

Epistemic Paradox

Author(s): Tyler Burge

Source: *The Journal of Philosophy*, Vol. 81, No. 1 (Jan., 1984), pp. 5-29

Published by: Journal of Philosophy, Inc.

Stable URL: <http://www.jstor.org/stable/2026102>

Accessed: 11-04-2017 02:23 UTC

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://about.jstor.org/terms>



*Journal of Philosophy, Inc.* is collaborating with JSTOR to digitize, preserve and extend access to *The Journal of Philosophy*

---

---

# THE JOURNAL OF PHILOSOPHY

VOLUME LXXXI, NO. 1, JANUARY 1984

---

---

## EPISTEMIC PARADOX\*

**S**EMANTICAL paradox first appeared cloaked in a that-clause. The Liar emerged from an indirect-discourse report of what Epimenides, the Cretan, said. This origin suggests that the paradox might have implications for our understanding not only of truth and other semantical notions, but also of indirect discourse and the propositional attitudes.

Apart from Russell's ramification of the theory of types, most work on paradox in this century has ignored that-clauses. One reason for this neglect is a predominance of mathematical motivations. Attention has centered either on set theory or, following Tarski, on semantical notions like truth, denotation, satisfaction, definition, with little attention to cognitive notions.<sup>1</sup> In fact, most semantical work relevant to the paradoxes has concerned direct discourse. This focus has the advantage of being philosophically related to fewer notions of warm dispute or ill repute. But it has tended further to distract attention from the effect of the paradoxes on propositional-attitude notions.

Another source of distraction has been an argument by Richard Montague that a certain extension of the reasoning in the Liar Paradox demonstrates that expressions like 'is necessary' cannot be counted predicates of sentences or of propositions with logical structure, unless one sacrifices widely accepted axioms of modal logic. The argument can easily be transposed to apply to "factive" propositional attitudes like 'knows'. And several philosophers have concluded that the argument supports the further view that neces-

\*A version of this paper was read at the University of Southern California in Spring 1980. A later version was read at the University of California, Irvine, in Spring 1982. Both occasions occasioned improvements.

<sup>1</sup>Bertrand Russell, "Mathematical Logic as Based on the Theory of Types" in *Logic and Knowledge*, R. C. Marsh, ed. (London: Allen & Unwin, 1956). Alfred Tarski, "The Concept of Truth in Formalized Languages," in *Logic, Semantics, Metamathematics*, J. H. Woodger, trans. (New York: Oxford, 1956).

sity, knowledge, indirect discourse, and propositional attitudes generally must be represented in terms of operators on sentences—or at any rate, not by predicates of anything with something like syntactical structure (sentences, traditional propositions, Fregean thoughts).<sup>2</sup>

A thorough discussion of these issues is impossible here. But it can be said simply that Montague's argument does not establish what it purports to. For the argument unaccountably ignores the possibility that the intuitive (iteration) laws of modal logic might be expressed in a predicational treatment of modality by using a hierarchy of modal predicates—analogueous to the hierarchy used in

<sup>2</sup>“Syntactical Treatments of Modality with Corollaries on Reflexion Principles and Finite Axiomatizability,” *Acta Philosophica Fennica*, xvi (1963): 153–167. This paper extends to necessity the formal result for knowledge established in David Kaplan and Richard Montague, “A Paradox Regained,” *Notre Dame Journal of Formal Logic* 1, 3 (July 1960): 79–90. Montague's philosophical argument has been accepted and applied to other factive propositional attitudes by Richmond H. Thomason in “Indirect Discourse Is Not Quotational,” *Monist*, LX, 3 (July 1977): 340–354. Thomason's argument explicitly assumes that an unqualified truth schema applies to the contents of indirect discourse. Montague and Thomason also assume that, in a system in which necessity (or indirect discourse) is treated as a predicate of sentences, the iteration laws of modal logic (and certain plausible analogs for factive propositional attitudes) will not be expressed in terms of a hierarchy of predicates in a way analogueous to the hierarchy of truth predicates one might invoke to handle the standard Liar. Each of these assumptions is tantamount to begging the question—since the question is whether necessity and the propositional attitudes are to be treated as predicates of sentence-like entities on an analogy to the truth predicate for sentences. For criticisms of Montague, see my “Buridan and Epistemic Paradox,” *Philosophical Studies*, xxxiv, 1 (July 1978): 21–35, note 5; and Brian Skyrms, “An Immaculate Conception of Modality,” this JOURNAL, LXXV, 7 (July 1978): 368–387. The issue raised by Montague's paper is not limited to *syntactical* treatments of modality and indirect discourse. Substantially similar issues arise for treatments that represent these notions as predicates of propositions, where propositions are specified as having structures analogueous to the syntax of sentences (quantification, conjunction, and so forth). There are treatments along this line which are cognizant of the threat of paradox and which represent the relevant predicates hierarchically. Cf. Alonzo Church, “Outline of a Revised Formulation of the Logic of Sense and Denotation,” *Noûs*, vii, 1 (March 1973): 24–33, and viii, 2 (May 1974): 135–156; David Kaplan, *Foundations of Intensional Logic* (dissertation, UCLA, 1964). Neither of these systems seems to me to interpret natural language—for a number of reasons, but fundamentally because of the rather traditional form of their hierarchies. Still, they illustrate how predicate treatments of modality (and in Church's work, belief as well) can respond to Montague-like variants of the Liar. For a more congenial development along these lines, see C. Anthony Anderson, “The Paradox of the Knower,” this JOURNAL, LXXX, 6 (June 1983): 338–355.

I believe not only that Montague's approach has failed to refute predicate approaches, but that its claimed advantage is in fact a defect as an account of our common conceptions of modality, knowledge, and the like. For natural and I think coherent reasoning involving these conceptions occurs in and around the edges of paradox. Montague's approach blocks the paradoxes, but also blocks the opportunity to understand deep features of these conceptions which are revealed in our reasoning about the paradoxes.

treatments of truth as a predicate of sentences. Traditional hierarchical accounts are not without difficulties. But citing difficulties with specific accounts does not constitute a general argument against any approach along traditional lines. Thus Montague's argument boils down to the point that "syntactical" accounts of modality must do something about variants on the Liar paradox. Indeed they must.

In my view, there are overwhelming reasons to regard propositional-attitude expressions as predicates, and indeed as predicates of entities with some analog of logical structure. I will not argue for these views here; I will have to presuppose them. For present purposes I need not take a position on the ontology of the contents of propositional attitudes. As far as our discussion goes, they might be symbols or some more abstract intensional entity, like Fregean thoughts.

I shall be guided by the assumption that the paradoxes are best approached as resources for understanding deep and subtle features of our language and concepts, rather than as symptoms of contradiction or incoherence in them. Insofar as the paradoxes are not resolved, they *are* symptoms of confusion or mistakes in our assumptions about our language and concepts. Since these assumptions have seemed obvious, the paradoxes are a source of theoretical illumination. But I believe that careful attention to intuitive reasoning in particular cases reveals a coherent and interesting pattern in our employment of semantical and propositional-attitude notions.<sup>3</sup>

The best known epistemic paradoxes fall into one of two categories. Either they make use of the predicate 'know' (as do the Kaplan-Montague paradox and the Hangman), or they utilize a predicate like indirect-discourse 'says', 'believes', or 'thinks' together with a truth predicate (as does the Epimenides). The former sort of paradox can, with care, be formulated using highly general quasi-a-priori principles governing knowledge. Or it may depend partly on empirical assumptions. The latter inevitably uses empirical assumptions; these can, however, be made overwhelmingly plausible. In either case, the formulation of paradox utilizes a predicate whose application to an entity entails that entity's truth. One is thus tempted to see epistemic paradoxes as close relatives to the direct-discourse version of the Liar discussed by Tarski. Some have

<sup>3</sup> The present paper is intended to be self-contained. But its methods and even some of its conclusions—particularly the theoretical proposal alluded to at the end of section 1—will be better understood against the background of my work on the semantical paradoxes that arise for direct discourse: "Semantical Paradox," this *JOURNAL*, LXXVI, 4 (April 1979): 169–198; "The Liar Paradox: Tangles and Chains," *Philosophical Studies*, xxxviii, 4 (May 1981): 353–366.

gone further—holding that such paradoxes arise only through *conflating* that-clause locutions with direct-discourse applications of ‘true’.

This further step is doubly mistaken. In the first place, paradox arises for that-clause locutions as well as for direct discourse. As we have noted, the oldest paradoxes are of this form. In section I, I discuss a case that combines the notion of truth with a propositional attitude. I shall show that such cases are not satisfactorily dealt with by claiming that no “proposition” is thought. Paradox-engendering contents are quite thinkable. I shall suggest that such cases are best handled by seeing the truth predicate (or other factive predicate) as undergoing context-dependent shifts of extension.

A second reason why it is mistaken to see epistemic paradox as simply the result of a conflation with the direct-discourse truth-theoretic paradox is that relevantly similar pathology can be produced using nonfactive propositional-attitude predicates and no semantical predicate at all. I shall discuss such a case in section II. Paradoxes of this kind have received little or no attention in the recent literature. I shall argue that they are appropriately treated by construing predicates like ‘believes’, ‘has once accepted (the proposition)’, and so forth as analogous to ‘is true’ in involving shifts of extension, though these shifts do not form the same hierarchy. In section III I draw some semantical morals from the preceding. I conclude in section IV with remarks on more general philosophical issues.

## I

Let us begin with an intriguing paradox derived from an example of A. N. Prior’s.<sup>4</sup> Suppose that Galileo and his inquisitor are returning from a session along a corridor. Galileo is frustrated and thinks his companion a fool. They separate into different rooms just before 6 A.M. Galileo thinks that his companion has retired to room 13 and himself into room 12. And he disgustedly thinks that nothing thought in room 13 at 6 A.M. is true. But in fact, Galileo himself has entered room 13 and the inquisitor has occupied room 14. If we assume that Galileo thought no other thought at 6 A.M.

<sup>4</sup> Cf. “On a Family of Paradoxes” *The Notre Dame Journal of Formal Logic*, II, 1 (January 1961): 16–32. Prior’s paper is exceedingly rich in interesting examples. As Prior notes, one cannot reasonably escape the present problem by insisting that Galileo held a *de re* belief of room 14. Galileo may have held such a belief, but citing it does nothing to diminish the plausibility of his holding the belief ascribed in the argument. A note on terminology: *thoughts* (or *thought contents*) are the oblique referents of that-clauses; no particular ontological status is presupposed. *Notions* are the analogs in thoughts of predicates; again, no particular ontological status is presupposed.

(the case could easily be altered to relax this assumption), then we have the makings of paradox. The only thought thought in room 13 at 6 A.M. is that nothing thought in room 13 at 6 A.M. is true. Is Galileo's thought true or not? Suppose it true. Then nothing thought in room 13 at 6 A.M. is true: in particular, Galileo's thought is not true. Suppose then that it is not true. Then something thought in room 13 at 6 A.M. is true; and this can only be Galileo's thought, there being no other. But now we are apparently caught in a contradiction.

We can schematize the reasoning as follows:

1. The only thought thought in room 13 at 6 A.M. = that nothing thought in room 13 at 6 A.M. is true. (Assumption)
2.  $g$  = that nothing thought in room 13 at 6 A.M. is true. (Abbreviative Definition)
3.  $g$  is true  $\leftrightarrow$  nothing thought in room 13 at 6 A.M. is true. (Assumption)
4. *Suppose*:  $g$  is true.
5. Nothing thought in room 13 at 6 A.M. is true. (By 3, 4)
6.  $g$  is not true. (By 1, 2, 5)
7. Something thought in room 13 at 6 A.M. is true. (By 3, 6)
8.  $g$  is true. (By 1, 2, 7)

The reasoning depends only on the most straightforward, elementary logical laws, applied in what, apart from the paradoxes, would normally be taken to be unproblematic, extensional contexts. The application of the truth predicate does not occur *within* the that-clauses themselves and thus is not bound up with issues involving nonextensionality. The truth predicate is applied to the thoughts indicated by the that-clauses.

There are a number of possible maneuvers to block the contradiction. Some reinterpret the logical expressions (such as 'not') or revise the logical laws (such as substitutivity of identity). But I shall ignore these possibilities. I have discussed them elsewhere in a context that differs but slightly (cf. note 3). The source of the trouble is most plausibly sought either in the initial supposition that, in the context, Galileo thought what I described him as thinking (in step 1), or in the application of the truth schema to Galileo's thought (in step 3).

*Denial of step 1* may take one of two forms. According to the first, we simply deny that Galileo really thought anything in the context. This position can be seen as analogous to a position commonly taken about indirect discourse. Not just any utterance counts as saying something in the indirect-discourse use of 'say'. And there is indeed something peculiar about the expression we

utilized to attribute a thought content to Galileo in the relevant context. For *in effect* the expression evaluates itself as untrue. But such expressions as 'This expression is untrue', seem, in some sense, not to say anything. There is no "proposition" that one can evaluate as true or untrue. The expression attributed to Galileo is similarly empty. Therefore, contrary to first appearances, there is really nothing—no proposition—for Galileo to think. Galileo did not think anything. So goes the reasoning.<sup>5</sup>

Despite the reasoning, the position is unsatisfactory. It seems natural to hold that whatever the consequences, Galileo could, in the context, have thought what I described him as thinking. Whether a perfectly normal person, with ordinary training and a good record in ratiocination, can think anything at all on a given occasion is claimed by the view to depend on empirical contingencies on the other side of a door or (for that matter) the globe. The view would also make whether a person can think anything on a given occasion depend on the whims of another person's thought. This is because of two-person paradoxes. If *A* appears to think at 5:30 A.M. (only) that everything thought by *B* at 6 A.M. is true and *B* thinks at 6 A.M. that there have been dogs, then *A* has thought something. But if *A* appears at 5:30 A.M. to think the same and *B* appears to think at 6 A.M. that everything thought by *A* at 5:30 A.M. is false (or true), then neither *A* nor *B* has thought anything. Thus whether *A* thinks anything at a given time may depend on what someone else thinks later. In fact, the situation is slightly worse. Since 'appear to think' is a predicate that appears formally analogous to 'think' in its capacity to produce paradox, *A* cannot, on the view being considered, even appear to think anything under relevantly adjusted circumstances if *B* does not happen to be right thinking. Motivated by such considerations, let us scrutinize the rationale for the view.

The rationale is greased by the notion of proposition, a notion notorious for its slipperiness. I would be willing to say that, in a certain special sense, Galileo failed to think a "proposition." But one must note what purpose is being served by such a remark. The informal notion of proposition has been pressed into serving three primary purposes: indicating what is evaluated as true or false; characterizing what is meant by a sentence or sentence utterance; and applying to what someone thinks, believes, or states.

Clearly, the first of these purposes is exploited in the rationale I offered a couple of paragraphs back. The question of whether Gali-

<sup>5</sup> This view is maintained by Prior, *op. cit.*; and by William Kneale, "Propositions and Truth in Natural Language," *Mind*, LXXXI, 232 (April 1972): 225-243.

leo's purported thought is true or false seems, in some sense, not to have a good answer. There is nothing for Galileo's purported truth evaluation to base itself upon. In some sense, he did not succeed in evaluating anything as true or not, and, in some sense, he is not related to or characterized by anything that is true or false. These intuitive observations ground our intuition that Galileo was not related to a proposition.

What should we say of the other two purposes that the notion of proposition has been asked to serve? The expression in terms of which we characterized Galileo's thought is perfectly meaningful. We can paraphrase it, reason in terms of it (even in the context), imagine cases in which it would be unproblematically true. So broadly speaking, meaninglessness is not involved in any failure of Galileo to think or express a proposition. This point renders tenuous the analogy to cases in which we refuse to use indirect discourse despite direct-discourse utterance. For we tend to resist such use primarily when the speaker did not utter expressions with conventional meaning or when the speaker did not at all understand the expressions he mouthed. Neither meaninglessness nor failure of understanding is at issue here.

It is, of course, the third role of the notion of proposition which is crucial to evaluating the attempt to deny that Galileo thought what I described him as thinking. Here propositions serve as thought contents. Could Galileo, in the relevant context, think that nothing thought in room 13 at 6 A.M. is true? I think it extremely plausible, indeed obvious, that he could.

Let us explore why. One purpose of thought attributions is to provide a standard whereby we can understand and assess the epistemic viewpoint of other people. The need to understand how Galileo sees things, how he reasons, is no less present in our problem case than in others. We want to ask whether Galileo's thought is justified, how he arrived at his view, what errors led to his adopting it. These questions are not undermined by the unfortunate circumstances Galileo has got himself into.

Another purpose of attributing thoughts is to explain a person's sincere asserted utterances, his actions, and other pieces of his behavior. There is nothing in the context that renders this purpose null and void. The psychological background of Galileo's dispositions to utter words or assert things that are reported in indirect discourse is no different from what it would have been if the inquisitor, not Galileo, had been in room 13. We can still explain Galileo's actions and behavior in terms of what he is thinking. The thought content most useful in giving such explanations may be exactly as I have described it in step 1.

A philosopher may, of course, introduce a sense of 'think' or 'say' according to which Galileo did not think or say anything. But the introduction does nothing to help answer the questions of what Galileo was doing, how he regarded the situation, how his actions are to be explained, and whether they are reasonable—and to answer them in a way that fills descriptive, communicative, explanatory, and evaluative functions of ordinary mentalistic attributions.

The difficulty in squaring this denial of step 1 with the role of propositional-attitude attributions in characterizing someone's epistemic point of view and in explaining his behavior is serious because of the arbitrariness and contingency of the matters on which paradox can depend. If Galileo is in the wrong room or if someone else thinks the wrong thought, then we are suddenly deprived of our ability to apply our scheme of characterization and explanation to Galileo. I think it clear that this result is unacceptable. It is a measure of the isolation of semantics from epistemology and psychology that so many philosophers have embraced it.<sup>6</sup>

A more nearly acceptable way of denying step 1 in our argument takes its cue from F. P. Ramsey's criticism of ramified type theory. On this view, one regards the notion of thinking as systematically ambiguous, or better, as indexical.<sup>7</sup> Thus one might assign the occurrence of 'thought' in the that-clause which indicates Galileo's thought a subscript 1: that nothing thought<sub>1</sub> in room 13 at 6 A.M. is true. But when we judge that Galileo thought this thought, the outer occurrence of 'thought' has a wider extension and a higher subscript; Galileo thought<sub>2</sub> that nothing thought<sub>1</sub> in room 13 at 6 A.M. is true. The idea would be that Galileo did not think<sub>1</sub> anything, but did think<sub>2</sub> something. On this reconstruction, what Galileo

<sup>6</sup> Although the view I have been criticizing may be marginally more plausible as applied to indirect discourse ("says that"), I find this application unacceptable also. If Galileo had sincerely and with understanding asserted Latin words equivalent to the expression we attributed to him, English speakers who understood the words would certainly have reported him in indirect discourse as having said that nothing thought in room 13 at 6 A.M. was true. Of course, one could give a revisionistic rationale against such attributions analogous to the one we discussed earlier—with equally unconvincing results. Galileo's words clearly had import, meaning. The practice of indirect-discourse reporting is fitted to conveying the import of the speaker's words to a home audience. Moreover, his words, given some appropriate interpretation, serve in facilitating our psychological explanations (say, in terms of his thoughts). There is no apparent reason why these needs should not be served in the present instance.

<sup>7</sup> Ramsey, "The Foundations of Mathematics" in *Foundations*, D. H. Mellor ed. (Atlantic Highlands, N.J.: Humanities Press, 1978); Charles Parsons, "The Liar Paradox," *Journal of Philosophical Logic*, III, 4 (October 1974): 381-412. I am indebted to Parsons' fine paper and to exchanges with him nearly ten years ago, for considerable stimulation.

leo thought<sub>2</sub> was true: nothing thought<sub>1</sub> in room 13 at 6 A.M. was true, since nothing was thought<sub>1</sub> there-then at all. Since Galileo did think (think<sub>2</sub>) something, our objections to the previously discussed denial of step I in our argument are inapplicable.

There are many questions that might be raised about the viewpoint I have just sketched: Why should we believe that anything like subscripts characterize the application of predicates like 'think'? How are the extensions fixed? These are good questions, but I shall not discuss them here. I believe that they do not defeat the approach. In fact, I shall support a relative of this strategy in section II.

I believe, however, that the strategy is inappropriate for dealing with the paradox we have been discussing. Although the approach yields some unintuitive results, its main problem concerns motivation. The problem is to motivate the claim that Galileo's application of a propositional-attitude notion did not apply to his own thought. Such a motivation must rely on considerations involving schemata like (t), since the present paradox depends essentially on the notion of truth. The approach we are considering must motivate some connection between schema (t) and a failure of Galileo's application of the notion of thought to his own thought.

An attempt to provide the needed motivation is suggested by Charles Parsons.<sup>8</sup> The initial idea is that the thinker must presuppose a certain "scheme of interpretation" for the mental goings on that he is judging. The reasoning that leads to paradox is supposed to show that the interpretation cannot include or "give sense to" his own performance, and that only subsequent reflection on his scheme can do so. (I shall pass over the peculiarity of holding that only reflection on Galileo's mental act gives sense to it.) Even if this general line is accepted, it does little in itself to locate the shift of extension in the propositional-attitude notions rather than in the notion of truth. In Parsons' own discussion, the burden of the argument for locating extension shifts in the indirect-discourse notions is borne by an argument playing on the notion of proposition.

Transposing slightly, we can characterize the argument as follows. The truth schema applies to propositions; apparent failure of the schema indicates that no proposition was being thought by Galileo; but one thinks only propositions; so Galileo did not think any-

<sup>8</sup> *Ibid.*, pp. 385-388, 399-406. Parsons' remarks apply to assertion rather than to thought. Thus I am extrapolating. I do not claim that he would endorse the extrapolation. I would, however, criticize his own argument in a way similar to the way I respond to the extrapolation.

thing, or, at least insofar as we say he thought something, the notion of thought is not being used univocally with Galileo's own use. As we have seen above, this sort of argument employs 'proposition' in a dubious way. There is no obvious reason why one should not think thoughts that are relevantly deficient in truth value. There remains no clear motivation for focusing on thought rather than truth or for differentiating between Galileo's abortive application of the notion of thought and our successful application. It is more straightforward to focus on the truth schema.

*Denial of step 3:* We now consider the second of the two possible sources of trouble in our original argument to contradiction. This was the assumption, in step 3 of our semiformal summary, that Galileo's use of the notion of truth is normal and may be seen as coextensive with our use of the notion in the schema (t) that \_\_\_\_\_ is true if and only if \_\_\_\_\_. Qualifying this assumption forms the basis of a resolution of the paradox which I wish to explore. Our motivation stems from seeing Galileo's truth evaluation as nonnormal, relative to the *condition for normal application* laid down by schema (t).

Concentrating on step 3 in the argument carries with it a methodological advantage. It enables us to unite discussion of the present paradox with the main-line logical work on the semantical paradoxes concerning the attribution of truth to sentences. Virtually all approaches agree in placing some limitation on substitutions in analogs to schema (t). It is sometimes held that this unification of approaches is undesirable on the ground that the logician's use of the term 'true' to apply to sentences, as distinguished from thoughts, indirect-discourse statements, and the like, is a nonnatural, purely technical use. I see little merit in such claims. It is not in the least abnormal for ordinary language users to evaluate a piece of direct discourse as true or false. Sentence (and word) types or tokens as well as thoughts are quite normally subjected to semantical evaluation. Moreover, as I have just argued, that-clause discourse produces paradoxes structurally analogous to those of direct discourse.

Tarski restricted substitution into the analog of schema (t) by counting the offending natural-language sentences *ungrammatical* in his formal system. This measure has been widely recognized as inappropriate insofar as one wants to account for semantical notions in natural language. For the problematic sentences are not in general meaningless, much less ungrammatical. Paradox depends sometimes on empirical facts, not merely on the meaning of the

sentence.<sup>9</sup> A parallel point applies to paradoxes concerning thoughts. As the Galileo example shows, paradox may stem not from anything intrinsic to the thought Galileo thinks, but from the empirical facts—facts that may be unknown to Galileo or even to the reporter of his thought. One cannot restrict the relevant truth schema by placing context-free conditions on the meaning of the sentence or the nature of the thought.

Nevertheless schemata like (t) must be restricted. Let us return to the Galileo example. We do not want to take Galileo's thought to be true—where 'true' is to be used univocally in our evaluation and Galileo's. (I shall abbreviate this remark by saying that we do not want to take Galileo's evaluation to be *true on its own terms*.) If the truth schema (t) applied and Galileo's thought were taken to be true on its own terms, then we would have to concede that nothing thought in room 13 at 6 A.M.—not excepting Galileo's thought—is true. Here we would have succumbed to contradiction. We also do not want to count the negation of Galileo's thought as true on its own terms. For if Galileo had thought the negation under the circumstances, we would still have a thoroughly pathological evaluation, though not strictly one that would lead to contradiction. If we ask whether that thought is true, we must answer fatuously that it is if it is, and otherwise not. Galileo is not in a position to evaluate his thought on its own terms.

These considerations suggest that it would be reasonable to restrict schemata like (t) to cases where either the thought (equally, what is stated, meant, believed in a context) or its negation is true. And as far as I have gone so far, the restriction accords with, or is closely analogous to, restrictions proposed by almost all authors who have seriously discussed the semantical paradoxes of direct discourse.

Now we must note an aspect of the paradox which we have been suppressing. Given that something goes seriously wrong with Galileo's attempted truth evaluation, we have concluded that it is not true. But then since nothing else is thought in room 13 at 6 A.M., it follows that nothing thought in room 13 at 6 A.M. is true. But now we have asserted something in the very words we have been using to characterize Galileo's ill-starred thought. Indeed, we have just thought that nothing thought in room 13 at 6 A.M. is true. We are also committed to the truth of our thought. Have we not contradicted ourselves?

<sup>9</sup> Cf. Tarski, *op. cit.*; Parsons, *op. cit.*; Kripke, "Outline of a Theory of Truth," this JOURNAL, LXXII, 19 (Nov. 6, 1975): 690-716; my "Semantical Paradox," *op. cit.*

No, I think we have not. Each step we have taken in our reasoning has seemed reasonable. The proper response to the situation is to interpret the reasoning so as to justify it.

The stages in the reasoning are three. (a) Galileo thought his thought, but his application of the notion of truth misfired—as evinced by our inability to apply schema (t) to it. (b) Just because it misfired, we may judge that nothing thought in room 13 at 6 A.M. is true. To avoid equivocation and irrelevance, the notion of truth in this judgment of ours must be taken as coextensive with the occurrence of the notion in Galileo's own thought content. For it is the failure of this occurrence which we are judging. (c) We count this judgment of ours true. In so doing, we are ratifying Galileo's own thought content, since the occurrences of the notion of truth in our judgment and his thought are coextensive and since (by hypothesis) there is no equivocation between other notions in his content and ours. We have counted Galileo's thought [in stages (a) and (b)] not true and subsequently [in stage (c)] true.<sup>10</sup>

Since this reasoning is justified, it involves a shift in the extension of the notion of truth between stages (a) and (c). Since the shift does not involve a shift in linguistic meaning, it should be seen as indexical in a weak sense: a contextual shift in extension without a shift in meaning. Using subscripts to make vivid the idea: Galileo's thought content is that nothing thought in room 13 at 6 A.M. is true<sub>j</sub>. This content (whether he or we think it) is not true<sub>j</sub>; its negation is also not true<sub>j</sub>; and since nothing else is thought at the relevant place and time, nothing thought in room 13 at 6 A.M. is true<sub>j</sub> [stage (b)]. So  $g$  is true<sub>k</sub> [stage (c)].

Underlying this reasoning is the assumption that schema (t) in step 3 of our original argument to contradiction must be seen as involving a truth predicate whose extension undergoes contextual shifts. Call the truth schema whose application of the notion of truth is coextensive with that in a given thought content the schema *associated* with that thought content. Galileo's thought content cannot be evaluated by its associated schema. Neither it nor its negation is true on its own terms. But the thought can be evaluated via an occurrence of schema (t) whose application of the notion of truth has a different (broader) extension: Galileo's thought is true<sub>k</sub> if and only if nothing thought in room 13 at 6 A.M. is true<sub>j</sub>.

How are the levels marked by the subscripts contextually fixed? Under what conditions is neither a thought nor its negation true

<sup>10</sup> Parallel reasoning as applied to direct discourse is treated in far greater detail in my works cited in note 3.

on its own terms? I think that the answers to these questions run parallel to those I have developed for the application of truth schemata to direct-discourse sentences. (cf. note 3.) I shall not go into these matters here.<sup>11</sup>

## II

Let us now consider a problem that arises in the use of nonfactive propositional predicates unsupplemented by a truth predicate. Imagine that Galileo, at 7 A.M., entertains (with a view toward accepting or not) the thought that Galileo does not accept the thought being thought in room 13 at 7 A.M.<sup>12</sup> Still unapprised of his whereabouts, he is assuming he is not in room 13, but in room 12. Thus, the first thought being thought in room 13 at 7 A.M. is the thought Galileo entertains. We represent this assumption:

(1) The first thought = that  $\sim A(G, \text{the first thought})$

It would seem possible for Galileo either to accept or not to accept this thought. Suppose that Galileo does not accept it:

(2)  $\sim A(G, \text{that } \sim A(G, \text{the first thought}))$

Let us imagine that Galileo considers the matter and decides to abstain. (The temporal application of 'accept' will play a part in discussing this case. Although I think that such application is not finally at the root of the matter, it is important to keep track of it. We shall read ' $\sim A$ ' as the tenseless 'does not accept at any time'.)

<sup>11</sup> The levels are fixed by general pragmatic principles that I stated in "Semantical Paradox," *op. cit.* pp. 192-195. The conditions under which one can evaluate a thought with its associated truth schema have, I think, the structure of the Kleene strong tables extended to quantification in the natural way, and further equipped with iteration principles. These are stated in *op. cit.* pp. 185-190. (Cf. alterations cited in "The Liar Paradox," *op. cit.*, pp. 355-359.) Intuitive motivation for substantial parts of this structure derives from reflecting on remarks by Alonzo Church about an empirical version of the Epimenides paradox. The interested reader is invited to test the structure of the strong Kleene tables against Church's intuitive remarks. Cf. Church's review of Alexandre Koyré, "The Liar," *Journal of Symbolic Logic*, II, 4 (December 1946): 131.

<sup>12</sup> I intend 'accept' to apply to occurrent thoughts that involve belief. 'Occurrently believe' would do as well, except that it is barbaric. 'Think' is often used this way, but it also may bear the unwanted interpretation 'entertain a thought'—which lacks the element of epistemic commitment. 'Think' is also used as a mere synonym of 'believe'. There are those who sharply distinguish what they call "belief" from what they call "acceptance"; one is said to "accept" sentences but believe propositions. The distinction is occasionally associated with a view which I consider to be philosophically as well as intuitively mistaken: that one cannot believe necessary falsehoods; one can only "accept" necessarily false sentences. This view is subject to some of the same objections raised against the first response to the paradox of section 1, as well as to further objections. I note it here simply to disassociate my use of 'accept' from it.

Galileo then reflects and accepts that he does not accept that Galileo does not accept the first thought:

(3)  $A(G, \text{that } \sim A(G, \text{that } \sim A(G, \text{the first thought})))$

Again, this seems possible, even plausible, for a sufficiently reflective person.

Suppose next that Galileo is told (and comes to believe) that the first thought thought in room 13 at 7 A.M. *is* the thought he entertained. That is, Galileo now comes to believe what is represented by (1). It would seem plausible and at least possible that he would then realize that his nonacceptance, which he correctly believes in and which we will imagine him maintaining, is an absence from the first thought. We can represent this realization as:

(4)  $A(G, \text{that } \sim A(G, \text{the first thought}))$

But (4) contradicts (2).

I will want to argue that our way of representing the premises is the source of the trouble. But there is a simple criticism of the argument as it stands. When Galileo accepts that he does not accept the first thought (represented by (4)), his act undoes the claim that he does not (at any time) accept the first thought. The informal premise represented by (2) cannot then be true.

One might go on to dismiss the informal paradox. Once Galileo learns that the thought he entertained *is* the first thought—i.e., once he learns what is represented by (1)—the question of whether to accept that thought is *reopened*. In accepting it, Galileo just shifts his original position.

The main problem with this view is that it presents Galileo's own epistemic situation in the wrong light. Suppose we put ourselves in Galileo's shoes after he has learned what is represented by (1). He has already entertained the thought

(a) that Galileo does not (at any time) accept the first thought

and now he learns that this is the first thought. *Should* he now accept (a)?

If Galileo does accept (a), it cannot be true—so ideally, he should not have accepted it. If he avoids accepting it (as he has up until the present), it is true—and ideally, he should accept it. On this view of the matter Galileo cannot be right. (The difficulty is somewhat different if 'accept' is tensed or temporally indexed, but the intuitive problem can still be made serious by pursuing the question whether Galileo should change his mind.)

Galileo is in a position to *see* that it cannot be right to accept (a). Before accepting (a), thereby making (4) true, he could review the assumptions that seem to be forcing him into what is (on this account of the matter) clearly a mistake. Which assumption is he to give up? (1) represents an empirical fact. (2) represents his present position of nonacceptance of (a), and (3) is simply acknowledgment of that nonacceptance. So (2) and (3) represent self-conscious *avoidance* of accepting the self-defeating (a). But if he continues to avoid accepting it, he can see very well that it will be true. None of Galileo's options seem viable.

Dismissing the problem by claiming that at (4) Galileo has simply changed his mind thus fails to explicate what is puzzling about the situation. It simply shifts the difficulty from us to Galileo. Burden-shifting may beat whistling in the dark, but it is not the method of true philosophy.

If Galileo believes that (a) is the first thought, his reasoning about whether or not to accept it cannot reasonably be independent of the effect his acceptance or nonacceptance would have on the truth value of the thought. The problem for Galileo in our paradoxical case is, of course, not just that he cannot be right. The same point would apply to the first thought

(b) that Galileo accepts (at some time) the first thought

By accepting it, Galileo would make it true. But if Galileo's reason for acceptance is that acceptance would make (b) true, then any evaluation he tried to make would be palpably empty.

A condition on actually evaluating a thought in a given context is that the reasons one marshalls to support an evaluation (acceptance, rejection, suspension of belief) of the thought in that context not concern any effect that that evaluation might have on the truth value or acceptability of the thought. Once Galileo learns and makes use of what is represented by (1), he is in some sense not in a position to evaluate the thought (a). I think that the proper characterization of the situation is that, in such a context, the notion *accept*, as it occurs in (a), cannot apply to any evaluation Galileo makes of (a). He cannot accept (a) *on its own terms*. (Here we carry over this expression by obvious analogy from section 1.) His nonacceptance of (a) on its own terms is conceptually enforced.

Galileo's *reflection* on his nonacceptance of (a) is intuitively in a different position. Once Galileo learns that the thought he entertained is the first thought, his nonacceptance of that thought becomes unalterable. He does not accept it (on its own terms) by de-

fault: it is impossible for him to evaluate it in any way on its own terms. Reflecting again on his nonacceptance now forced upon him, he is still capable of realizing that he does not, cannot, and never did accept the thought (on its own terms). This reflective acceptance seems both possible and reasonable. But it is precisely an acceptance of the content of the first thought. Its intuitive reasonability is grounds for thinking that, despite Galileo's knowledge of what is represented by (1), his *reflective* evaluation is independent of the truth or falsity of what he evaluates.

What I have just argued is that the intuitive, *informal* reasoning that we began with is sound. The semiformal representations of premises (1)-(4) fail to capture an equivocation or indexicality in the informal reasoning. No expressions seem to change their linguistic meaning during the course of the argument. So the culprit is indexicality rather than ambiguity: contextual shifts of extension without shifts of meaning. Indexicality resides in nonfactive propositional-attitude predicates as well as in factive and semantical predicates.

To schematize the argument, let '1' mark the context of the first occurrence of 'accepts' and '2' the context of the occurrences marking reflective acceptance. Then the argument can be represented.

- (1') The first thought = that  $\sim A_1(G, \text{the first thought})$
- (2')  $\sim A_1(G, \text{that } \sim A_1(G, \text{the first thought}))$
- (3')  $A_1(G, \text{that } \sim A_1(G, \text{that } \sim A_1(G, \text{the first thought})))$
- (4')  $A_2(G, \text{that } \sim A_1(G, \text{the first thought}))$

The shift of extension in 'A' between (3') and (4') is not a function of Galileo's immediate intentions. It stems from conditions on the notion of evaluation. I now wish to step back from our example and attempt to articulate some principles about the indexicality of evaluative propositional-attitude terms like 'accept'.

### III

A first principle derivable from our examples:

- (A) Thoughts about evaluative attitudes can be self-referential.

The first thought that Galileo entertains (section II) is an instance. This principle is in conflict with traditional ramified type theory and with Ramsey's proposal. As an account of natural language or of actual thoughts, these theories are clearly too restrictive. Galileo's ability not only to entertain but to accept the thought just cited illustrates a second, stronger principle about self-reference:

- (B) An evaluation of a thought can itself make the thought true or false.

We now repeat a crucial principle enunciated earlier:

- (C) An evaluation of a thought is not possible by a given person in a given context if the reasons that the person has in that context that would support an evaluation concern the effect that the evaluation would have on the truth value or acceptability of the thought.

Once Galileo realizes that the thought he is entertaining is the first thought, and given that his reasoning to support an evaluation is not independent of the effect the evaluation would have on the thought's truth value or acceptability, he cannot accept or otherwise evaluate the question of whether he accepts the first thought.

Let us say that a person's chance to evaluate a thought (content) in a given context is *aborted* if the person's grounds for an evaluation in that context would concern the effect that the evaluation would have on the truth value of the thought.

Galileo's evaluations that inadvertently make the evaluated thought true or false illustrate an unseemly, incestuous relation between a thought's evaluation and its truth value or acceptability. I want now to characterize an idealized notion of *independence* that applies to actual and would-be evaluations. The notion is a rough analog of the familiar notion of groundedness (or, in our terminology, nonpathologicality or rootedness) appropriate to the notion of truth. It is intended to exclude evaluations, such as those by Galileo in section II, which would have been aborted had the evaluator had relevant knowledge or reasons.

- (a) Evaluations of nonevaluative thoughts (thought contents that contain no evaluative propositional-attitude notion) are independent.
- (b) Evaluations of disjunctive (conjunctive, conditional) thoughts are independent if the truth value of the thought is fixed by logical parts whose evaluations are independent.

I assume that a true disjunct (false conjunct) fixes the truth value of a disjunction (conjunction), and so forth.

For example, any evaluation of the thought

that thrice four is twelve or Galileo does not accept the first thought is independent, because the truth value of the thought is fixed by the truth value of the thought that thrice four is twelve, evaluation of which is independent.

- (c) Evaluations of negations are independent if evaluation of the un-negated thought is independent.

- (d) Evaluations of a quantified thought are independent if the truth value of the thought is fixed by an instance whose evaluation is independent.

A false instance fixes a universal quantification; a true instance fixes an existential quantification. Otherwise, all instances must be evaluation-independent for the quantification to be evaluation-independent.

- (e) Evaluations of evaluative thoughts are independent if the evaluation characterized by the evaluative thought is independent.

Thus my evaluation of whether you accept a certain thought is independent if your (actual or would be) evaluation is independent. Or, to illustrate clauses (d) and (e): Suppose that a Cretan is in room 13 without knowing it. That Cretan's evaluation of the thought

that all Cretans accept all thoughts they think in room 13

is independent if the thought is falsified by a Cretan's nonacceptance of another thought, where this latter nonacceptance is an independent evaluation; otherwise, the first-mentioned Cretan's evaluation is nonindependent.

- (f) That is all.

Call evaluations that are not independent *incestuous*. Incestuous evaluations would have been aborted if the evaluator had had relevant information and reasons.

It is natural to connect (B) with the notion of independence by holding that *incestuous evaluations are exactly those in which the evaluation affects the truth value of the evaluated thought*. It is also natural to conclude from (C) that where a person's evaluation does affect the truth value of the evaluated thought, this must occur inadvertently. Neither of these natural viewpoints is clearly correct. It seems arguable that a person can accept a thought, *knowing* that his acceptance will make it true—if his reasons support the evaluation independently of its effect on the thought's truth value.

Suppose that Galileo is told, as he enters room 12 on Thursday, that he is the first person in room 12 on Thursday. He is then asked to consider whether twice seven is fourteen. Immediately thereafter, he is asked whether

- (i) Galileo is the first person in room 12 on Thursday to consciously accept that he accepts that twice seven is fourteen.

Clearly, Galileo should and might well accept (i). Clearly, such acceptance is independent in the sense we defined. Arguably, his acceptance of (i) makes it true.

Without question, Galileo's acceptance of (i) makes it true that

- (ii) Galileo accepts that Galileo is the first person in room 12 on Thursday to consciously accept that he accepts that twice seven is fourteen.

But it seems plausible that in accepting (i) and making (ii) true, Galileo *thereby* consciously accepts that he accepts that twice seven is fourteen. There is no evident need for him to engage in a "separate" conscious act, accepting that

- (iii) Galileo accepts that twice seven is fourteen.

The relation between Galileo's acceptance of (i) and his acceptance of (iii) might, in the context, be quasi-logical: a manifestation of first-person authority. Galileo might realize this and realize that he is the first person there then to accept (iii). Thus Galileo might know that his acceptance of (i) makes it true. Yet the evaluation is both possible and nonincestuous (independent). This is because his reasons for accepting (i) do not concern the effect his evaluation has on its truth value—in accord with (C). It remains true that

- (B') An incestuous evaluation of a thought can itself make the thought true or false, but only if the evaluator's reasons for the evaluation do not concern the effect of the evaluation on the truth value or acceptability of the thought (roughly, only if the evaluator does not realize the incestuousness of the evaluation).

A person may come to realize that an erstwhile evaluation was incestuous. The person may come upon information that leaves him with no other relevant ground on which to evaluate the thought (on its own terms) than one that concerns the effect of the erstwhile evaluation on the thought's truth value. In such a case, further opportunity to evaluate the thought (on its own terms) is aborted. Nevertheless one may evaluate the thought in the light of one's new viewpoint. Call this *informed evaluative reflection*. When, in section II, Galileo realizes what the first thought is, he recognizes the incestuousness of his original suspension of judgment (nonacceptance). He did not accept the thought on its own terms; and now he cannot; any evaluation that would affect truth value is aborted. Nevertheless, he can understand that the thought was made true by his nonacceptance. So, on reflection, he appropri-

ately accepts the thought. The informed reflective evaluation is intuitively from an independent perspective; his evaluation does not affect the truth value of the first thought.

This shift from an unreflective, incestuous evaluation to an informed evaluative reflection is a source of shifts in extension in the evaluative, propositional-attitude predicate. It is perhaps not immediately obvious that this must be so. The point can be argued, however. Suppose that the first thought is that Galileo accepts (at some time) the first thought. Suppose that, being mistaken about what the first thought is, Galileo accepts it (on its own terms)—making it true. Later Galileo is apprised of what the first thought is. Now Galileo would not be in a position to accept it for the first time on its own terms. But since he once did so, it is true—and on reflection, he is capable of realizing that it is true. He now accepts it from his new, independent perspective. His original acceptance and his informed reflective acceptance both relate him to the *same* (first) thought content. So, it might be asked, how is it that occurrences of ‘accept’ that apply to the two acts have different extensions?

The answer is that independence of perspective derives from a difference in reasons. Even if both occurrences relate Galileo to the first thought, the occurrence that describes the informed reflective acceptance will relate him to thoughts that the former does not. Galileo could not have accepted the reasons that revealed to him how the first thought is made true and yet accepted it. For those reasons would have *aborted* the original acceptance.

These remarks support the following principle governing shifts in extension:

- (D) Occurrences of a predicate like ‘accept’ (‘reject’, ‘suspends judgment’) that characterize a person’s actual incestuous evaluation of a thought and occurrences that characterize informed evaluative reflection on that incestuous evaluation always differ in extension.

The argument for (D), which is the principal result of this section, is a relatively simple generalization from the case we just described. For expositional convenience, suppose that an occurrence of ‘accept’ has as its extension a single thinker and a set of thought contents. If a thought content is a member of the set, let us say that that member is *contained* in the extension of the occurrence. Now an occurrence that applies to an incestuous evaluation and one that applies to informed evaluative reflection on that evaluation would have the same extension only if they both contained the incestuously evaluated thought content in their extensions. Suppose they

do. This could come about only if the thinker obtained information relevant to his informed evaluative reflection on his incestuous evaluation which enabled him to identify the incestuous evaluation that he reflected upon as the thinking of the incestuous thought. In other words, in order to engage in informed evaluative reflection on an incestuous evaluation the thinker must come upon reasons that essentially concern the effect of the incestuous evaluation on the truth value of the evaluated thought. But these reasons are not (all) contained in the extension of the occurrence that characterizes the incestuous evaluation. For if they were, the thinker could not have evaluated that thought content on its own terms. The incestuous evaluation would have been aborted.

We shall now consider briefly what sort of structure might be discerned in these shifts in extension. I think it obvious that the structure has vastly weaker logical constraints than the corresponding structure governing shifts in the extension of the truth predicate. One reason for this derives from (B')—the possibility of incestuous evaluation, which has no semantical analog. Conditions on derivativeness of evaluation (more exactly, on nonpathologicality or rootedness) are simultaneously conditions on the extension of the occurrence of the truth predicate. A relevantly rootless application of the predicate 'true' never includes in its extension the sentence in which it occurs or the thought containing the notion of truth that it expresses. In the case of evaluative propositional attitude predicates, conditions on independence of evaluation are *not* simultaneously conditions on the predicate's extension [cf. (B')].

Another structural difference between semantical and nonfactive propositional-attitude predicates concerns reflection. Individuals differ in the extent to which they reflect on their evaluations. Iterative properties of truth are, by contrast, fixed by lower-level applications—and ultimately by nonsemantical contents. Moreover, the difference between independence and rootedness, noted in the previous paragraph, seems to carry over under iteration. Reflection on incestuous evaluation does not seem to automatically engender extension shifts. [Cf. (2')–(3'), section II.] Not all reflections on incestuous evaluations are “informed evaluative reflections.”

By (C) and (D), extension shifts are forced when a person's reasons supporting an evaluation would concern the effect of the evaluation on the truth value of the evaluated thought. It would enrich an account of these shifts if one could generalize about what sorts of thoughts a person has when he applies such reasons. Unfortunately, I see little hope of providing such a generalization. The

problem was pointed out by Lewis Carroll.<sup>13</sup> There are no beliefs one might have (or thoughts one might accept) that logically guarantee that one will see and apply the relevant relations among those beliefs. Galileo might even learn the identity of his thought [learn what is represented by (1') in section II], and yet not realize the significance of this information for evaluating his thought. Thus it seems to me questionable whether one can provide a substantially richer characterization of the conditions under which extensions shift than (C) and (D) already provide.

We should note an apparent tension between the Lewis Carroll point and our earlier argument for (D). Could a person move from an incestuous evaluation to an informed evaluative reflection on it because of a change of reasons, unaccompanied by a change of thoughts accepted? Could a person *simply* realize a connection between his beliefs which he had not previously realized, obtaining new insight without new information? I think this is not possible. Having thoughts does not guarantee using them as reasons. But an acquisition of reasons generates an acquisition of thoughts. The new information might be meta-level, an articulation of the new connections. Or, it might be irreducibly *de re* (that *that* was the first thought). But "realization," or "insight" into reasons, cannot be wholly nonconceptual in a person who already has higher-order concepts like acceptance.<sup>14</sup>

There is at least one further elementary question about the shifts in extension. When such a shift does occur, how are the extensions related? Often the new information and reasons that lead to informed reflection will involve *changes of mind*. (For example, Galileo accepted that he was in room 12 at the relevant time, but came to realize that this was mistaken.) Often informed evaluative reflection

<sup>13</sup> "What the Tortoise Said to Achilles," *Mind* (1895): 178. For a discussion of Carroll's point along these lines, see Barry Stroud, "Inference, Belief, and Understanding," *Mind*, LXXXVIII, 349 (January 1979): 179-196.

<sup>14</sup> Although I find this claim nearly obvious, I suspect that making it compelling would require more argument and a wider array of thought experiments than I wish to expound here. Thus, we might consider briefly the consequences of denying the claim. Such denial would undermine the generality of the argument for (D). One would have to construct a logically weaker analog of (D) which would characterize mandatory shifts in the sets of thoughts that function as *reasons* for incestuous evaluations and informed evaluative reflection, respectively. The evaluative predicates like 'accept' would normally undergo extension shifts, but sometimes only the development of the potential for different extensions containing different reasons. Once a person does articulate the shift in reasons that his thinking has undergone, following informed evaluative reflection, there will be a difference in thoughts that he accepts. I do not believe that this weakening of the viewpoint is justified. But I sketch it in lieu of a detailed justification of the stronger view.

tion constitutes a *forced position change*. [For example, if Galileo had incestuously accepted the first thought (section II), he would reasonably have not-accepted (rejected) it under informed evaluative reflection.] Because of these two factors, the shifts in extension will certainly not be monotonic. (Shifts in the extension of the truth predicate are monotonic increasing.) So far, I see no reason to expect any interesting *mathematical* regularities governing the shifting extensions of *particular* propositional-attitude predicates. What does seem plausible is that under informed evaluative reflection, the union of the extensions of evaluative predicates ('accepts', 'rejects', 'suspends judgment whether') enlarges, if we lay aside forgetfulness. Thus a weak analog of the cumulative hierarchy remains.

## IV

I wish to conclude with some more broad-brushed philosophical remarks. There is a group of propositional acts and states that do not involve evaluation by the subject—such as entertaining or considering a thought. Are these “nonevaluative” notions context-sensitive and derivative in the way the evaluative notions (belief, disbelief, hoping, fearing, and so on) are? There are grounds for thinking that they should be interpreted differently. Thus ‘I am accepting this very thought’ is empty and highly problematic. ‘I am thinking this very thought’, though awkward, is intuitively self-justifying. It seems to differ only stylistically from ‘I am thinking that I am thinking’. Both depend on self-referential intentions by the subject, traditionally expressed in Descartes’s *cogito*. As far as I can see, there is no analog of the conundrum of section II for nonevaluative propositional attitudes. This does not, of course, demonstrate the nonexistence of nonevaluative, mentalistic conundra. But I know of none that clearly derive from these notions, as opposed to, say, semantical notions. The paradoxes of grounding, for example, arise for thought contents generally. But these depend on notions like aboutness, denotation, and satisfaction. I conjecture that nonevaluative mentalistic notions are not indexical or derivative in the way the evaluative notions are. This is a reflection, I suspect, of the depth of the fact-value distinction even as it recurs at second intension.

The special role of evaluation in the puzzles suggests a problem for certain philosophical reductions. The indexical and derivative features of the notions we have been discussing resist recursive or even arithmetical specification. So it appears that the use of these evaluative notions cannot be captured by ordinary nonintentional syntactical or causal specifications (commonly called “functionalist” specifications) of procedures or processes within an individ-

ual.<sup>15</sup> Since evaluative predicates are counted indexical by our theory, this result may not seem to raise any problem for nonintentional, “functionalist” reductionisms that other indexicals do not already raise. I believe this appearance is misleading.

The requirement of derivativeness [expressed by (C) for propositional-attitude notions and via the notion of grounding or rootedness for semantical notions] and the practice of reflecting on failures of derivativeness and, in a sense, rising above them [expressed by (D) and the analogous extension shifts in semantical predicates] are fundamental to our concepts of cognitive evaluation. More simply, derivativeness and reflection on nonderivativeness are essential to our semantical and attitudinal notions. Any reductionist account of these notions must explain these features. But it is hard to see any of the standard reductionist strategies as promising any illumination on this score.

Syntactical and causal theories of mind and meaning have purported to give recursive or mechanistic specifications of notions the rules for whose use are essentially not recursively specifiable. But this fact barely suggests a deeper difficulty. It is that standard reductionisms have here, as elsewhere, largely ignored normative features of the normative or partly normative notions they have purported to explain. Philosophically interesting reductions of mentalistic and semantical notions do not seem to me to be in the offing—or even to be expected. But those with contrary viewpoints must explain such principles as (C) and (D)—and their analogs for semantical notions—in nonnormative (or at least nonmentalistic and nonsemantical) terms.

The epistemic conundrum of section II derives from assumptions formally analogous to naive reflection principles known to produce paradoxes in the use of the notion of truth. Other paradoxes can be produced using analogs of standard consistency principles.<sup>16</sup> We do not always reflect, nor are we always consistent, much less infallible or omniscient. So the logical constraints on nonfactive

<sup>15</sup> For a paper that develops a version of this general point, see Thomason, “Some Limitations to the Psychological Orientation of Semantical Theory” (forthcoming). It is unclear to me whether Thomason’s argument assumes that his opponents’ theories treat propositional-attitude notions as nonindexical.

<sup>16</sup> In fact, the argument of section II is in some respects more similar to that which leads to Gödel’s incompleteness theorems than to that which underlies Tarski’s theorem. The problem, both in the present case and in the case of the device that leads to the incompleteness theorems, arises with an epistemically motivated attitude or procedure. Given antecedent intuitions about the relation between the attitude or procedure and truth, the attitude or procedure is infected by the nonrecursive specifiability of truth.

propositional-attitude notions lack the power of those on semantical notions, which are more purely normative.

But the epistemic principles that are instrumental in producing the puzzle of section II do play an indispensable normative role in our use of epistemic notions. Indeed, they demand some empirical application. In judging a person's evaluations of thought contents, we must see such evaluations as often being accurately self-conscious and logically coherent. For attributing minimal consistency and some epistemic self-mastery is part of treating someone as a responsible rational being. This pressure to find self-consciousness and coherence in evaluations by people yields epistemic analogs of the semantical paradoxes. Both sorts of knot are unraveled by recognizing the indexical and derivative character of evaluative notions—semantical and attitudinal.

TYLER BURGE

University of California/Los Angeles

### COGNITIVE ABILITIES, CONDITIONALS, AND KNOWLEDGE: A RESPONSE TO NOZICK\*

**A**NALYSES of knowing that include subjunctive conditionals have received wide attention, but have typically been flawed by various fallacies concerning the use of conditionals.<sup>1</sup> Robert Nozick's recent account of knowing in *Philosophical Explanations*<sup>†</sup> is one of the most sophisticated conditional analyses of knowing yet developed. He argues that it may succeed in avoiding the particular mistake that I call the "conditional fallacy," thanks to the elegant simplicity of its conditional clauses. Of

\* I am grateful to Steven Luper-Foy, Michael A. Slote, and Ernest Sosa for comments on an earlier version.

<sup>1</sup> I have provided numerous illustrations in *The Analysis of Knowing: A Decade of Research* (Princeton, N.J.: University Press, 1983), and in "The Conditional Fallacy in Contemporary Philosophy," this *JOURNAL* LXXV, 8 (August 1978): 397-413, reprinted in *The Philosopher's Annual*, II (1979): 173-190.

<sup>†</sup> Cambridge, Massachusetts: Harvard University, 1981. Parenthetical page references, unless otherwise noted, are to this book.