

Default Values, Criteria and Constructivism

ALAN GARNHAM*

University of Texas at Austin

Wittgenstein, in his later writings, gave an account of the meaning of expressions in terms of *criteria* for their application. As with many of Wittgenstein's later ideas the notion of a criterion has proved difficult to explicate. A recent account, which ties criteria to the philosophical doctrine of constructivism, provides a link between the concept of a criterion and a series of ideas about language understanding which have emerged in the past few years. It is shown that these ideas can be made to cohere within a general constructivist framework, and that an alternative realist framework is inadequate in this respect.

Many people working in the field of natural language understanding have recognized the potential importance of Wittgenstein's later writings for their work. However, the difficulty in divining Wittgenstein's overall line of thought in *Philosophical Investigations* (1953, henceforth *PI*) has prevented this potential from being fully realized. Some writers on Wittgenstein have laid too much stress on his apparent eschewal of theorizing (e.g., *PI* sec. 109). Specifically, it is often concluded that Wittgenstein had no *theory* of meaning, and held that such a thing was a chimera, which we should not chase after. If this is the correct characterization of Wittgenstein's thought then it would seem that his ideas could only be of relatively indirect relevance to work in the empirical sciences. In this paper I point out that there are reasons for rejecting this view of Wittgenstein's later work, and mention a different analysis of what he said about meaning. I go on to show how this new account is closely related to some ideas on language processing which have recently been mooted in psychology, artificial intelligence (AI) and linguistics, and suggest that constructivism, the philosophical position which Wittgenstein arguably takes up in *PI*, ties together and justifies these ideas.

One of the things which is clear about *PI* is that Wittgenstein very often argues against positions which he took in the *Tractatus Logico-Philosophicus* (1922). In the *Tractatus* Wittgenstein had developed the idea, put forward by

*Present address: Experimental Psychology, Sussex University, Brighton BN1 9QG, England.

Frege (1892) and Russell (1903), that languages should be analyzed as logical calculi, and that natural languages are inadequate in-so-far as they do not conform to the pattern of such calculi. As part of his overall thesis Frege had argued that the *sense* of a word should be thought of as a set of necessary and sufficient conditions which determined for any object whether it fell under the denotation of the appropriate concept. The sense of a sentence was that which determined whether it was true of any state of affairs. Wittgenstein's discussion of games in *PI* (secs. 69–71) strongly suggests that he came to believe that there was no set of necessary and sufficient conditions for the application of many concepts, so what alternative account of meaning did he give?

Many people believe that Wittgenstein's last word on meaning was that 'meaning is use.' However, it is clear from the text of *PI* that Wittgenstein remained concerned, if not with necessary and sufficient conditions for the application of concepts, then at least with *criteria* for their application (e.g., *PI* sec. 182). But what are criteria? Very little of the discussion of this difficult concept of Wittgenstein's has been perspicuous. One view which has the merits of being both simple and relevant to the present discussion is that of Putnam. In a number of papers (e.g., 1963) he has come to the conclusion that criteria must, after all, be members of sets of necessary and sufficient conditions. To see why he was forced to adopt this position it is helpful to consider the question what else could criteria be.

Putnam is a self-avowed realist and hence would adopt the thesis that there are only two kinds of rational inference, the inductive and the deductive. The move from a statement that certain criteria are satisfied to the statement that an object falls under some concept is usually taken to be a rational one, so the inference must be inductive or deductive. It is not usual to think of conceptual relations as holding empirically, so the inference must be deductive. But deductive logic is just the study of necessary and sufficient conditions, therefore criteria must be necessary and/or sufficient conditions for the application of concepts.

Interpreters of Wittgenstein face a dilemma. On the one hand Wittgenstein rejected the Fregean notion of sense, but he proposes to replace it with an account of meaning in terms of criteria, which apparently turn out to be necessary and sufficient conditions after all. Baker and Hacker (Baker, 1974; Baker & Hacker, 1980; Hacker, 1972) have shown a possible way out of this impasse. They claim that Wittgenstein rejected the tenets of realism, and in particular the idea that there are only two kinds of rational inference. Baker (1974) outlines a case for taking the *crierial relation* to be both *a priori*, and therefore noninductive, and yet *context-dependent* and *defeasible*, unlike the relation of deductive support. This case is developed at greater length in Baker and Hacker (1980).

The properties of the *crierial relation* can be illustrated using one of Wittgenstein's own favorite examples, pain behavior as a criterion for someone's being in pain (*PI* secs. 281 ff). Unless there is specific evidence to the contrary,

we do not doubt that a person who is crying out, holding the affected part or wincing is in pain. Such evidence, under these circumstances, is conclusive. Furthermore, it is no mere empirical fact that pain behavior is connected with pain, the relation is a conceptual one (Descartes and others have, of course, denied this, but see Baker & Hacker, 1980, for a discussion of Wittgenstein's position). However, an actor in a play may exhibit exactly the same pattern of behavior as a person who has been stabbed, but we do not take this as evidence for his being in pain. As a function of context, criterial evidence can be defeated.

Baker (1974) argues that Wittgenstein's position is best characterized as constructivist. Constructivism is a generalized version of intuitionism (see e.g., Dummett 1977, for an account) minus the psychological trappings. Intuitionists spoke of mathematical entities as aspects of the human mind, but Wittgenstein specifically rejected the idea that meanings are mental entities (*PI* p.54, note a). Constructivism is diametrically opposed to realism in a number of respects. The most important for present purposes is that it claims there is a close connection between what is true and what can be known to be true (*cf.*, *PI* sec. 514). For realists truth is determined simply by the state of the world, and there is no recognition of the fact that the possible states of the world are, in some sense, determined by the way in which we choose to conceptualize things. Intuitionists demand that every mathematical truth should have a *constructive* proof. In particular the proof that some object exists must comprise a way of constructing that object. Wittgenstein made a related request that all explanations of meaning be surveyable (*übersichtlich*, see *PI* sec. 122).

The way in which such a general issue as that between realism and constructivism is resolved will obviously have manifold consequences for the way in which meaning is studied in a number of disciplines. The two doctrines may appear to be metaphysical, but adopting one view does affect the light in which other kinds of questions are viewed, not least empirical questions. The rest of this paper is an attempt to show that the version of constructivism developed by Wittgenstein, the one which employs the criterial relation as an evidential relation connecting how we find things out with whether they are true, provides a useful perspective from which to view a large body of work of language understanding.

I will not try to argue that realism actually rules out any particular piece of theorizing about cognitive behavior, but it cannot be simply concluded that there is nothing to choose between realism and constructivism. In the physical sciences traditional mechanics modified to include the Lorentz-Fitzgerald contraction accounts for exactly the same data on the movement of bodies as the theory of special relativity, but the latter was developed from a simpler set of initial assumptions, and was easier to integrate with the rest of physics. There was very little question about which theory should be adopted. The motivation behind a theory is often an important consideration in deciding whether to accept it.

Those familiar with recent work in AI may have recognized a connection

between Wittgenstein's concept of a criterion and some of the ideas of frame-system theory (Minsky, 1975), in particular those of *default value* and *default reasoning*. Frames organize knowledge around conceptual entities. Such entities may have structural properties (Woods, 1975) which take values from a more or less well-specified domain. Frames contain place-holders for such values, called *terminals* or *slots*. Particular objects are represented as instances of concepts by the creation of instance frames in which (some of) the terminals are filled by specific values. The slots may be filled by information which is external to the frame-system, but when such information is not available typical values, called *defaults* can be assigned. Thus a HUMAN has a NUMBER-OF-LEGS, which can be taken to be two, unless there is specific information to the contrary.

Many of the structural relations represented in a frame-system are best thought of as conceptual relations, yet the system will rarely insist that some slot must be filled, or that the filler must take a constant value, for every exemplar of the concept. The lack of such stipulations amounts to saying that it is neither a necessary nor a sufficient condition for an object's being an instance of a particular concept that it has some property. However, as with criteria, default values are only overridden by specific information to the contrary.

Frame-system theory can also be used to explicate the notion of a prototype, which has been discussed in the literatures of psychology (e.g., Rosch & Mervis, 1975), linguistics (e.g., Fillmore, 1975) and philosophy (Putnam, 1970, 1975, discusses the closely related concept of a *stereotype*). The connection between prototypes and some of Wittgenstein's ideas has often been noted. Wittgenstein frequently talked about typical objects of various kinds (e.g., the discussion of samples in *PI* sec. 73, and of the importance of normal cases in fixing meaning, e.g., *PI* sec. 142). The parallel between Wittgenstein's ideas and those of frame-system theory allows an explanation of why his thoughts relate to prototypes. A prototype can be thought of as a frame in which all the slots which can be filled by default values are so filled, whilst the others remain empty. It is important to realize that frames with unfilled slots can be used in making inferences. For Wittgenstein a prototypical exemplar of a concept would tend to satisfy most or all of the criteria for being a member of the relevant class.

So far we have seen a fairly specific parallel between some of Wittgenstein's ideas and recent developments in AI. I now want to show how the constructivist framework allows us to make sense of a number of more general findings in the area of language understanding, which can only receive *ad hoc* explanations from the realist point of view. Constructivist semantics focuses on the connection between knowledge (what is true) and understanding (how we find out that it is true). A large number of recent studies have shown that a person has to deploy his knowledge of the world in order to determine the meaning of an utterance. However, more than a decade of previous research had been based on the assumption that language understanding should be studied by the manipulation of narrowly defined syntactic and semantic variables. Perhaps the best

known demonstration of how knowledge can guide language understanding is Winograd's SHRDLU program (1972), but work in psychology (e.g., Johnson, Bransford, & Solomon, 1973, and the much earlier work of Bartlett, 1932) and linguistics (e.g., Fillmore, 1975) has led to the same conclusion. The explicit connection drawn by constructivists between the meaning of an expression and knowledge of its typical contexts of use makes these findings seem wholly natural. An interaction between knowledge and understanding is not incompatible with the basic tenets of realism, but it is not predicted by them. In order to explain such an interaction additional assumptions must be introduced, which inevitably makes the realist account less parsimonious than that of the constructivists.

Wittgenstein proposed context-dependent criteria for the application of expressions. The obvious psychological implication of this principle is that an expression will receive different encodings when it occurs in different contexts. Anderson and Ortony (1975) and Garnham (1979) have shown that this is true of nouns and verbs respectively. Similar effects have been found for larger units up to the level of whole passages. Dooling and Lachman (1971) and Bransford and Johnson (1972) showed that the way in which a short text is encoded can vary as a function of context. Such results do not fit neatly into a realist framework, which proposes that the meaning of an expression is determined by a set of necessary and sufficient conditions for its application.

Bransford and his colleagues (see Bransford & McCarrell, 1975) have demonstrated in several experiments that people frequently cannot distinguish information that was actually presented to them from further facts that they have inferred. For example, subjects who heard the sentence:

Three turtles rested on a floating log and a fish swam beneath them.

assumed that they had been told that the fish swam under the log. Wittgenstein stresses the importance of a shared "form of life" (*PI* sec. 241, Pt. 2 sec. xi) for mutual understanding. Language alone does not carry all the information we wish to communicate. It would be tiresome and pointless, in most cases, to remember what information was actually stated, and what was presupposed or made available from context. Constructivism makes this "forgetting" of what inferences have been made seem very natural.

One other recent development in the study of language understanding fits very neatly into the constructivist framework. This is the suggestion that in order to understand a text a person must construct a *mental model* of the situation described (Johnson-Laird, 1978; Johnson-Laird & Garnham, 1980). There is no very simple parallel between the constructivist sense of *construct* and its sense in "to construct a mental model." The constructivist concept is a formal one, developed in the context of logic and mathematics. The idea that mental models can be constructed is based on the computer-brain analogy, and the possibility of representing the world by programming a machine. Nevertheless, the notion that

language understanding involves constructing a finite and, hence, presumably surveyable representation of the situation described conforms to the spirit of constructivism, whilst bearing no particular relation to realist principles.

It is tempting to think that the phenomenological content of mental models is of great importance. But to say that one thing models another is to say that the two have *structural* similarities. As Wittgenstein pointed out (*PI* sec. 53) an image of the color red and a sample of the color on a piece of card can both be used by someone who has been sent to fetch a red flower. There is no essential difference between physical and mental models. Furthermore, physical models do not always look like what they model. Models are not always isomorphic with what they represent, but there must be some structural similarity. One valuable property of models is that they need not reflect every detail of what they model. By deliberate simplification models can be made more easily surveyable, and allow their users to come to grips with domains which might otherwise prove too difficult to understand.

Many of the major ideas about language understanding which have emerged in the last decade are explained if we adopt the general principles of constructivism. Realism is not incompatible with these ideas, but it fails to make them cohere. It may seem surprising that Wittgenstein's ideas can provide a framework for empirical research into broadly psychological problems, given his own condemnation of scientific psychology (*PI* Pt. 2, sec. xiv). However, Wittgenstein's main complaint about psychology was that conceptual confusion has lead psychologists to try to apply empirical methods to questions which such methods cannot answer. This complaint should be viewed in the light of his more general claims that the philosopher's job is to clear up conceptual confusion, "to shew the fly the way out of the fly bottle" (*PI* sec. 309), and that every question which can legitimately be asked has an answer (*Tractatus* sec. 6.51, *PI* sec. 516). Wittgenstein would have held that the conceptual confusion which has bedevilled empirical research into cognitive behavior can be cleared up, and that when this has been done, it will be possible to see what empirical questions can legitimately be asked. Recent trends in AI, psychology and linguistics suggest that we are, at last, beginning to clear up some of the conceptual confusion.

ACKNOWLEDGMENTS

The work presented in this paper was undertaken whilst the author was in receipt of funding from the Science Research Council of Great Britain and the Sloan Foundation (Grant No. 79-6-4 to the Center for Cognitive Science, University of Texas at Austin). I owe what grasp I have of the later philosophy of Wittgenstein to the teaching of Drs. Peter Hacker and Gordon Baker of St. John's College, Oxford. I would like to thank Phil Johnson-Laird for many helpful comments both on matters of style and on matters of substance which he made on a much earlier version of this paper.