apart from the fact that examples can be constructed that are immune to this objection,[9] it would render the definition of 'logical truth' too restrictive. Consider

$$(7)\quad (\exists x)(\exists y)[Qx \cdot \sim Qy] \equiv (\exists x)Qx \cdot (\exists y) \sim Qy$$

Certainly (7) has the form of a logical truth, but the triple (x, y, Q) does not occur vacuously in it; for again the substitution of (x, x, Q) generates a falsehood.

Thus far we have shown that Quine's "substitutional" definition bloats the extension of 'logical truth' well beyond that specified by the more familiar model-theoretic account. The sentences on which the two characterizations diverge are, by our lights, unwelcome additions to the class of logical truths. Here we must conclude that Quine's intuitions differ from our own, for the failure of the two definitions to coincide is surely not news to him.[10]

We do not, however, rest our charge of inadequacy merely on the intuitive oddness of classifying such sentences as (5) and (6) above as logical truths. For central to the notion of logical truth as we understand it is the principle already cited, that classification of a

[9] One such is $(x)Px \vee (x) \sim Px$ in a quantificational language in which P, a predicate true of everything, is the only nonlogical atomic predicate.

[10] Cf. Quine's *Elementary Logic* (rev. ed., Cambridge, Mass.: Harvard, 1966), p. vi. The divergence of the two definitions was perhaps first noted by Tarski. Cf. "On the Concept of Logical Consequence" in A. Tarski, *Logic, Semantics, and Metamathematics* (Oxford: Clarendon, 1956). The article appeared in Polish in 1936.

sentence as a logical truth should depend only on the interpretation and arrangement of its logical components. Quine's substitutional account violates this principle in two directions.

First, the definition is sensitive to the interpretation of the non-logical as well as the logical components of a sentence. Consider the language L_3 with the same grammar and logical structure as L_1, differing from L_1 only in that the symbol P is now interpreted as a predicate true of exactly one object. Then the sentence (5) occurs in both languages and, on Quine's account, is a logical truth in L_1 but not in L_3.

Second, the definition is sensitive to the richness of the extralogical apparatus of the language. Consider the language L_4 obtained by adding to L_1 a second one-place predicate Q which is interpreted as a predicate true of exactly one object. Then L_1 is a sublanguage of L_4, and (5) is a logical truth in L_1 but not in L_4.

It is hard to see the appeal of a notion of logical truth so curiously ill behaved.[11]

PETER G. HINMAN
JAEGWON KIM
STEPHEN P. STICH

The University of Michigan

COMMENTS AND CRITICISM

ON INTERPRETING DOXASTIC LOGIC

IN the logic of belief ("doxastic logic") developed in Hintikka's *Knowledge and Belief,** the statement that a person b believes the statement p is "indefensible" if p itself is logically inconsistent. Following Hintikka, we represent this statement by the expression '$B_b p$'. The result mentioned follows by the condition (C.B*). This condition provides that if "$B_b p$" is a member of a model set μ and if μ^* is a doxastic alternative to μ (with respect to b) in some model system, then p is a member of μ^*. Since model sets represent (partial) descriptions of possible states of affairs, they cannot contain logically inconsistent statements. Thus, (C.B*) entails that the statement "$B_b p$" is not a member of a model set if p is inconsistent. A statement (or set of statements) that cannot be imbedded in a model set is said by Hintikka to be "indefensible." The rules and conditions that constitute doxastic logic determine which sets of statements

[11] We are indebted to Professor Quine and to John R. Wallace for their helpful comments on an earlier draft of this paper.

* Jaakko Hintikka, *Knowledge and Belief* (Ithaca, N.Y.: Cornell, 1962).