Chapter 9   Wittgenstein and Categorization Research in Cognitive Psychology

Eleanor Rosch

Research in psychology tends to reflect, sometimes self-consciously, prevailing philosophical viewpoints. Categorization is the area in cognitive psychology which deals with the ancient problem of universals, that is, with the fact that unique particular objects or events can be treated equivalently. Prior to the 1970s, categorization research tended to mirror the simplified worlds described in early Wittgenstein and in logical positivism. However, Wittgenstein's later philosophy has revolutionary implications for many aspects of human thought, among them issues in categorization. In this paper, I will argue that modern research in natural categories is actually derived from Wittgensteinian insights, but ambivalently so: It has tended to work with the symptoms rather than the root of his challenge.

Background

For the early Wittgenstein of the *Tractatus,* categories were the objects of reference of words. Language functioned as a picture of reality; it was made up of propositions, each word of which stood for an object. The relations among the words stood for the ways in which the objects were related. This can be seen as a linguistic version of the long-standing tradition in philosophy to treat categories as objects of knowledge. In Plato these objects were the Forms; in Aristotle, the formal causes of the categories, for example, their definitions in terms of genus and differentia. For the British empiricists who followed Aristotle in this respect (see Fodor, 1981), concepts consisted of a *connotation* (meaning, intension), which was a specification of the qualities that a thing must have to be a member of the class, and a *denotation* (extension), which was just those objects in the world which belonged to the class. Because categories were objects of names (in the special sense of the *Tractatus*) and/or objects of knowledge, they had to have certain properies. First, they had to be exact rather than vague; that is, boundaries had to be clearly defined. One cannot have vague knowledge. Second, category members had to have something in common; after all, that was the object of reference. And that which the members had in common had to be the necessary and sufficient conditions for membership in the category. Third, following from the other properties but never explicitly stated, was the assumption that all members of a category were equally good with regard to membership; either they had the necessary common features or they didn't. Thus categories were seen as a

common set; all positive instances should manifest the common characteristic(s) defining membership, and negative instances should lack it.

It is somewhat paradoxical that the issue of categories did not enter into experimental psychology until the 1950s. From the 1920s American psychology had been dominated by behaviorism. The natural analog of the philosophical problem of universals in the conditioning paradigm is stimulus generalization. However, none of the behaviorists appeared to make this connection nor to speak of generalization gradients as relevant to concepts or categories (this, despite the fact that Clark Hull had performed a concept learning experiment for his doctoral research in 1920; see Brown, 1979). We might speculate that this was due not only to the mentalistic flavor of the word *concept* but also to the fact that generalization gradients violate all of the requirements for a proper category of empirical reference as laid down by logical positivism, which was the philosophical position explicitly espoused by the behaviorists (Marx, 1968). The *Tractatus* was seminal in the formation of logical positivism and the criteria for the reference of a category term in empirical discourse were just those stated above.

It was not until the publication of Bruner's *A study of thinking* (Bruner, Goodnow, & Austin, 1956) that research in categorization began to come into its own. Bruner's emphasis was on the active rather than passive aspects of learning and on the necessity of reintroducing cognitive theoretical terms such as *concept*. The main body of the book is the report of a specific program of experimentation in concept learning. It is in the design of these experiments that we can see a reflection of the prevailing philosophical assumptions about the nature of categories. Stimulus arrays typically consisted of items which represented all possible orthogonal combinations of an arbitrary set of attributes. For example, there might be forms which were squares and circles, each one of which was either red or blue and each one of which had one or two borders. The concepts which subjects learned were defined by specific attributes combined by a logical rule; for example, *red, blue and square, round or blue*. For such concepts, once the subject had learned the rule(s) defining the positive subset, boundaries of the concept could only be well-defined and any instance which fit the rule(s) was equivalent to any other. In fact, in the terms of the early Wittgenstein work, these tasks and the concepts derived from them are as close as one can imagine to simplified language-games in which reference and the mapping of simple elements onto reality has its clearest portrait. Research on this type of task burgeoned (Erickson & Jones, 1978; Neimark & Santa, 1975).

In developmental psychology, concept learning also became an area of interest. Input came from the traditions of Piaget (1972) and Vygotsky (1962) as well as from the concept identification paradigm outlined above. Much of the emphasis was on how children's concepts and learning strategies differ from the adult mode; the guiding image seemed to be one in which children's "irrational" and ill-structured concepts could be seen to develop into the clear and logical concepts of adulthood (Bruner, Olver, & Greenfield, 1966).

The later philosophy of Wittgenstein entirely reverses the position of the *Tractatus* and offers a profound criticism of virtually all previous philosophy.

His central point is that it is false (actually meaningless) to claim that language is necessarily a description of the world and that words and propositions get their meaning through the objects to which they refer. It is this image of language which has produced the false problems of philosophy – metaphysics, skepticism of all sorts, atomism, logical positivism, and so on. By dissolving the root position, the "false picture which had us in its grip," he can dissolve these problems. He has an alternative view of language: Rather than being referential, rather than giving us the "facts," language is part of our actions, part of the most basic practices which make up our physical and social "forms of life."

This view has radical implications for categories. They are no longer objects of words or knowledge but are part of our delicately shifting forms of life. Therefore, they need not meet any of the conditions for categories outlined previously.

1. They need not be precise; boundaries can be ill-defined. In Wittgenstein's words:

Frege compares a concept to an area and says that an area with vague boundaries cannot be called an area at all. This presumably means that we cannot do anything with it. – But is it senseless to say: "Stand roughly there?" Suppose that I were standing with someone in a city square and said that. As I say it I do not draw any kind of boundary, but perhaps point with my hand – as if I were indicating a particular *spot*. And this is just how one might explain to someone what a game is. (*PI*, 71)

2. They need not have anything in common, any common defining attributes:

Consider for example the proceedings that we call "games." I mean board-games, card-games, ball-games Olympic games, and so on. What is common to them all? – Don't say: "There *must* be something common, or they would not be called 'games'" – but *look and see* whether there is anything common to all. – For if you look at them you will not see something that is common to *all*, but similarities, relationships, and a whole series of them at that.... And the result of this examination is: we see a complicated network of similarities overlapping and criss-crossing:...
I can think of no better expression to characterize those similarities than "family resemblance"; for the various resemblances between members of a family: build, features, colour of eyes, gait, temperament, etc. etc. overlap and criss-cross in the same way. – And I shall say: games form a family. (*PI*, 66–67)

3. All members need not be equally good members. Wittgenstein does not explicitly discuss this issue in relation to categories but he seems aware of it:

Someone says to me: "Show the children a game." I teach them gaming with dice, and the other says "I didn't mean that sort of game." Must the exclusion of the game with dice have come before his mind when he gave me the order? (*PI*, p.33 e)

These are some of the implications of Wittgenstein's refutation of the object of reference view of categories. My contention in this paper is that it is these implications *only* which have received the attention of Wittgenstein-oriented categorization research in cognitive psychology. The root which produced them, Wittgenstein's challenge to the object-of- reference view of language, has tended to disappear from focus. Thus the Wittgensteinian ideas (prototypes, family resemblances) which have replaced criterial attributes of categories have tended

to be reified and eventually simply substituted for criterial attributes as that to which category terms refer – by the critics, at least, if not also by the proponents of the now well-known "non-classical view" of categories. I will attempt to trace, and perhaps place in a broader context, this tendency.

The Entrance of Wittgenstein's Ideas into Categorization Research

Prototypes. By the end of the 1960s, a new cognitive psychʋlogy and new approaches to language were rapidly appearing (Kessel & Bevan, in press). Chomskian linguistics, Neisserian cognitive psychology, developmental psycholinguistics, cultural relativity, ethnoscience – there were a wealth of new ideas. Since it is my own categorization research, growing out of issues in cultural relativity, which was the most explicitly influenced by Wittgenstein, I will describe something of its origins.

In 1969, I was involved in research on color categories. Colors, which had once seemed the clearest imaginable case of linguistic relativity (Brown & Lenneberg, 1954), now appeared to have a universal aspect (Berlin & Kay, 1969). I was working on the hypothesis that for basic color terms, categories formed around physiologically salient points in the color space. These points were most easily remembered, and thus first became attached to color names. Adjacent colors became part of the category by stimulus generalization. Thus color categories universally became structured around these same salient points. Evidence from the naming, memory, and learning of colors both in the United States and for a New Guinea people, the Dani of West Irian, who lacked basic hue terms, seemed to support this story (Heider, 1971; Rosch, 1973). What had this to do with the general issue of categorization?

Colors seemed markedly unlike the kinds of categories studied in the concept attainment literature. First, they do not have definite boundaries; not only is there experimental evidence that people disagree with each other and with themselves about the category boundaries (Berlin & Kay, 1969; McCloskey & Glucksberg, 1978) but, in fact, it was this very aspect of colors that had given rise to the view of their cultural relativity. Second, there are no identical attributes in common to all members of the category; physical properties of light, such as wavelength, vary continuously. Wittgenstein comments on color: What do all blue things have in common? – Just that they are blue (*PI,* 72). Third, some members of the category (for example the central, universal, salient points) are clearly better examples of their category than others. How could we have concepts like this?

"But is it senseless to say: 'Stand roughly there'? ... point[ing] with my hand – as if I were indicating a particular *spot*" (Wittgenstein, *PI,* 71). As the result of a dispute between two neighbors in the street on which I lived as a child, it

became apparent that none of our property lines were exactly known, and I remember thinking how remarkable it was that we could know that our houses were firmly in the middle of their yards without knowing where any of the yards ended. Might color categories, and other categories as well, be of this nature? I called the salient points at the centers of the color categories *prototypes* and must admit that I was ambivalent from the start about whether I thought of "proto-type" simply as a placemarker indicating the center of a "Wittgensteinian" cate-gory or whether I thought of it as an actual something, for example, a mental code to which the category name might refer.

Upon returning from New Guinea, I set about operationalizing the extension of the color work to other natural categories. Did semantic categories in general, like colors, have best examples and gradients of membership? The first concern was whether subjects would agree on typicality ratings of items. The categories used were common superordinates such as "bird" and "vehicle," and their instances were the words listed as examples of the category in the Battig and Montague (1969) norms. Experimental instructions are often revealing. The sub-jects were being asked to rate, on a 7-point scale, the extent to which each instance "represented their idea or image of the meaning of the category term." And how were these instructions conveyed to them?

This study has to do with what we have in mind when we use words which refer to categories. Let's take the word *red* as an example. Close your eyes and imagine a true red. Now imagine an orangish red ... imagine a purple red. Although you might still name the orange red or the pur-ple red with the term *red,* they are not as good examples of red (as clear cases of what *red* refers to) as the clear "true" red. In short, some reds are reder than others. The same is true for other kinds of categories. Think of dogs. You all have some notion of what a "real dog" a "doggy dog" is. To me a retriever or a German shepherd is a very doggy dog while a Pekinese is a less doggy dog. Notice that this kind of judgment has nothing to do with how well you like the thing; you can like a purple red better than a true red but still recognize that the color you like is not a true red. You may prefer to own a Pekinese without thinking that it is the breed that best represents what people man by dogginess. (Rosch, 1973, pp. 131–132)

(Needless to say, neither *red* nor *dog* were categories in the experiment.)

These instructions (even the phrase "what red refers to") were meant quite innocently. They were intended, and in fact accomplished, the task of communi-cation to subjects – by whatever experiences or presuppositions or folk theories we shared – what I meant by good examples and typicality. In fact, subjects had no trouble in understanding these instructions. Typicality ratings of items were highly correlated and reliable, and were found to predict reaction times in a vari-ety of categorization tasks (Rips, Shoben, & Smith, 1973; Rosch, 1973). Yet note what is happening: It is remarkably easy to describe – and think of – prototypes as particular "things."

In subsequent research the reification of prototypes continued – though it was never complete. The idea of cognitive representations was by then estab-lished in cognitive psychology, and it became natural to think of categories as represented by a mental code (Posner, Boies, Eichelman, & Taylor, 1969). It was an unknown code about which, with suitably ingenious experimental methods,

one could make discoveries, perhaps finding answers to age old questions such as the nature of abstraction or imageless thought (Posner, 1969; Rosch, 1975a). What was the intension (meaning) of a category actually like? I conceived of this code as the "mental representation generated by hearing the category name," and by use of the technique of priming (Beller, 1971; Posner et al., 1969) in a series of experiments (Rosch 1975a, 1975b) drew the following types of conclusions: (1) that such a representation was more like good than bad examples of the category; (2) that it was in a form common to both words and pictures but somewhat closer in format to pictures; and (3) that it was manipulable through practice for semantic but not color categories. Although the representation was treated as a definite thing, no specific claims were made for its functions or for its relation to meaning. However, any proposed concrete code, as we shall presently see, makes an easy target for such claims.

Family Resemblances. Where do prototypes and typicality orderings come from? With colors the origin can be argued to be physiological. But for semantic categories? At the same time as the emergence of the concept identification paradigm there had been developing a strain of research on abstract mental representations called schemas (Bartlett, 1932). Using primarily artificial stimuli, schemas tended to be operationalized as the central tendencies, such as means and modes, of the quantifiable dimensions on which members of the category differed. Schemas could also be centers of axes in a multidimensional scaling space. And there is an extensive literature on schemas derived from gestalt configurations having no definable attributes, such as families of random dot patterns (see reviews in Posner, 1969; Mervis & Rosch, 1981; and Medin & Smith, 1984). Here is a clear convergence of schema and prototype research: One type of prototype would seem to be explainable by the organism's computation of central tendency. But what of categories that do not seem to consist of quantifiable dimensions or even gestalt configurations – as perhaps *games?*

Wittgenstein enters, at this point, as the guide to a specific research proposal. The concept of family resemblances played the role in Wittgenstein's argument of a counterexample. His "opponent" claims that categories *must* have something in common in order to be proper referents of words. Wittgenstein points out that, in fact, categories such as game do not have anything in common; all that they have is a complicated network of similarities which we may call family resemblances. Philosophically it is a relatively peripheral part of the argument – snipping at the leaves and branches of the object of reference model rather than attacking the roots as he does elsewhere. Might it, none the less, form a central part of the empirical investigation of categories? Wittgenstein (in *PI*) says of family resemblances "look and see," and Carolyn Mervis and I decided to look and see.

In order to investigate family resemblances one needs a set of categories, many instances of these categories, and a computation of the attributes of the instances. Accordingly, we had subjects list all the attributes they could think of for 20 instances of each of six superordinate categories (an example of a category

is *furniture* and of an instance *chair*) and for 15 pictures of instances of each of six basic level categories (an example of a category is *chair* and of an instance a picture of a chair). Independent ratings were obtained of how good an example each instance was of its category. From the lists of attributes we then computed family resemblance scores for each item, for example, the number of attributes which each item shared with other members of its category (Rosch & Mervis, 1975).

The first part of the family resemblance claim is that all instances of categories need not have attributes in common. It was in fact the case that, for the superordinate categories of this study (as in a previous study: Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976), few attributes were given which were true of all members of any category – for four of the six categories, there was only one such item; for two of the categories, none. Furthermore, the single attribute which did apply to all members, in three cases was true of many other items besides those within that superordinate (for example, "you eat it" for fruit). But, of course, if all you can say about a group of items is that they have no attributes in common, one might well ask, why should there be a category here at all?

The second part of the family resemblance claim is the positive one that instances do share family resemblances. We next analyzed our family resemblance measures. The results both for superordinate and basic level objects were very clear: Family resemblance scores correlated highly with ratings of typicality. The better subjects rated an item as a member of its category, the more attributes it shared with other members of that category. Furthermore, thinking that this finding could account for the persistent "illusion" that all members of a category do have attributes in common, we further analyzed the five most typical items in each category. These did, indeed, prove to have many attributes in common: For example, chair, sofa, table, dresser and desk do share a number of attributes with each other, and they do not share these with mirror, stove, clock, picture, vase, telephone or other marginal items of furniture. We reasoned that if, on hearing the category name, people tend to think of the most typical category members, then – even when not under the sway of philosophical reasoning – one would have the impression of commonality of attributes.

There is a third aspect of family resemblance that Wittgenstein does not mention in the *Philosophical investigations* but which played an important role in his rejection of the logical implications of the *Tractatus* (Waismann, 1979). Categories do not occur in isolation. Any time one places an item into one category one is simultaneously not placing it into other contrasting categories. We also had subjects list attributes of members of contrasting categories. Negative correlations were obtained between typicality ratings and an item's possession of attributes characteristic of contrasting categories.

In summary, we found that the most prototypical items in categories have most attributes in common with other members of their own category and least attributes in common with other categories. To establish this finding irrefutably we conducted two more experiments using artificial categories consisting of letter strings in which family resemblance relationships could be built in and possi-

ble confounding factors excluded. Results were the same as for the natural cate-
gories. In fact, family resemblance not only correlated with typicality ratings but
also with ease of learning the item and with reaction time to identify the item
after learning – the standard independent variables with which typicality ratings
had correlated in other research on prototypes.

Growth of Nonclassical Categorization Research

Although the empirical evidence challenging the criterial definition view of cate-
gories has aroused wide interest, the concept of prototypes and family resem-
blances has seemed underspecified or in other ways inadequate to many
researchers. A host of new models which specify what the prototype is and how
it functions in decision procedures have been proposed and investigated. These
models can be classified roughly into (1) exemplar models, in which specific
exemplars are stored and membership for new instances computed by means of
a similarity function, and (2) probabilistic models, in which the category is stored
in terms of a core on which operate probabilistic functions for classifying new
members (see Smith & Medin, 1981, for a detailed account of these models).

    Noncriterial attribute accounts of categories (often called "nonclassical,"
whether they are prototype, exemplar, probabilistic or otherwise) have generated
an enormous body of research (see Mervis & Rosch, 1981; and Medin & Smith,
1984, for reviews). On the whole, typicality (used now in a generic sense) has
been found to correlate highly with verification times for category membership,
order and probability of production of category exemplars, and many important
variables in the learning and development of categories. It has been applied to
the natural language use of words, and in linguistics, Lakoff (in press) has
recently compiled an enormous body of evidence for nonclassical effects in lan-
guage, which he uses as an attack on Chomskian linguistics. The nonclassical
view has been incorporated into models of semantic memory and into artificial
intelligence programs. In philosophy, it has become an inevitable part of the dis-
cussion of natural kind terms (see, for example, Kelley & Krueger, 1984). And it
is being applied to various issues in social and clinical classification. What has
all of this to do with Wittgenstein?

Prototypes, Etc. as Objects of Reference

Prototype was initially conceived as a noncommittal designation for the central
regions of categories such that nonclassical category structure could be investi-
gated empirically. The term gradually became reified as indicated. The result is
that a prototype account or any other model of nonclassical category structure is
expected to fulfill the very functions of the classical criterial attribute view –

namely, to be an object of reference that gives meaning to words and makes true knowledge possible. This may be most clearly seen in some of the criticisms of prototypes.

1. *Formal Semantic Conditions.* If it is the prototype (exemplars, etc.) that are to provide the meaning for words, then prototypes must fulfill the requirements of a formal semantic model; for example, account for synonymy, contradiction, and conjunctive categories. Osherson and Smith (1981) modeled prototype theory with fuzzy set logic and showed that prototypes of conjunctive categories did not follow the rule of conjunction as maximization; for example, a pet fish is neither a prototypical pet nor a prototypical fish. This was taken as a refutation of prototype theory. Hampton (1985) accepting these same requirements, argued that prototypes do account for pet fish.

2. *Context Effects.* The *meaning* of a word must not change with conditions of its *use*. One of the great virtues of the criterial attribute view is that critical attributes are just that which is unchanging over contexts. If prototypes are to fulfill this function, they must be unaffected by context. But many studies have shown changes in comprehension or memory of category terms as a function of context (Anderson & Ortony, 1975; Anderson, Pichert, Goetz, Schallert, Stevens & Trollip, 1976; Barclay, Bransford, Franks, McCarrell & Nitsch, 1974; Barsalou, in press; Potter & Faulconer). For example, typicality ratings for animals are different in the context of African than American animals. And different words will cue memory for *piano* when it occurs in the context of *playing* versus *moving* a piano. Context effects are often taken as a refutation of protoype theory.

3. *Typicality Effects Are Too Universal.* If the classical definition of a category is taken as the meaning to which the term refers and if the prototype is taken as the meaning to which a term refers, then no category can have both a classical definition and a prototype (unless it is a strange case of polysemy – because a prototype and a classical definition are not the same). Armstrong, Gleitman, and Gleitman (1983) demonstrated that all of the typicality effects associated with empirical categories such as *furniture* and *fruit* are also found for categories which are argued to have, by necessity, a criterial definition, such as *odd number*. It is only by assumptions such as we have outlined that this could possibly be taken as a refutation of prototypes.

4. *Core Concept and Processing Heuristics.* A class of models of categorization have been proposed in which the actual meaning for category terms is a classical definition onto which is added a processing heuristic or identification procedure that accounts for typicality effects (Armstrong et al., 1983; Caramazza, 1979; Glass & Holyoak, 1975; Hampton, 1979; McClosky & Glucksberg, 1979; Smith, Shoben & Rips, 1974). This is, in fact, the way in which Armstrong et al. (1983), resolve the finding that *odd number* can "have" both a classical definition and a prototype. By this masterful stroke, data are consigned to prototypes and other

nonclassical models for their explanation and criterial definitions, freed at last, can perform unhampered the philosophical function which is their heritage.

But what is wrong with meanings as objects of reference? Couldn't prototype or other such models perhaps fulfill this function for category terms in the way that classical definitions could not?

Wittgenstein's Criticisms of Objects of Reference

Suppose that prototypes are the objects to which category names refer, the objects which are their meaning. To have meaning in the reference view a word must (1) have an object for which it stands – in this case the prototype, and (2) knowing what the word stands for should be a sufficient condition for its having meaning, for understanding it. How will this work with prototypes? In the first place, if prototypes are to serve as the objects of reference, the meanings of words, they cannot simply be hypothetical constructs in the theory (or mind) of the experimenter; they must also be in the world or in the mind of the individual using the word. (The four criticisms of prototypes just enumerated would not be coherent without such an assumption.) So we may ask *where* they are.

In present cognitive science there are two basic possibilities for the location of such "entities," the external world or the cognitive representation. Since cognitive variables are usually placed in the cognitive representation let us first follow the logic of that possibility. If prototypes are in the cognitive representation, there is a sense in which they can be said to be private; my prototypical *dog* or *red* need not be identical to anyone else's and only I can know what my prototypes are like. There are two possibilities regarding this privateness: one in which the representation of the prototype is *inherently* private and the other in which both public and private criteria are applicable.

Suppose the representation to be inherently private. To set up the meaning for a category term and to remember or use the term subsequently I need simply associate (connect) the term with something in my cognitive representational system – a sensation, image, exemplar, prototype, rule, or anything else that is in it. Then I can use the term; for example, I can keep a record of the occurrence of that thing in my representational system – perhaps I could keep an actual diary (*PI*, 258). And *how* do I associate the word; how do I set up the connection? Perhaps one concentrates one's attention on the cognitive thing while writing down the word and thus impresses on oneself the connection between the word and the thing. But, Wittgenstein explains, "'I impress it on myself' can only mean: this process brings it about that I remember the connection *right* in the future" (*PI*, 258). Being inherently private, I can only use private modes of justifying that I have remembered the connection right and am using the term correctly.

Wittgenstein makes two objections to private justification. The first is that it leads to an infinite regress: "'Well I *believe* that this is the [cognitive thing] again' – perhaps you *believe* that you believe it" (*PI*, 260). The second is that it isn't jus-

tifying. Wittgenstein (*PI*, 258-267) asks us to compare: checking our memory by looking up a word in a real dictionary compared with an imagined dictionary; checking a real timetable compared to checking a remembered timetable; buying several copies of the morning paper to assure ourselves that what it said was true; the difference between the result of an experiment and the result of an imagined experiment; and the difference between what we would call justifying an imagined choice of dimensions for a bridge and imagining what is called justifying the choice of dimensions. This is summed up thus:

Why can't my right hand give my left hand money? - My right hand can put it into my left hand. My right hand can write a deed of gift and my left hand a receipt. - But the further practical consequences would not be those of a gift. When the left hand has taken the money from the right, etc., we shall ask: "Well, and what of it?" And the same could be asked if a person had given himself a private definition of a word; I mean, if he has said the word to himself and at the same time has directed his attention to a sensation. (*PI*, 268)

In the present case, one may object that one didn't mean to claim that kind of strict privacy for one's cognitive representations. But notice that in this example, we had all of the conditions necessary and sufficient for a term having meaning in the object of reference view - we had an object for which the term stands and we had knowing that the word stood for it. But without something more than this we did not have meaning.

Wittgenstein then presents a case where we *lack* the necessary and sufficient conditions for meaning under the object name view yet the word still has meaning:

"Imagine a person whose memory could not retain *what* the word 'pain' meant - so that he constantly called different things by that name - but nevertheless used the word in a way fitting in with the usual symptoms and presuppositions of pain" - in short he uses it as we all do. (*PI*, 271)

To reiterate: Given the necessary and sufficient conditions for meaning under the object of reference view, Wittgenstein has shown a case in which a word met those conditions yet we could not say it had meaning as well as a case in which it did not meet the conditions and yet did have meaning. The next move of the adherent of the object of reference view might be to remove the claim of inherent privacy and instead suppose that both public and private criteria are applicable. Wittgenstein has this to say:

Suppose everyone had a box with something in it: we call it a "beetle." No one can look into anyone else's box, and everyone says he knows what a beetle is only by looking at *his* beetle. - Here it would be quite possible for everyone to have something different in his box. One might even imagine such a thing constantly changing. - But suppose the word "beetle" had a use in these people's language? - If so it would not be used as the name of a thing. The thing in the box has no place in the language-game at all; not even as a *something:* for the box might even be empty. - No, one can 'divide through' by the thing in the box; it cancels out, whatever it is.
That is to say: if we construe the grammar of the expression of sensation on the model of 'object and designation' the object drops out of consideration as irrelevant. (*PI*, 293)

At this point, the proponent of the object of reference view might wish to abandon altogether the claim that the object of reference is in the cognitive rep-

resentation, and place it instead in the world. This is not an unreasonable position; after all, we might argue, when we speak or point it is usually directed to *things* not to hypothetical cognitive representations. We point to public exemplars, signposts, rules telling us how to proceed with a category word. Wittgenstein has this to say:

A rule stands there like a sign-post. - Does the sign-post leave no doubt open about the way I have to go? Does it shew which direction I am to take when I have passed it; whether along the road or the footpath or cross-country? But where is it said which way I am to follow it; whether in the direction of the finger or (e. g.) in the opposite one? - And if there were, not a single sign-post, but a chain of adjacent ones or of chalk marks on the ground - is there only *one* way of interpreting them? (*PI*, 85)

This argument may perhaps be seen in an even purer form when it is applied to the "certainties" of mathematics. Suppose that the mathematical Platonist is right and that we do have some kind of external "archetype" of mathematical truths - of the number series, let us say. Suppose even an extreme case in which mathematical objects can be literally perceived; they are a shadow world in which the answers are already written down faintly:

And continuing the series just means copying them out ... [but] this cannot explain how we know that what we are copying is the correct answer. The Platonist's problem is like that of the schoolboy who cheats. He has to know who has the right answers to copy from. (Bloor, 1983, p. 86)

Or in Wittgenstein's words, there are "an infinity of shadowy worlds ... we don't know which of them we're talking about" (*LFM*, 145).

Wittgenstein is often taken as simply attacking privacy, that is, mentalistic interpretations of phenomena such as meaning. In fact he is equally critical of public, external objects, rules, pictures, behaviors or anything else claimed to be the object of reference or necessary meaning of language or knowledge. Simply changing the putative object of reference - whether prototypes are substituted for definitions or exemplars for prototypes or rules of probabilistic inference for exemplars - or changing the location of those objects from private to public or to some combination of both will have no effect upon the arguments.

What Might be a Wittgensteinian Investigation of Categorization?

Wittgenstein does not leave us without any account at all. Language and knowing are part of the activities, language-games, meanings in use, conventions, and forms of life of an entire people. "Only in the stream of thought and life do words have meaning" (*Z*, 173). What implications might this have for the study of categorization?

Wittgenstein left the idea of conventions and forms of life, perhaps deliberately, unspecified. Almost any proposal that one tries to derive from it seems to violate the delicate and precise balance of understanding that Wittgenstein

creates. For example, from the pervasiveness of context effects one might want to claim that categories are created anew each time on the spot (Barsalou, in press); but then, as Barsalou himself points out, one must still distinguish between categories that are already part of the language like *chair* and those created by a particular novel goal like *things to carry out of the house in a fire*. One might wish to argue with Gleitman, Armstrong and Gleitman (1983) that there is no single field of categorization to be studied; but then organisms still do treat discriminably different objects and events equivalently. It might be tempting to argue that the criterial definition view of categories has no place at all in cognitive psychology if its philosophical underpinnings have been removed and that we should concentrate on the investigation of processing heuristics. But criterial definitions certainly have a place in our life and culture; indeed, we cling to them tenaciously. Wittgenstein himself acknowledges this in his remarks on the distinction between *justifications* and *symptoms* of category membership. There is even some empirical evidence about justification (Landau, 1982).

There are at least three current programs of research in categorization that have a somewhat Wittgensteinian flavor (see also Neisser, in press). Perhaps the narrowest of these is my own work on basic objects (Rosch, 1978; Rosch et al., 1976). In taxonomies of material objects there is often one level of abstraction at which we feel we have the true category, the real name of the object (*chair* as opposed to *furniture* or *kitchen chair*). The basic object hypothesis was that this level actually maps real world structure; it is the level at which attributes cluster – physical, perceptual, functional, and social. It is here that we have our most basic language-games with the everyday material world. It is at this level that similar motor movements are used when interacting with the objects in a category; this level at which objects are imaged and at which they are first categorized in perception of the environment; and this is the earliest level categorized by children and the most necessary level for vocabulary in languages. The idea of basic level applies also to social objects (Brown, 1986; Cantor & Mischel, 1979; Tversky & Hemenway, 1984). Presently, work on the development of basic level categories as a flexible negotiation between the mother, child, and environment is being carried out by Mervis (1984; also Mervis & Pani, 1980).

Most category learning occurs not in a specialized categorization language-game, but as part of the ordinary events of daily life. Interest in such events in cognitive psychology was spurred some years ago by the use in artificial intelligence of the concept of scripts (Schank & Abelson, 1977). A programmatic study of the development of scripts and children's knowledge of events is being carried out by Nelson (in press). Nelson argues that language itself can be seen as arising from participation in scripted events and that it is generalized event representations which are the building blocks of cognitive structure. Analysis of the relation of language and events at a more micro level has proved quite generative in developmental psycholinguistics (Slobin, 1981).

Being human, our language-games have a special quality – we have theories. Carey (1982) argues that what may look like a single category such as animal is actually part of an entire theory of living things; it is part of our beliefs about

biology. As early as age four, children who usually say that a mechanical monkey is more similar to a person than is a worm, when told that people have a spleen inside them, readily attribute the spleen to the worm but not the monkey. This cannot be accounted for, Carey argues by reference to a simple similarity metric; it requires a knowledge of folk biology. In fact it is possible to argue, though this may be committing the opposite of one of those skeptical errors, that all categories are theories of sorts (Murphy & Medin, in press).

In conclusion, one might wish to ask - how to proceed with this research? Can we look to Wittgenstein for advice, if not on topic, at least on strategy? Perhaps not, but Wittgenstein's own tactics might be described as perpetually trying to find the correct "focus" from which to analyze a problem (as Merleau-Ponty, 1962, claims we always do in the very act of perception). If one gets "too close" to the problem, looking for what "lies hidden" beyond the uses of language, one's subject matter begins to look queer - and perhaps metaphysical. "In order to find the real artichoke, we divested it of its leaves" (*PI*, 164). If one fails to get close enough, Wittgenstein asks "Do you mean *this?*" giving a further definition, a distinction, a specific case, or constructing a language- game. For Wittgenstein the criterion for having reached the right focus point is the disappearence of the philosophical problem: "For the clarity that we are aiming at is indeed *complete* clarity. But this simply means that the philosophical problems should *completely* disappear" (*PI*, 133). For empirical issues, perhaps this means rather that the problems are brought completely into view. And that is the beginning of the investigation.

References

Anderson, R.C., & Ortony, A.(1975). On putting apples into bottles. - A problem of polysemy. *Cognitive Psychology, 7,* 167-180.

Anderson, R.C., Pichert, J.W., Goetz, E.T., Schallert, D.L., Stevens, K.V., & Trollip, S.R. (1976). Instantiation of general terms. *Journal of Verbal Learning and Verbal Behavior, 15,* 667-679.

Armstrong, S.L., Gleitman, L.R., & Gleitman, H.(1983). What some concepts might not be. *Cognition, 13,* 263-308.

Barclay, J.R., Bransford, J.D., Franks, J.J., McCarrell, N.S., & Nitsch, K.(1974). Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behavior, 13,* 471-481.

Barsalou, S.W. (in press). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts reconsidered: The ecological and intellectual bases of categories.* Cambridge: Cambridge University Press.

Bartlett, F.C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge: Cambridge University Press.

Battig, W.F., & Montague, W.E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology, 80* (Monograph Supplement 3, Part 2).

Beller, H.K. (1971). Priming: Effects of advance information on matching. *Journal of Experimental Psychology, 87,* 176-182.

Berlin, B., & Kay, P.(1969). *Basic color terms: Their universality and evolution.* Berkeley: University of California Press.

Bloor, D. (1983). *Wittgenstein: A social theory of knowledge*. New York: Columbia University Press.

Brown, R. (1979). Cognitive categories. In R. A. Kasschau & C. N. Cofer (Eds.), *Psychology's second century: Enduring issues*. New York: Praeger.

Brown, R. (1986). *Social psychology: The second edition*. New York: The Free Press.

Brown, R., & Lenneberg, E. (1954). A study in language and cognition. *Journal of Abnormal and Social Psychology, 49*, 454-462.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York: Wiley.

Bruner, J. S., Olver, R. R., & Greenfield, P. M. (1966). *Studies in cognitive growth*. New York: Wiley.

Cantor, N., & Mischel, W. (1979). Prototypes in person perception. *Advances in Experimental Social Psychology, 12*, 3-52.

Caramazza, A. (1979). *Are concepts ill-defined or vague?* Paper presented at the Conference on Word and Concept, Stanford University, Stanford, CA.

Carey, S. (1982). Semantic development: The state of the art. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art*. Cambridge: Cambridge University Press.

Erickson, J. R., & Jones, M. R. (1978). Thinking. *Annual Review of Psychology, 29*, 61-90.

Fodor, J. (1981). The current status of the innateness controversy. In J. Fodor (Ed.), *Representations*. Montgomery, VT: Bradford Books.

Glass, A. L., & Holyoak, K. J. (1975). Alternative concepts of semantic memory. *Cognition, 3*, 313-339.

Gleitman, L. R., Armstrong, S. L., & Gleitman, H. (1983). On doubting the concept 'concept.' In E. K. Scholnick (Ed.), *New trends in conceptual representation: Challenges to Piaget's theory?* Hillsdale, NJ: Erlbaum.

Hampton, J. A. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning and Verbal Behavior, 18*, 441-461.

Hampton, J. A. (1985). *Overextension of conjunctive concepts: Evidence for a unitary model of concept typicality and class inclusion*. Unpublished manuscript.

Heider, E. R. (1972). Universals in color naming and memory. *Journal of Experimental Psychology, 93*, 10-20.

Kelley, D., & Krueger, J. (1984). The psychology of abstraction. *Journal for the Theory of Social Behavior, 14*, 43-67.

Kessel, F., & Bevan, W. (in press). In C. W. Buxton (Ed.), *Points of view in the modern history of psychology*. New York: Academic Press.

Lakoff, G. (in press). *Women, fire and dangerous things: What categories tell us about the nature of thought*. Chicago: University of Chicago Press.

Landau, B. (1982). Will the real grandmother please stand up? The psychological reality of dual meaning representations. *Journal of Psychological Research, 11*, 47-62.

Marx, M. H. (Ed.) (1968). *Theories in contemporary psychology*. New York: Macmillan.

McCloskey, M. E., & Glucksberg, S. (1978). Natural categories: Well defined or fuzzy sets? *Memory and Cognition, 6*, 642-672.

McCloskey, M. E., & Glucksberg, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology, 11*, 1-37.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Review of Psychology, 35*, 113-138.

Merleau-Ponty, M. (1962). *Phenomenology of perception*. London: Routledge & Kegan Paul.

Mervis, C. B. (1984). Early lexical development: The contributions of mother and child. In C. Sophian (Ed.), *Origins of cognitive skills*. Hillsdale, NJ: Erlbaum.

Mervis, C. B., & Pani, J. R. (1980). Acquisition of basic object categories. *Cognitive Psychology, 12*, 496-522.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology, 32*, 89-115.

Murphy, G. L., & Medin, D. L. (in press). The role of theories in conceptual coherence. In U. Neisser (Ed.), *Concepts reconsidered: The ecological and intellectual bases of categories*. Cambridge: Cambridge University Press.

Neimark, E. D., & Santa, J. L. (1975). Thinking and concept attainment. *Annual Review of Psychology, 26,* 173–205.

Neisser, U. (in press). From direct perception to conceptual structure. In U. Neisser (Ed.), *Concepts reconsidered: The ecological and intellectual bases of categories.* Cambridge: Cambridge University Press.

Nelson, K. (in press). Social cognition in a scripts framework. In L. Ross & J. Flavell (Eds.), *The development of social cognition in children.* Hillsdale, NJ: Erlbaum.

Osherson, D., & Smith, E. E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition, 9,* 35–58.

Piaget, J. (1972). *The principles of genetic epistemology.* New York: Basic Books.

Posner, M. I. (1969). Abstraction and the process of recognition. In G. H. Bower & J. T. Spence (Eds.), *The psychology of learning and motivation.* New York: Academic Press.

Posner, M. I., Boies, S. J., Eichelman, W. H., & Taylor, R. L. (1969). Retention of visual and name codes of single letters. *Journal of Experimental Psychology Monograph, 79* (1).

Potter, M. C., & Faulconer, B. A. (1979). Understanding noun phrases. *Journal of Verbal Learning and Verbal Behavior, 18,* 509–521.

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior, 12,* 1–20.

Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language.* New York: Academic Press.

Rosch, E. (1975a). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104,* 192–233.

Rosch, E. (1975b). The nature of mental codes for color categories. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 303–322.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and Categorization.* Hillsdale, NJ: Erlbaum.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573–605.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382–439.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures.* Hillsdale, NJ: Erlbaum.

Slobin, D. K. (1981). The origins of grammatical coding of events. In W. Deutsch (Ed.), *The child's construction of language.* London: Academic Press.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review, 81,* 214–241.

Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General, 113,* 169–197.

Vygotsky, L. S. (1962). *Thought and language.* New York: Wiley.

Waisman, F. (1979). *Wittgenstein and the Vienna Circle.* New York: Barnes & Noble.