The true fly is dipterous, but it has halteres which have evolved from posterior wings. It had been thought that the astonishingly rapidity with which a fly maneuvers must be due to a direct neural connection between the eyes and the wings, but recent experiments at Berkeley suggest something more sophisticated (Chan, Prete, and Dickinson 1998). It was known before that the halteres, which beat antiphase to the wings, act as gyroscopes which stabilize flight on all three axes by feeding information directly to the wing muscles. Remove a fly’s halteres and it crashes. What is new is that apparently the visual system is directly connected to the halteres, which then control the wing muscles. This fancy setup distinguishes between the aerodynamic forces and the Coriolis forces acting on the wings, permitting the fly to evade the flyswatter with marvelous ease.

The fly serves to remind us that an organism can discriminate aspects of its environment with superb accuracy, and make use of the resulting information in complex ways that help keep it alive, without anything we would, or should, call thought.

Leibniz, who believed animals are machines, was asked why he was so reluctant to kill a bothersome fly. A fly is just a machine, Leibniz replied, but what a wonderful machine; it certainly would not be right, he thought, to destroy a manmade machine of comparable complexity (Guhrauer 1846, Vol. 2, p. 364). Leibniz was perhaps more right than he realized: designing such a machine is surely beyond the dreams of even today’s technology. Leibniz also saw a profound difference between the fly (or any non-human creature) and man: man thinks, the fly does not. The fly’s reaction to visual input is far too rapid to involve thought. On the other hand, if a man managed to design such a machine, he might well be inclined to view his mechanical fly as calculating aerodynamic and Coriolis forces in order to maintain its stability during a double Immelman. But of course a machine that could do all that a fly can do, and no more, would not be calculating in the sense of giving conscious thought to the matter. So we need to ask what would turn calculation, in the sense in which a fly or a computer can calculate, into conscious thought?

We are just machines that are complex in ways flies are not, so the problem isn’t one of transcending mere physical devices. I do not doubt that an artificer could, at least in principle, manufacture a thinking machine. The problem, for philosophy anyway, is what to aim for; what would show that the artificer had succeeded? I assume that you and I can tell, given enough time, and the right sort of environment, whether an object can think, and we can tell this without any clear idea of what is inside the skin. In this respect, Turing had the right idea, though his test was not conclusive for a variety of reasons. But what, more exactly, is it that we detect when we recognize an object as a thinking being?

Animals show by their behavior that they are making fine distinctions, and many of the things they discriminate we do too. They recognize individual people and other animals, distinguish between various sorts of animal, find their way back to places they have been before, and can learn all sorts of tricks. So it is important to reflect on why none of this shows they have propositional attitudes: beliefs, desires, doubts, intentions, and the rest. They see and hear and
smell all sorts of things, but they do not perceive that anything is the case. Some animals can learn a great deal, but they do not learn that something is true.

Why doesn’t the fact that a horse or a duck discriminates many of the things we do strongly suggest that they have the same concepts we do, or at least concepts much like ours? This is a suggestion many find persuasive, and it is apparently unavoidably seductive for most of us when we describe the activities of dumb beasts. But there is little reason to take the suggestion literally. Someone could easily teach me to recognize a planet in our solar system (smaller than the sun and moon, untwinkling) without my having any idea what a planet is. A horse can distinguish men from other animals, but if it has a concept of what it is distinguishing, that concept is nothing like ours. Our concept is complicated and rich: we would deny that someone had the concept of a man who did not know something about what distinguishes a man from a woman, who did not know that fathers are men, that every man has a father and a mother, and that normal adults have thoughts. It is easy to be misled by the fact that creatures with propositional attitudes and creatures without can both be conditioned to respond differentially to many of the same properties, sorts of object, or types of event into thinking the same sort of thing is going on in the brains of both creatures. A creature with propositional attitudes is equipped to fit a new concept into a complex scheme in which concepts have logical and other relations to one another. Speechless creatures lack the conceptual framework which supports propositional attitudes.

I think this is enough to ensure that some degree of holism goes with having concepts. Many concepts are fairly directly connected, through causality, with the world, but they would not be the concepts they are without their connections with other concepts, and without any relations to other concepts, they would not be concepts. To say this is neither to suppose holism is so pervasive that no two people could, in any sense required for communication, have the same concept, nor is it to deny that the contents of some concepts are more directly attached to sensory moorings than others. We can appreciate why holism is not the disaster it has sometimes been portrayed as being if, instead of asking how the content of a concept or judgment looks from the inside, we ask instead how an observer can size up the contents of the thoughts of another creature. This is, again, the Turing approach. So here I want to say something about how I think it is possible for one creature with a full basic set of concepts to come to understand another, for I think this will throw light on the central question I raised, which is how we can tell when a creature has a genuine concept.

There is no distinction to be made between having concepts and having propositional attitudes. To have a concept is to class things under it. This is not just a matter of being natively disposed, or having learned, to react in some specific way to items that fall under a concept; it is to judge or believe that certain items fall under the concept. If we do not make this a condition on having a concept, we will have to treat simple tendencies to eat berries, or to seek warmth and avoid cold, as having the concepts of a berry, or of warm, or of cold. I assume we don’t want to view earthworms and sunflowers as having concepts. This would be a terminological mistake, for it would be to lose track of the fundamental distinction between a mindless disposition to respond differentially to the members of a class of stimuli, and a disposition to respond to those items as members of that class.

Given the task of deciphering a language we do not know, we will perforce start with perception sentences, sentences which a speaker will assent to or dissent from given a stimulus.
we too can sense. As Quine put it, “Linguistically, and hence conceptually, the things in sharpest focus are the things that are public enough to be talked of publicly, common and conspicuous enough to be talked of often, and near enough to sense to be quickly identified and learned by name.” (Quine 1960, p. 1) Our first guess as to what is meant by a perception sentence will be a shot in the dark, but given how much alike people are, getting it right is like hitting a barn door; the most casual guess is often correct. The simplicity of this mode of entry into an alien language should not leave us thinking that a concept so identified is defined by its external causes, without the aid of theory or a supporting nexus of further concepts. A concept is defined by its typical cause only within the framework of a system of concepts that allows us to respond to certain stimuli as tables, friends, horses, and flies. A concept is defined for us by its typical causes, given that we are already in the world of language and conceptualization. But patterns of stimulation do not, in themselves, delineate the content of any sentence or concept. Only a very modest degree of holism is enough to lead to the conclusion that no simple story about the causal relations between mental states and the world can account for intentionality, much less specify the intentional contents of thoughts or utterances.

Concepts, and the sentences and thoughts that employ them, are in part individuated by their causal relations to the world and in part by their relations to each other. Thoughts, because they have propositional content, are unlike everything else in the world except for utterances in having logical relations to each other. There is only one way for an interpreter to spot these relations, and that is by noting patterns among the utterances to which a speaker awards credence. Thus the interpreter will note that a speaker who assents to “It rained in Spain and we all got wet” will also assent to “It rained in Spain” and “We all got wet”; that a speaker who assents to “John is taller than Sam” will also assent to “Sam is shorter than John”; and so on. These examples illustrate the routes to different discoveries. The pattern of the first example holds no matter what sentences are substituted for “It rained in Spain” and “We all got wet”, and so leads to the identification of the truth functional connective for conjunction, one of the logical constants. The second pattern holds no matter what names are substituted for “John” and “Sam”, and so leads to the recognition of a logical relation between the two two-place predicates, “taller than” and “shorter than”. The former discovery is far more important, since it uncovers one of the most basic sources of the creativity of language, a recursive rule. One can easily see how the other logical constants, at least those involved in the first-order predicate calculus, can be identified.

A more subtle problem is that of discerning relations of evidential support among sentences, as viewed by the speaker, of course. These relations can be uncovered, but only by invoking a version of decision theory which, by finding the subjective probabilities of sentences, allows the computation of conditional probabilities. Degrees of evidential support, while more variable from speaker to speaker than matters of logic and logical form, are essential to the identification of theoretical terms less directly keyed to perception than perceptual sentences, for they provide the ties that give substance, along with the structure provided by theory, to theoretical concepts.

I have been pursuing the twin questions of the relations between thought and language and the world on the one hand, and the sort of structure thought and language require on the other, in order to evaluate claims that one analysis or another of thought or language is satisfactory, or to decide what criteria to employ in judging whether a creature or device is
thinking. How much structure should we demand? Here the fact that the structure of language mirrors the structure of propositional thought is a help. Possession of a concept already implies a degree of creativity, since the point of a concept is that it is applicable to any item in an indefinitely large class. The fixed singular terms of a language are presumably finite in number, but demonstrative devices, whether combined with sortal predicates or not, provide the means for picking out an unlimited number of items. The truth-functional connectives, with their iterative powers, supply a further form of creativity. But is creativity enough? There is a good reason to think not. Consider a language consisting only of names, predicates, and the pure sentential connectives. Such a language has a finite vocabulary, but a potential infinity of sentences, so it is creative. But it easy to give the semantics of such a language without introducing a concept of reference, and so without matching up either names or predicates with objects. The explanation is simple: given a finite vocabulary of names and predicates, the truth conditions of each of the finite number of sentences formed without the aid of connectives can be stated without considering the roles of the parts of sentences; the rest is truth tables. There is no reason to credit a creature with so simple a language with an ontology.

Should we nevertheless say a creature with these conceptual resources and no others has what we would call a language, or thoughts? It seems to me not. A creature without the concept of an object, however good it is at discriminating what we call objects, is a creature without even the rudiments of the framework of thought. What calls for ontology is the apparatus of pronouns and cross reference in natural languages, what we represent by the symbolism of quantifiers and variables in elementary logic. These devices provide the resources for constructing complex predicates, and at this point semantics must map names and predicates on to objects.

If I am right that language and thought require the structure provided by a logic of quantification, what further conceptual resources is it reasonable to consider basic? I have no definite list in mind, but if the ontology includes macroscopic physical objects, including animals, as I think it must, then there must be sortal concepts for classifying the items in the ontology. There must be concepts for marking spatial and temporal position. There must be concepts for some of the evident properties of objects, and for expressing the various changes and activities of objects. If such changes and activities can be characterized in turn, then the ontology must also include events, and among the concepts must be that of the relation between cause and effect. I am inclined to make some major additions to this list, as I shall indicate in a moment, but this is enough to suggest that the domain in which thought can occur is fairly complex. It is the domain each of us inhabits, but one we have good reason to suppose is inhabited by no other animal on earth, or machine.

Much of what I have said about the complexity and specificity of thought may be thought to be appropriate in connection with human thought, but applicable only in that context. In other words, I am revealing a provincial attitude toward intensionality. Well, perhaps. There is a further consideration, however, that may reinforce the anthropomorphic perspective. The important question, after all, isn’t whether some animals have a simpler or degraded set of concepts; it is the question whether they have concepts at all. There is a clear difference between being disposed to react in different ways to Vs and Ws, as octopodes can be trained to do, and having concepts, however vague and poor, of those letters. To have a concept is to classify items as instantiating the concept or not, to judge, however implicitly, that here is a V and there is a W. The difference, as is well known, lies in the idea of error. We can say, if we like, that the octopus
has erred when it reacts to the V as it was trained to react to the W. That was not what we had in mind, and its action may deprive the octopus of a tasty reward. But on what grounds can we claim that the octopus did not grasp our concept? As Wittgenstein says, whatever the octopus does is in accord with some rule, that is, some concept or other, which is a way of saying there is no reason to suppose it has any concept. What the octopus did, when it chose a V when we had trained it to chose the W, was not in accord with our idea of what resembles prior stimuli. But the judgment of resemblance is ours, not that of the octopus. So far as I can see, no account of error that depends on the classifications we find most natural, and counts what deviates from such as error, will get at the essence of error, which is that the creature itself must be able to recognize error. A creature that has a concept knows that the concept applies to things independently of what it believes. A creature that cannot entertain the thought that it may be wrong has no concepts, no thoughts. To this extent, the possibility of thought depends on the idea of objective truth, of there being a way things are which is not up to us. I do not see how any causal story about the sequence of stimuli reaching an isolated creature can account for this feature of conceptualization or intentionality, provided the story is told in the vocabulary of the natural sciences.

There is, I believe, a direction in which to look for a solution, and that direction has been pointed out by Wittgenstein. What is needed is something that can provide a standard against which an individual can check his or her reactions, and only other individuals can do this. To take the simplest case, consider two individuals jointly interacting with some aspect of the world. When the pair spot a lion, each hides behind a tree. If the individuals are in sight of one another, each also sees the other hide. Each is therefore in a position to correlate what he sees (the lion) with the other’s reaction. After a time, a consequence is that if one individual sees a lion when the other does not, the one who does not see the lion is apt to treat the first’s reaction as a conditioned stimulus, and also hide. Now consider a situation in which each sees the same lion, but one of the individuals, because the light is poor, or a tree partially obscures the lion, reacts as he normally reacts to a gazelle. This turns out to be a mistake. This little skit cannot, in itself, explain conceptualization or grasp of the idea of error on the part of either observer. It does no more than indicate the sort of conditions in which the idea of error could arise. Thus it suggests necessary (though certainly not sufficient) conditions for conceptualization.

Tyler Burge has argued (Burge 1986) that the content of a perceptual belief is the usual or normal cause of that belief. Thus the cause of the belief that a lion is now present is past correlations of lions with stimuli similar to the present stimulus. The difficulty with this proposal is that equally good answers would be that beliefs about lions are caused by the appropriate stimulation of the sense organs, or by the photons streaming from lion to eye, in which case the beliefs would be about stimulations or photons. There are endless such causal explanations, and each would dictate a different content for the same perceptual belief. It is natural to reply, and Burge does reply, that we have no idea how to characterize the various patterns of stimulated optic nerves that would be caused by a lion, aside from the way I just did it, by appealing to the role of lions. The force of this reply depends on the fact that we happen to have a single lion-concept, but no single concept for patterns of lion-caused firings of neurons. But nature with its causal doings is indifferent to our supply of concepts. When it is conceptualization that is to be explained, it begs the question to project our classifications on to nature.
Burge’s suggestion fails two tests: it fails to pick out the relevant cause, and so gives no account of the content of perceptual sentences, and it fails to explain error. Adding a second person helps on both counts. It narrows down the relevant cause to the nearest cause common to two agents who are triangulating the cause by jointly observing an object and each other’s reactions. The two observer don’t share neural firings or incoming photons; the nearest thing they share is the object prompting both to react in ways the other can note. This is not enough to define the concept, as I said before, since to have the concept of a lion or of anything else is to have a network of interrelated concepts of the right sorts. But given such a network, triangulation will pick out the right content for perceptual beliefs. Triangulation also creates the space needed for error, not by deciding what an error is in any particular case, but by making objectivity dependent on intersubjectivity.

It is clear that for triangulation to work at all, the creatures involved must be very much alike. They must class together the same distal stimuli, among them each others’ reactions to those stimuli. In the end, it is just this double sharing of propensities that gives meaning to the idea of classing things together. We say: that creature puts lions together into a class. How do we tell? The creature reacts in relevantly similar ways to lions. What makes the responses similar? Our concepts do; we have the concepts that define these classes. It takes another creature enough like the first to see and say this. The sharing of many discriminatory abilities explains why a considerable degree of holism is no obstacle to communication. This is also why Turing had the right idea about how to tell if a device (or animal) is thinking.

Here we have a reason why the third person approach to language is not a mere philosophical exercise. The point of the study of radical interpretation is to grasp how it is possible for one person to come to understand the speech and thoughts of another, for this ability is basic to our sense of a world independent of ourselves, and hence to the possibility of thought itself. The third person approach is yours and mine.

Triangulation depends not only on a minimum of two creatures, but equally on shared external promptings. For this reason, among others, I think Kripke’s account (Kripke 1982) of what he takes to be Wittgenstein’s “skeptical solution” to the puzzle of rule-following is inadequate to serve as the whole story about conceptualization. The problem is just the one we have been discussing: how to account for failure to apply a concept correctly, given that what one person might count as an error may just be another person applying a different concept. Kripke’s suggestion is that if a learner fails to apply a concept as his teacher thinks correct, the learner has made a mistake. Unfortunately this does not distinguish between failure to apply a concept incorrectly and applying a different concept correctly — the very distinction in need of explication. But Kripke’s examples have another, related, flaw: they concern mathematical examples, and so lack the shared stimulus to provide the possibility of a shared content.

Ostensive learning, whether undertaken by a radical interpreter as a first step into a second language, or undergone by the acquirerer of a first language, is an example of triangulation. The radical interpreter has, of course, the idea of possible error, and so do his informants, and he can assume he and they share most basic concepts. Thus a first guess is apt to be right, though there can be no assurance of this in particular cases. Someone being initiated into the wonders of language and serious thought is also being initiated into the distinction between belief and knowledge, appearance and reality — in other words, the idea of error. Like triangulation, ostensive learning runs the risk of leaving unclear not only how the next step
should go, but also what constitutes a wrong application of a concept at earlier steps. But neither the novice nor the sophisticated radical interpreter is in a position to question a teacher’s or informant’s early applications of a concept or word new to the learner. Teacher or informant may not be applying her own concepts correctly, but learner and interpreter must accept wrong steps as right until later in the game, since for them a concept is being given content. Erroneous ostensions on the teacher’s part just lead the learner to learn a different concept from the one the teacher wished to introduce, and so will promulgate misunderstandings.

How will the learner or interpreter discover when he is applying a different concept than the one his teacher or informant had in mind, and when one of them is misapplying the same concept? There are various possible answers. Some will appeal simply to the power of consensus; but this cannot be conclusive. Of course, consensus of use, where use is assumed to reflect what the teacher or society means, is just what the learner or radical interpreter needs to recognize, but consensus of application does not distinguish the two varieties of error. As far as I can see, nothing in the observable behavior of teacher or learner with respect to an isolated sentence can sort this out. The distinction depends on relations among uttered sentences. The relation of evidential support among sentences provides powerful clues. When the learner says “That’s a cow” when faced by a bull, is she erroneously applying the concept cow, or correctly applying a concept that covers both cows and bulls? If she also learns what may be the truth-conditions of “That’s an udder”, one can test whether she assents to “That’s an udder” when presented with a cow but usually not when presented with a bull.

But the real test, in my opinion, is learning to explain errors. It is when one has learned to say or to think, “That looks green,” “That man seems small,” “I thought it was an oasis” when one has said or thought that something blue was green, or that the large man in the distance was small, or that what looked like an oasis was a mirage, that one has truly mastered the distinction between appearance and reality, between believing truly and believing falsely. It is also at this point, of course, that the distinction becomes clear between falsely thinking a bull is a cow, and simply applying the word “Cow” to both.

* * * * *

Cognitive science aims, among other things, to deal with thought and thinking. Up to this point I have been chiefly concerned to speculate about the conditions thought must satisfy, about what constitutes the subject matter, or part of the subject matter, of cognitive science. But cognitive science also aims to be a science, or at least to be scientific. Is thought, as I have described it, amenable to scientific study?

One reasonable demand on a scientific theory is that it should be possible to set forth a structure in such a way that it can be tested. This requires laws, or generalizations, which predict what will be observed given observed input. Some of the most impressive early work in psychology satisfied this condition. It was found, for example, that if a subject — just about any subject who was not deaf — was repeatedly asked to adjust a variable tone so that it sounded half way between two fixed tones, the subject made decisions that consistently defined an interval scale, in other words, a scale formally like the ordinary scales for temperature. This is not psychophysical measurement, since it does not relate a physical magnitude with subjective judgments; the tonal scale simply relates subjective judgments with other subjective judgments.
Patterns of such judgments instantiate the laws specified by the axioms which define an interval scale.

Bayesian decision theory, in the form which Frank Ramsey gave it (Ramsey 1990), is more subtle, but it is similar in that it relates judgments to judgments. Ramsey showed how, given only choices between wagers, it was possible to construct two scales, one for degrees of belief (sometimes called subjective probabilities), and one for comparative degrees of perceived value. It is a question, of course, whether anyone’s choices satisfy the conditions necessary for constructing such scales. As is well known, this is a very tricky question because of the mixture of normative and descriptive elements that enter into an attempt to give empirical application to the theory. Actual tests of Bayesian decision theory seldom show perfect consistency with the conditions; on the other hand, neither do they show many absurd deviations from them, and there are often persuasive arguments to explain the deviations. What one can say is that “given the right conditions”, ceteris paribus, the laws of decision theory do describe how people make real choices. The fact is that we all depend on this. People will seldom risk their lives for a small reward, will pay quite a lot for a good chance at a large prize, and so forth. These are laws of human behavior, difficult as they sometimes are to apply.

The laws of logic are, in the same way, laws of thought — always, of course, given the right conditions, and so forth. Tarski-type truth definitions, modified to fit natural languages, describe the basic semantic structure that informs the human language ability. We do not know how to fit all the idioms of natural languages into the format Tarski provided, but a very impressive core can be handled. These three structures, of logic, decision theory and formal semantics, have the characteristics of serious theories in science: they can be precisely, that is, axiomatically, stated, and, given empirical interpretation and input, they entail endless testable results. Furthermore, logic, semantics and decision theory can be combined into a single unified theory of thought, decision, and language, as I have shown (Davidson 1990). This is to be expected. Decision theory extracts from simple choices subjective scales for probabilities, i.e., degrees to which sentences are held to be true, and for values or the extent to which various states of affairs are held to be desirable. Radical interpretation, as I briefly described it above, extracts truth conditions, that is, meanings, and belief from simple expressions of assent and dissent. Formal semantics has logic built in, so to speak, and so does decision theory in the version of Richard Jeffrey (Jeffrey 1965). Uniting the theories depends on finding an appropriate empirical concept, and one such concept is the relation between an agent, the time and circumstances of utterance, and two sentences, one of which the agent would rather have true than the other. The protocols for testing such a theory are like the protocols in the testing of decision theory except that the choices which express preferences are treated as awaiting interpretation rather than as already interpreted. Given the richness of the structure of the unified theory, it is possible to derive the usual scales for subjective probability and desire, applied to sentences, and from these to determine the meanings of the sentences.

There is a widespread feeling among philosophers — a feeling with which I have a good deal of sympathy — that we will not really understand the intensional attitudes, conceptualization, or language, until we can give a purely extensional, physicalistic, account of them. Unless we can in this sense reduce the intensional to the extensional, the mental to the physical, so the theme runs, we will not see how psychology can be made a seriously scientific
subject. Jerry Fodor argues that if intensional and semantic predicates “form a closed circle”, that is, can’t be reduced to physical predicates, this

appears to preclude a physicalistic ontology for psychology since if psychological states were physical then there would surely be physicalistically specifiable sufficient conditions for their instantiation. But it’s arguable that if the ontology of psychology is not physicalistic, then there is no such science. (Fodor 1990, p. 51.)

Fodor seems to indicate in a footnote that he is aware that psychological states and events may be physically describable one by one even though mentalistic predicates are neither definitional nor nomologically reducible to the vocabularies of the physical sciences. But if this is the case, as I have argued, then the issue is not ontological; the question just concerns vocabularies. Whether or not the ontology of psychology is physicalistic, my guess is that Fodor believes there can’t be a science of psychology if its subject matter can’t be reduced, either definitionally or nomologically, to that of the physical sciences.

I think there is a lot to this. Since psychology wants to explain perception, for example, it wants to explain how certain events physically described cause beliefs through the agency of the senses. Any laws concerning such interactions would, it seems, amount to partial nomological reductions of the mental vocabulary to the physical. (Psychophysical measurement has produced plenty of laws, but these typically deal with the relations between physical quantities and sensations, not thoughts.) Any really complete scientific psychology would have at many points to relate the mental and the physical, by which I mean events and states described both in psychological and in physicalistic terms. A lot depends, of course, on how strict one wants the laws of such a science to be. With an ample sprinkling of “other things being equal” and “under normal conditions” clauses, we constantly utilize “laws” that relate the mental and the physical in everyday life. But here nothing like the laws of physics are in the cards.

I didn’t finish discussing the unified theory of thought and language which I mentioned a few minutes ago. How much like a serious science could it be? Formally it’s as clear and precise as any science. The difficulties lie in the application, the empirical interpretation. One trouble springs from its holism, though not quite in the way one would expect. All serious science is holistic. Whenever we assign a number to a physical magnitude we assume the correctness of the conditions which must hold to justify the form of measurement involved. In the ordinary measurement of length, for example, we assume that the relation of longer than is transitive. This assumption has no empirical content until we give an interpretation to this relation, and once we do this, we are assuming that the operation we have specified for determining that one object is longer than another holds for all objects under study. If the law of transitivity fails in a single case, the entire theory of measurement of length is false, and we are not justified talking of physical lengths. Once one considers the further conditions imposed by the theory, one appreciates the thoroughly holistic character of almost any physical theory. What makes the empirical application of decision theory or formal semantics tricky is that the norms of rationality apply to the subject matter. In deciding what a subject wants or thinks or means, we need to see their mental workings as more or less coherent if we are to assign contents to them at all, given that the contents are partly defined by their relations to one another. As in any science, we must be able to describe the evidence in terms the relevant theory accepts. The trouble with
the study of thought is that the standards of rationality, outside of decision theory and logic at least, are not agreed upon. We cannot compare our standards with those of others without employing the very standards in question. This is a problem that does not arise when the subject matter is not mental.

In one respect, the unified theory of thought and meaning which I described is a little better off than one might think. The important primitive term in that theory is the one expressing the attitude of preferring one sentence true rather than another. This is certainly a psychological concept, and a pretty complicated one. So there is no chance that the theory can be specified in physical terms. On the other hand, the theory is entirely stated in extensional terms. The relation of preferring true is a three-place relation between an agent and two sentences, and it holds no matter how those three entities are described. Of course, propositional attitudes are involved; they just aren’t expressed, in the theory, in a way that individuates attitudes generally, and in the way that would make the theory circular. In testing the theory, one would have to devise a way of telling when an agent preferred one sentence true rather than another. This is not such a bad deal, for if the operation one hits on at first is wrong, the theory will yield nothing intelligible. The richness of the structure of thought and meaning will necessarily tease out a workable interpretation. This is the attitude we take to physical measurement, and, in ordinary life, the attitude we actually take to the understanding of others.
REFERENCES


