Flexibility of Categorization in Children and Adults

Gedeon O. Deák[1]

Part I. Introduction

Categorization is treating two or more discriminable entities as equivalent. Cognitive psychologists typically study categorization by looking at sorting, labelling, identification, and membership judgments. However, categorization includes any behavior that is generalized across several different entities or items. Flexibility of categorization is categorizing entity [a] with set [X] on one occasion, and with a different set [Y] on another occasion. Flexibility can be distinguished from stochastic errors in categorization if recategorizations do not occur randomly but are predicted by pre-specified variables and forces. These forces affect both the likelihood and nature of recategorizations.

Because research and theory of categorization has seldom focused on flexibility, there are several outstanding questions regarding flexibility and recategorization: What is the evidence of flexibility in human categorization? What variables govern recategorization? Are there substantial changes in the flexibility of categorization from early childhood to adulthood? This last question is important because an examination of the development of flexible categorization will address question such as, To what extent does flexible categorization rely on extensive factual knowledge? Before addressing these questions, however, it is necessary to delve into some thorny theoretical issues in order to frame these questions and establish the importance of flexible categorization.

What is most noticeable about flexibility is its underrepresentation in both theory and research. In cognitive psychology the modal description of concepts and categories[2] is fairly uniform.

---

[1] Unpublished Master's Thesis, University of Minnesota, 1994.

[2] Most authors use these terms interchangably (assuming that "categories" refers to mental categories), leading one to wonder whether concepts and categories are the same or different. In fact, concepts are sometimes viewed as subsets of categories. That is, a concept is a singular representation of a unitary entity, and a psychological category is either a concept or a group of concepts. For example, the category "dog" refers to the concept "dog," but the category "animal" refers to the concepts "dog," "cat," "horse," etc. Other authors treat concepts as mental abstractions of objectively existing categories. Both distinctions are vague and vagarious, and I therefore will use the terms interchangeably. This decision is of no real consequence for my purposes. Similarly,

Concepts are treated as unitary, fixed entities consisting of intensions (definitions) and extensions (referents), such that the intension completely determines the extension (e.g., Katz & Fodor, 1963). The intension may entail necessary and sufficient conditions, probabilistic conditions, or a combination of necessary and probabilistic conditions. In any event, the intension is the criterion for category membership, and once acquired it is static. This means that an entity either is or is not an instance of a concept, although goodness-of-membership may be graded (i.e., some entities are better members of a category than others; Rosch, 1973; however, see Armstrong, Gleitman & Gleitman, 1983). Furthermore, many theorists (e.g., Smith & Medin, 1981) distinguish between a concept's core and its identifying function. The core is its true intension or essence, and purportedly is difficult to access. The identifying function is the process by which we decide whether an entity is a member of a category. The identifying function, then, is responsible for graded membership or typicality.

Accounts of this sort imply two important features: category membership is fixed and determined, and categories have fixed represention (fixed, that is, within a relatively constrained period of the development, although presumably few would disagree that these might change as the organism acquires new knowledge). Equally important is the (usually implicit) assumption that taxonomic systems are fixed and determined. For example, Keil's (1979) ontological tree and resulting M-constraint implies that our concepts are organized in a static hierarchical taxonomy. Even Barsalou (1991), who has extensively studied the lability of categories, suggests that ad hoc categories, that is, categories that are constructed on the spot for a particular purpose, are deviations from a static primary taxonomic system. Of course, each theorist presents the details differently, and some theorists oppose this basic framework (e.g., Lakoff, 1987; Olson, 1970), but the picture is fairly consistent.

Given these presuppositions, it is perhaps less surprising that so little work has focused on flexibility of categorization: if categories are stable and membership is determined by a core

---

no systematic distinction is made between categorization, classification, and generalization, since no deductively rigorous proposal has been put forward to differentiate these.

definition, then an entity fits into only one category. Of course, Fido is both a dog and an animal, but this is a relatively simple qualification because all dogs are animals, and the heirarchy which includes these concepts is static. In fact, many proposed semantic systems (e.g., Collins & Quillian, 1969; Keil, 1979) are hierarchically organized. This satisfies the intuition that concepts are interrelated, but it does not impart much flexibility because class-inclusion hierarchies only promote deductive inferences in one direction, that is, in ascention of the hierarchy. For example, the knowledge that all sparrows are birds promotes the inference that all sparrows have feathers, because all birds have feathers. Hierarchies do not, however, promote other kinds of common categorization judgments, for example inferences in which the direction (ascention or descention) is not specified, inferences that are not supported by inclusion relations (for example, those involving matrices), analogies, metaphors, and associations. So while this slightly caricatured summary of the dominant framework in cognitive psychology begins to explain the absence of flexibility research, it does not justify the absence, because it has led to the neglect of certain very interesting category-related judgments. In order to better understand this framework, then, as well as to underscore the potential importance of an emphasis on the flexibility of categorization, we must turn to the philosophy of concepts and categories.

A brief history of philosophy

Modern philosophy of concepts began 2,300 years ago. Plato suggested that things have essences: the chair that I am sitting on, for example, is a reflection of some metaphysical essence of "chairness." Our perceptions of objects are biased, degraded sketches of objectively-true category cores. (Note the overt dualism!) Although Platonism is still influential, our modern folk and scientific beliefs in essences are perhaps more directly attributable to Aristotle. Aristotle considered essences "final causes," or average forms from which individual instances deviated. This is the philosophical definition of essentialism: individual entities are imperfect variants of a true core of some category. In general, essentialism imposes a limitation on flexible categorization, because it implies that every entity belongs to a primary, objectively-correct category.

How is this relevant to cognitive psychology and theories of categorization? To begin with, we can distinguish metaphysical essentialism (e.g., Plato) from epistemological essentialism. Epistemological essentialism contends that we acquire ideal, essential forms either through empirical observation and abstraction (as Aristotle believed), through innate knowledge (Fodor, 1975), or both (Quine, 1977). One may, of course, believe in metaphysical essentialism but not in epistemological essentialism: John Locke, in Essays Concerning Human Understanding, suggested that natural kinds (i.e., classes occurring in nature such as animals and rocks) have essences, but humans are incapable of objectively knowing those essences. Although classes are truly alike in fundamental ways, naming reflects a mentally-constructed connection or association. This view is called nominalism. Nominalism holds that the fundamental unit of apprehension is not the class but the individual entity. Classes are therefore sets of individuals, each one of which must be treated as such. According to Oldroyd (1986), the nominalists believed "that universals were logically superfluous. Furthermore, humans could classify objects and qualities in any number of different ways" (p. 32). Thus, the dichotomy between (epistemological) essentialism and nominalism is central to any consideration of flexibility of categorization.[3]

In modern philosophy essentialism has its most powerful foothold in theories of natural kinds. Both Putnam (1975) and Kripke (1972) suggest that natural kind terms are "rigid designators," that is, they refer to a set of entities with a common essence. This common essence is not discernable by necessary and sufficient features (see Wittgenstein, 1953)--indeed, it may not be knowable at all (Putnam later disclaimed these views--see Putnam, 1988).

Although essentialism is the predominant view in academic psychology and folk psychology (a point to which I will return presently), the view is not without its detractors. Besides the nominalists, essentialism has been disputed by the Pragmatists, for example Dewey (1920) and Poincaré (1905). The pragmatists believed that scientific taxonomies are constructed for particular purposes, and it is fallacious to consider a classification "true" or "false" (Oldroyd, 1986). This is

---

[3] The metaphysical question is probably unanswerable, but it deserves mention because essentialist epistemology presupposes essentialist metaphysics.

nominalism with a twist: humans create taxonomies that best suit their purposes. Pragmatists turn a functionalist eye to categorization, allowing assumptions of flexibility.

Other critics of essentialist theories point out that it is simply incorrect to view, for example, biological kinds as having essences (Lakoff, 1987; Palmer & Donahoe, 1992; Sober, 1985). There are several reasons for this. First, the average phenotype of a population is not static; it changes as selection pressures shift and as genetic variation propagates from generation to generation. Second, variability between individuals in a population is not "noise," "error," or deviation from an essential average. Variation is a necessary causal force in evolution, since a population without variability would be decimated by small environmental changes. In fact, mechanisms for genetic variability (e.g., crossing over in meiosis) are themselves fundamental adaptations. Third, there are a number of instances that violate the species distinction--for example, cases in which two populations are morphologically identical but cannot interbreed, and cases in which two populations can interbreed but are morphologically distinct.

On the whole, it is difficult to defend pure essentialism. Nevertheless, it is very prevalent in our thinking: Medin & Ortony (1989) and Gelman & Medin (1993) have even proposed a label for this belief: psychological essentialism, the folk belief that categories are unified by underlying similarities. They assert that representations are based on belief in essences, and people look for causal theories to justify essences. Does essentialist thinking emerge full-blown, or does it develop? Keil (1989) asserts that preschoolers' concepts initially consist of probabilistic features and gradually shift to defining features, implying that psychological essentialism develops during early and middle childhood.

Medin and his colleagues are correct that people seem to believe in essences, at least for natural kinds. Contrary to Medin & Ortony's (1989) assertion, however, there is little compelling evidence that people <u>behave</u> as if things have essences. A person might believe, "A rose is a rose," yet place a real rose in a floral arrangement on Monday, and with a box of chocolates on Tuesday. Most of the evidence cited by Keil (1989), Medin & Ortony (1989), and others (e.g., Malt, 1988) comes from verbal tasks in which subjects are told fantastic stories (about, for example, birds that metamorphosize into insects) and asked to report what they would believe/say/do in that situation (e.g., Keil & Batterman, 1984; Rips, 1989 defends this method). The validity of such self-reports rests on subjects' ability to predict their own reactions to a hypothetical, bizarre situation. I suspect that this ability is quite unreliable.

On the whole, radical essentialism does not look like a realistic option. The critical question generated from philosophical debate, then, is this: Do people's concepts consist of a core that is static and that promotes one fundamental way of categorizing an entity, or are categorization schemas generated to achieve certain goals, implying that an entity can be categorized numerous ways? Before examining empirical evidence pertaining to this question, it may be fruitful to discuss the a priori likelihood of flexibility: that is, what does an organism gain from flexibility, as opposed to stable, immutable knowledge?

The ecology of categorical flexibility

Imagine an organism that can quickly learn and generalize new facts about the environment; that is, it is flexible. If this organism does not tend to learn and remember predictable regularities in the environment, it will be at the mercy of circumstance because its generalizations are not constrained. It is as likely to make spurious predictions based on idiosyncratic similarities as well as reasonable predictions based on important ones. Now imagine another organism possessing fixed, static associations that correspond to specific regularities in its environment. This creature is unable to exploit alternative, potentially important similarities. It is therefore extremely brittle; the poor beast cannot adapt to its environment.

Clearly, there is a trade-off between well-learned, stable knowledge, and flexibility. Stable knowledge allows inductive inferences: if something has fur and barks, it probably also likes meat and chases cars. Flexibility allows adaptation to unpredictable contingencies: if it has fur and barks it can be petted, unless it is growling or foaming at the mouth. Probably there are many factors that influence the trade-off between flexibility and stability. For instance, habit or social convention may compel use of well-learned information, whereas novelty or problem-solving situations may compel more flexible, creative use of information. Emotional investment in a theory or belief system may prevent consideration of alternative associations. Some educational or training experiences may lead to dogmatic rigidity, others to creative analogizing. Such factors suggest potentially exciting inquiries into the economics of categorization.

Although the tension between flexibility and stability is seldom acknowledged in the literature, it is indirectly addressed by several theorists. Barsalou (1982, 1989, 1991) suggests that we have a primary and secondary categorization system. The primary system governs object recognition and categorizes objects into static taxonomies. The secondary system generates ad hoc categories that fulfill specific goals, for example, "Inexpensive vacation spots for summer skiing." In this theory, radical flexibility is limited by a core of well-established knowledge. Another such theory is suggested by Harnad (1987), who proposes a three-tiered representational system, in which sensory input simultaneously forms an analog iconic representation and a digital categorical representation that extracts relevant invariants. Labels generated by the categorical representation system can recombine into symbolic representations. Symbolic representation, then, imparts flexibility via recombination of fundamental categorical elements. In Harnad's theory, as in Barsalou's, flexibility is secondary and derived.

Such proposals allow flexibility without sacrificing the comfortable familiarity of traditional semantic frameworks. Unfortunately, there is little research evaluating these models, thus they are quite speculative. Another breed of models, however, allows categorical flexibility more naturally. Network models assume that large semantic systems are richly interconnected, allowing a variety of associations. An early semantic network model (Collins & Loftus, 1975; Loftus, 1975) represented concepts and properties as nodes. Activation of a node spreads to adjoining nodes based on the degree of activation and decay rate, and the strength of connection between the activated and adjoining nodes. If multiple nodes are activated, intersections along the spreading paths of activation create a semantic "popout effect." This model can accommodate conceptual combination (a form of flexibility discussed in Part IV). Its major shortcoming is its lack of clarity about what a node represents. If a node is a discrete concept, then the model has difficulty accommodating the fact that word meanings change with context (e.g., Lakoff, 1987; Olson, 1970). If context effects are accounted for by intersections along paths of activation, it is unclear what is contained in nodes. The exact nature of activation and propagation of activation in such networks (see Ratcliff & McKoon, 1981; Rosch, 1975b) is also unclear.

A more promising network model is suggested by McClelland & Rumelhart (1985). Their parallel distributed model does not impart meaning to nodes; it is an associationist network that learns by reducing differences between input and output patterns over successive presentations of a stimulus. McClelland & Rumelhart (1985) showed that their system was able to produce patterns of output corresponding to a category "name" after receiving patterns corresponding to discrete exemplars from several categories. Of course, this was a "toy" network, and no attempt was made to allow recategorization or flexibility. However, the connectionist approach should be compatible with flexibility. That is, stable knowledge can be obtained by repeated presentations of relatively invariant input exemplars. Flexibility, on the other hand, may be achieved via inputs from "context" and "goal" units. The question is, will lawful and predictable patterns of flexibility emerge from such a network? There are many problems with such a proposal: artificial networks are not symbol grounded systems, the input and output are highly artificial, and it is unclear whether goals should be treated as input (since they are at least partially self-generated). Further, the architecture of such a system would necessarily be somewhat arbitrary. Most problematic with this proposal, though, is the fact that we do not have enough data about human flexibility to evaluate the performance of such a model. Consequently, it seems premature to consider serious attempts at modelling flexibility.

<u>Evidence for flexibility: three sources</u>

The purpose of this paper is to evaluate evidence regarding categorical flexibility, and attempt to draw specific conclusions about the extent and nature of flexibility in children and adults. Most of the evidence cited is indirect, since few critical experiments have been conducted. Most of the direct evidence for flexibility will be considered in the last section, Part IV. Part II explores one type of indirect evidence for flexibility--the range of different kinds of categorical organizations used by children and adults. Part III explores age-related changes in the kinds of criteria or sorting principles used by children. Since Parts II and III concern indirect evidence for flexibility, I will not attempt a thorough literature review, but rather a discussion of the relevance of selected studies to the development of categorization flexibility.

Part II. Range of available structures

Most studies have examined one kind of categorical structure: the taxonomic hierarchy. Researchers often assume that categorical information is organized into taxonomic trees. A similar assumption is made by developmental psychologists: although they question whether, and at what age, children can form hierarchical representations, they treat hierarchies as the natural end-state of development. There are of course good reasons for studying hierarchically-organized categorical knowledge. Hierarchies are economical: properties and relations are represented at their most inclusive level and are inherited by all subclasses at lower levels. Further, hierarchies allow convenient shifting among different levels of generality. For example, it is not always necessary to speak at the level of the species (to use Linnaean terminology); sometimes it is more sensible to speak at the level of the genus or the phylum.

These facts do not justify a priori assumptions that all information is organized into hierarchies. Actually there are a number of possible ways to organize categorical information--a fact seldom considered by categorization researchers. Several categorization schemes are described in the Appendix.[4] This "taxonomy of taxonomies" is not meant to be comprehensive, but to convey the fact that not all categorization systems are hierarchical. Other structures occurring in linguistic

---

[4] Note that there are two structurally distinct kinds of hierarchies. This distinction is not generally recognized, but it has important implications for the inheritance of properties and flexibility. With regard to flexibility, the "prioritized matrix" hierarchy, in which each successive level of the hierarchy is defined by a sortal feature or set of features that cuts across the entire horizontal level, may be easier to recategorize (by treating it as a matrix and "leapfrogging" a sortal up or down from one level to another level) than the Linnaeus-type hierarchy. The Linnaeus hierarchy introduces different sortals for different branches of the tree at the same level. For example, within the mammals, different orders are distinguished by different features: Bats (order Chiroptera) have particular morphological adaptations involving the hand and arm, face and mouth, and wings. These and other features distinguish them from the Insectivores and the Rodents. The Insectivores (order Insectivora), on the other hand, are distinguished from Rodentia by features such as leg anatomy, teeth, diet, and protective adaptations. In other words, within the mammalian orders there is no single distinguishing feature. Different features are diagnostic for different orders, and the set of diagnostic features depends on the relevant contrast classes.

categories are described by Lakoff (1987). Notice that the systems sketched in the Appendix and in Lakoff (1987) only describe the interrelations among categories in a category system. Categories also have internal structure (e.g., Barsalou, 1989; Lakoff, 1987; Medin, 1983, 1987; Mervis & Rosch, 1981). The internal structure describes the patterns of similarities and differences among members within a category.[5] Notably, not all categories have similar internal structures.

The fact that there is no single blueprint for category structure, nor a single kind of category system, has provocative implications for questions of flexibility. If humans exhibit different kinds of category structures and different kinds of category systems, it is possible that humans can embed the *same* entities in categories of different structures and in different category systems. On the other hand, it is possible that each entity has a "best" categorization, by virtue of which groups of related entities congeal into a "best" category system structure. In this case, alternative systems and structures may never be used or be used only with considerable effort. This view should be recognizable as essentialism, but more specifically it is objectivism, the view that objective discontinuities in the world are reflected in our categories. Lakoff (1987) musters persuasive arguments against objectivism (see also Putnam, 1988), and champions an alternative, "experiential realism." This is the view that we cannot induce objective categories, and our categories do not necessarily reflect objective continuities. Instead, categories are rooted in the contingencies of our biological and sociocultural reality. In this view, categories can be shared meaningfully to the extent that humans share biological characteristics and sociocultural experiences. Experiential realism can also accommodate flexibility and recategorization because human experience is not unitary or entirely predictable, and different experiences call for different construals of similarities and differences in the environment.

The range of category structures in adults

It is clear that adults can organize information into hierarchical taxonomies. This does not necessarily mean that such structures are commonly used outside of scientific nomenclature, but in

---

[5] Of course, the internal structure of a category is not independent of the structure of the system in which it is embedded. For example, if a category [A] is distinguished from a contrast category [B] by features (a, b, c), then that feature set determines membership in [A].

fact they generally seem to facilitate cognitive processing in adults, at least in Educated English-speaking adults. Collins & Quillian (1969) found that subjects verify sentences faster when the predicate and subject come from closer hierarchical levels (e.g., "A canary can sing") than further hierarchical levels (e.g., "A canary has skin"). The authors suggested that property information is stored at an appropriate level of abstraction, and verifying facts at different levels of abstraction requires moving vertically between levels, which takes time. Alba, Chromiak, Hasher & Attig (1980) found that subjects accurately report category frequency information even when this task is incidental. This implies that subjects automatically encode a superordinate category name when presented with a subordinate exemplar (a consistent finding is reported by Warren, 1972). Of course, there are alternative explanations for these data, but they suggest that subjects spontaneously consider the class-inclusion relations between words, at least in some contexts.

Numerous studies (e.g., Bower, Clark, Lesgold & Winzenz, 1969; Tulving & Pearlstone, 1966) have demonstrated the advantage of hierarchical categorical organization in recall. Two effects are robust and well-replicated. First, subjects recall more words from lists organized hierarchically (perhaps because words associated in some way, for example according to taxonomic relations, prime or cue each other). Second, subjects often "cluster" categorically-related words in recall; that is, words from one category are recalled sequentially. These findings indicate that adults find it fairly easy or "natural" to organize related words into hierarchies. At the very least they show that hierarchical organization is easy for adults to comprehend and reproduce.

Very few studies have explored adults' use of non-hierarchical categorical systems. Broadbent, Cooper & Broadbent (1978) presented subjects with lists of words organized in either hierarchies or matrices. Both kinds of organization facilitated recall equally well. The two did differ slightly, however: recall of hierarchical items depended more upon recalling superordinate names, whereas recall of matrix items did not depend on recalling relevant dimension-value labels. The fact that this study is perhaps the only one to examine the facilitating effects of a matrix structure on memory is puzzling, given the usefulness of matrices in organizing information (consider the periodic table). This neglect is exacerbated by the fact that there are actually several different kinds of matrices (besides those listed in the Appendix) that embody interesting distinctions. For example, dimension values can either be opposites (e.g., animals vs. not-animals) or non-opposite contrasts (e.g., animals vs. plants).

Several studies have explored subjects' tendency to organize information in associative systems. Some have focused on sets unified by a common script or theme, for example, "activities involved in skiing." Bower & Clark-Meyers (1980) presented subjects with a list of words organized into either script-related subsets or random arrangements. Subjects in the script-organized condition showed a considerable advantage in recalling and recognizing list words. Although the authors did not measure clustering in recall, foils related to scripts were falsely recognized in the script-organized condition, whereas high-associate words were falsely recognized in the random condition. Apparently subjects in the former group used scripts to organize the information and/or cue recall. Rabinowitz & Mandler (1983) compared the advantages of taxonomic organization to schematic organization in memory (schematic relations are defined by scripts or activities). Both organizations facilitated recall and clustering, but the schematic organization was more effective than the taxonomic organization. Khan & Paivio (1988), however, qualified this result by eliminating a confound (Rabinowitz & Mandler used the superordinate schema name as an item name in each schema), and finding no advantage of schematic organization over taxonomic organization.

It is clear from these studies that organizing information in terms of activities, events, or themes facilitates adult recall, and herein lies an interesting implication for flexibility. Specifically, schematic groupings have a vertical component (the activity or event is a superordinate concept) but no class-inclusion structure or effects (e.g., property inheritance). This suggests that adults can organize information in a number of ways, but they may need (or be aided by) a superordinate concept or idea. This does not mean, however, that a superordinate always helps. For example, Barsalou (1985) explored ad hoc categories, or categories that are constructed "on line" to fulfill a goal. Examples of ad hoc categories are "Ways to escape being killed by the Mafia" and "Things to take out of the house in case of fire."[6] Barsalou compared the effectiveness of common (taxonomic) and ad hoc categories in facilitating recall. He found that organized lists of ad hoc category members were not recalled better than random groups of words. Furthermore, when cued by the ad hoc category name, recall was barely better than recall of random words lists. In contrast, taxonomic categories showed typical free- and cued-recall advantages. However, this finding may not contradict the "superordinate" generalization: some of the ad hoc categories used were "has a smell," "can be walked on," "is a liquid," and "is manufactured by humans." These are unusual ad hoc categories--it is difficult to imagine what goals would compel their formulation. Even if a goal can be imagined, it is not clear that Barsalou's subjects made any effort to imagine a situation for using these categories. Furthermore, these categories are also taxonomic categories, since "taxonomic category" means, in its broadest sense, a group of items that all share some feature. Categories such as "liquids" are idiosyncratic only because they are uncommon and probably not widely useful (e.g., what function is shared by apple juice and bleach?). Perhaps the "superordinate rule," then, should specify the importance of a functional superordinate.

Finally, the prevalence of simple categorization systems should not be underestimated. Simple categories may have one sortal (i.e., an attribute that differentiates a category from its contrast class) or relevant dimension. For example, colors (see Heider & Olivier, 1972) are differentiated primarily by wavelength. Medin, Wattenmaker & Hampson (1987) constructed stimuli that could be sorted according to family resemblance (average within-category similarity minus average between-category similarity), or a single binary feature (e.g., short vs. long tail). Across a variety of stimulus sets, instructions, numbers of features, and feature variability ranges, subjects sorted according to a single dimension. Even when specifically instructed to attend to all features, given a story designed to compel family-resemblance sorting, and shown stimuli constructed to discourage unidimensional sorting, most subjects sorted according to one dimension (although a few used a simple conjunctive or disjunctive contingency rule). Only one condition led to a sizable minority of

[6] Presumably these categories are given names only for experimental purposes. That is, ad hoc categories should not merit names unless they are used repeatedly, at which point they are no longer ad hoc.

family resemblance sortings (9 out of 24 subjects). In that condition, some stimuli (artificial animals) possessed features diagnostic of "birdness" (e.g., feathers, beaks, webbed feet), and subjects were told that half the animals were "fliers" and half were "nonfliers." Clearly, subjects only used a family resemblance strategy when they were supported in making an analogy to a real, complex category (birds). This study is noteworthy because many researchers, following Rosch (1978, Mervis & Rosch 1981) have assumed that natural categorization proceedes by abstraction of family resemblance information (i.e., clusters of correlated features). Medin et. al. (1987) implies that in some situations a much simpler sorting strategy is more natural. The study, in conjunction with Kemler Nelson's research on family resemblance concept acquisition in adults (Kemler Nelson, 1984, Smith & Kemler Nelson, 1984) begins to sort out contingencies that do and do not induce more complex (family resemblance) categorization.

We must conclude that adults have a wide range of useable category system structures. Although I have not described all relevant structural variables, the basic conclusion should be clear. The fact that adults can utilize many category system structures allows the possibility that adults can categorize a group of instances in a variety of ways. It does not, however, evince that possibility, since most studies do not test recategorization. The exceptions are Broadbent et. al. (1978), Rabinowitz & Mandler (1983), and Khan & Paivio (1988), all of which used identical lists of items organized into alternative structures. These studies directly address questions of flexibility, since they each examine the effects of experimental contingencies on subjects' organization of a set of items. They certainly confirm the minimal position that related items can be recategorized, and the method of organization depends on task contingencies. However, they do not address other substantive questions, such as, (a) Is any categorization context-free (i.e., fixed and static); (b) Do some relationships among entities compel specific category system structures; and (c) How can we describe the conditions leading to recategorization? Unfortunately, there is currently almost no research that can address these questions. There is, however, another body of research relevant to questions of stability vs. flexibility. This is the work on basic-level categories, which will be reviewed shortly.

The range of category system structures in young children

Much of the research on children’s category structure has focused on acquisition of taxonomic hierarchies. It is often assumed that adult knowledge is represented hierarchically, and that children learn these (objectively correct) hierarchies as they progress into middle childhood. One question, then, is how early hierarchical knowledge is represented. This question has no simple answer, and a full discussion would be too lengthy for this paper. Nevertheless, we will attempt to narrow down the age at which some hierarchical knowledge is discernible.

Inhelder & Piaget (1964) asserted that true class inclusion relations are a hallmark of middle childhood. In class-inclusion relations a superordinate class [A] consists of two or more subordinate classes [B] & [C], such that all members of the latter are also members of the former. Inhelder & Piaget asserted that the addition and subtraction of classes necessary to simultaneously consider the superset and subsets depends upon concrete operations. This was supported by Piaget's well-known class-inclusion tasks in which children are asked to compare the size of a subset to its superset. For example, subjects are shown five roses and three daisies and asked, "Are there more roses or flowers?" Children younger than eight often fail this question. However, this task involves more than simply appreciating class-inclusion relations, and a flurry of research activity followed Piaget's claim. Trabasso, Isen, Dolecki, McLanahan, Riley & Tucker (1978) surveyed the fallout in a well-reasoned task analysis of class-inclusion research. Their conclusions are well worth noting:

> "preoperational children...can make quantitative comparisons between superordinate and subordinate classes. The major factor promoting successful performance seems to be unambiguous reference to the class as a whole and avoidance of procedures stressing distinctive features of the subordinate class.”
>
> “If we are interested in class concepts and their hierarchical structure, class inclusion is not a good place to start because it confounds perceptual, quantification, and decision-making skills" (pp. 177-178)

To the list of confounds we might add linguistic and pragmatic knowledge. Winer (1980) reaches a similar conclusion, although he is more reticent about granting preschoolers representation of class-inclusion relations.

The use of converging methods has reaffirmed Trabasso et. al.'s (1978) first conclusion. Smith (1979) gave children several tasks including inference questions (e.g., "A pug is a kind of dog. Does a pug have to be an animal?"). Her results indicated that first graders and kindergartners understood inclusion relations. Most preschoolers succeeded on at least one task, although their performance was more variable. It is unclear whether this was due to linguistic difficulties, fragile grasp of inclusion relations, or both. A different procedure was used by Steinberg & Anderson (1975). In a cued recall task with 6-year-olds, they found that for a target word ("apple"), a more distant superordinate ("food") was a less effective cue than a closer superordinate ("fruit"). These findings suggest that by 5-to-6 years children can comprehend and utilize class-inclusion relations.

There is less consensus about preschoolers' comprehension of inclusion relations. Markman, (1984; Callanan & Markman, 1982; Markman, Horton & McLanahan, 1980) suggests that knowledge of class-inclusion hierarchies begins as appreciation of collection relations (e.g., "tree-forest") rather than inclusion relations (e.g., "tree-plant"). Her data suggests that children occasionally misconstrue superordinate classes as superordinate collections. This misconstrual may be significant, since collection hierarchies do not allow inheritance of properties in the same way as class hierarchies. Nelson and her colleagues (e.g., Lucariello, Kyratzis & Nelson, 1992; Lucariello & Nelson, 1985; Nelson, 1986) assert that information about subordinate classes is embedded in knowledge of events. “Slot-filler” categories consist of subordinates that fulfill the same function in an event (e.g., macaroni and cheese, hot dogs, and sandwiches are members of "lunch foods"). Only as the child reaches the early school years are these categories reorganized into taxonomic categories. Evidence suggests that young children prefer slot-filler categories to taxonomic categories (Lucariello et. al., 1992; Lucariello & Nelson, 1985).

One problem with both of these propositions is that empirical evidence has relied on linguistic tasks. Even in the least "language-heavy" studies cited above, children are required to generalize labels. It is unclear how children understand such a task, and whether class-inclusion relations can be assessed in a non-linguistic context. Until such a task is devised, the lower age limits of class-inclusion understanding are doomed to be confounded with linguistic abilities.

Relatively few studies have explored preschoolers' use of matrix categories. Inhelder & Piaget (1964) hypothesized that simultaneously attending to two or more variables requires concrete operations. However, their data showed modest and irregular age-based improvement on a two- and three-variable matrix completion task--for example, more 4-to-5-year-olds solved a three-factor matrix task that did older children. The authors suggest that preschoolers and 7-to-8-year-olds use different strategies to solve the problem: preschoolers respond on the basis of overall similarity and then justify their responses by describing the perceptual features of their choice item, whereas older children reason about features in order to make their choice. Although the authors contend this pattern is "obvious" (pp. 155-157), it is difficult for a non-Piagetian to discern qualitative differences between the sample verbal protocols of younger and older children. It is not clear whether an objective distinction between older and younger children's responses can be made.

Halford (1980) provided explicit training to 3-, 4-, 5-, and 6-year-olds on a one- and two-variable matrix completion task . All except the three-year-olds learned the two-variable task within five trials. Halford concluded that by four years, the two-variable task was not more difficult than the one-variable task. Denney & Acito (1974) report a similar finding with even younger children. They tested two- and three-year-old children (non-spontaneous sorters only) on a multiple classification test. They then modelled two-dimensional sorting and asked children to sort the modelling set as well as a transfer set. All children sorted the modelling set along two dimensions, and most also sorted the transfer set along two dimensions. Thus, four-year-olds are capable of sorting objects into a two-dimensional matrix, and younger children probably are also capable although they do not do so spontaneously.[7] This conclusion should be treated cautiously, however:

---

7 Sharp, Cole & Lave (1979) found little spontaneous two-dimensional sorting at first-, third-, and sixth-grade levels, suggesting that sorting tasks using geometric stimuli rarely compel spontaneous two-dimensional sorting.

all of these studies used geometric forms varying on shape, color, and size. This limits the generalizability of conclusions about the prevalence of matrix classification in young children.

Associative categorization structures are probably at least as prevalent in young children as adults. Numerous theorists (e.g., Inhelder & Piaget, 1964; Vygotsky, 1962) have asserted that preschoolers' categories are organized around perceptual item-to-item associations, which Vygotsky termed "unorganized congeries." More recently, many theorists (e.g., Denney, 1974; Nelson, 1982) have suggested that preschoolers' categories are organized around events or schemas. Schematic structures have superordinate events or schemas, but the hierarchical organization is not mutually exclusive as in a class-inclusion relation: "bone" may be associated with "dog" and also with "soup." Many studies have demonstrated appreciation of thematic relations in children from 1-to-5 years: Ackerman (1988) in a cued recall task, Bauer & Mandler (1989), Markman & Hutchinson (1984), and Smiley & Brown (1979) in a forced-choice oddity task, and Mandler, Fivush & Resnick (1987) in an object manipulation task. What is unclear, however, is whether these associative clusters should be considered categories, or simply groups of high-associate items. This issue will be addressed in Part III.

Finally, it is clear that preschoolers can construct simple categories. Denney & Acito (1974), Halford (1980), and Sharp, Cole & Lave (1979) have found single-dimension sorting in children as young as 2-to-3 years. The only qualification to this generality is the hypothesis that preschoolers are holistic--that they treat separable features as integral and have difficulty selectively attending to such features (e.g., Shepp, Burns & McDonough, 1980; Kemler, 1983). However, this qualification must be one of degree rather than kind, since some evidence (Smith, 1979, 1984) shows that 3- and 4-year-olds can perceive separable dimensions (although two-year-olds may not). Deák & Pick (in review) have argued that holistic categorization may be a byproduct of the unmotivating stimulus materials and tasks used in traditional discrimination learning studies. Such findings do not negate the assertion that preschoolers can categorize on a single dimension, they merely implicate factors affecting children's tendency to do so.

Just as a wide range of category structures are available to adults, a range is available to children as young as 3-to-5 years. Having said this, we must admit the same caveat that followed the discussion of adults' structures: the fact that children have several structures available does not mean that they will apply more than one to the same set of stimuli. If particular sets of items compel particular organizations, it would seriously qualify categorical flexibility. However, the finding that a particular set of stimuli compels a particular organization in a specific situation does not qualify flexibility, because flexibility is not expected to be random or unconstrained. Rather, the task, context, and relations among the to-be-categorized items will dictate the manner in which they are grouped. Evidence of this kind of flexibility is reviewed in Part IV.

Basic-level phenomena

Rosch and her colleagues (e.g., Mervis & Rosch, 1981; Rosch, Mervis, Gray, Johnson & Boyes-Braem, 1976; Rosch, 1978) changed the course of Cognitive Psychology by asserting the importance of what they called "basic-level categories."[8] Basic-level categories were originally described as "...categories at the level of abstraction that maximizes cue validity and maximizes category resemblance [and] best mirror the correlational structure of the environment" (Rosch, 1978, p. 31). An impressive convergence of evidence supports Rosch's theory: subjects list relatively more shared attributes at the basic level than at a more inclusive (superordinate) or less inclusive (subordinate) level, and they describe common motor patterns for basic-level objects (Rosch et. al., 1976). Canonical, regularized two-dimensional outlines of basic-level exemplars overlap more than those of superordinate exemplars, and two-dimensional shapes averaged from two exemplars are more identifiable at the basic-level than the superordinate level (Rosch et. al., 1976). Items are identified more quickly at the basic-level (e.g., Jolicoeur, Gluck & Kosslyn, 1984; Murphy & Smith, 1982; Rosch et. al., 1976). Parents usually name objects for children at the basic level (Callanan,1985; Mervis & Mervis, 1982) and children interpret novel object names as basic level (Callanan, 1989). Finally, young preschoolers ($2^1/_2$ years) are better able to distinguish basic-level categories than superordinate or subordinate categories (Mervis & Crisafi, 1982) .

---

[8] The concept and phenomenon were previously suggested by Brown (1958).

The weight of evidence motivates the conclusion that there is something basic about basic-level categories. Does it also support the conclusion that basic-level categories are stable cognitive "reference points" reflecting predictable continuities in the environment? The answer is a definitive "sort of." I will briefly mention some qualifications of the primacy (hegemony?) of basic-level categories. First, as Rosch (1978) suggests, the basic level partly depends on experience, culture, and expertise. In support, Palmer, Jones, Hennessy, Unze & Pick (1989) report that the basic level of musical instruments differs for musicians and non-musicians. This means that discontinuities and correlated attribute clusters in the environment do not completely determine basic-level categories.

Second, some basic-level terms do not have a clear superordinate (e.g., stair, railing, door, seashell, stump, truth, beauty, love, etc.), others do not have obvious subordinates (e.g., icecube, headlight, crow, giraffe), and still others lead to inclusion fallacies (e.g., an ornamental sword is a sword but not a weapon). It is unclear, in these cases, what "basic-level" means, since these items may not fit into a hierarchy. Finally, there are usually taxonomic levels other than subordinate, basic, and superordinate, and intermediate levels may differ from person to person (Rosch, 1978).

Third, it is difficult to specify what is special about basic level categories. Murphy (1982) points out that cue validity, as often defined, cannot logically underlie the basic-level advantage.[9] Medin (1983) notes that the basic level cannot simultaneously maximize within-category similarity and minimize between-category similarity, because the former criterion calls for the most specific category, whereas the latter calls for the most general category.

Fourth, the finding that basic-level names are processed (i.e., verified and named) most quickly is qualified by two findings. First, Jolicoeur et. al. (1984) found that atypical exemplars do not show a basic-level advantage over the subordinate level. A similar finding is reported in White's (1982) investigation of parental naming and preschoolers' overextensions. Second, Murphy &

---

9 Murphy's (1982) argument is valid for the original definition of cue validity (Rosch et. al. 1976); that is, the conditional probability that an object with some feature [a] is a member of category A; or $N_A|N_a$. It can be argued that a more defensible and reasonable definition of cue validity is the ratio of the proportion of A possessing [a] to the proportion of non-A possessing [a].

Wisniewski (1989) convincingly demonstrate that the basic-level advantage disappears when pictures of objects are presented in scenes. In the context of scenes, subjects can verify superordinate-level names as quickly as basic-level names.

Finally, the primacy of basic-level words in acquisition is not absolute. Mandler and her colleagues (Mandler & Bauer, 1988; Mandler, Bauer & McDonough, 1991; Mandler & McDonough, 1993) have shown that within-superordinate distinctions are acquired later than between-superordinate distinctions. Thus, global contrasts (e.g., animals vs. vehicles) may be acquired before basic-level contrasts. Callanan (1989) found that preschoolers generalized a novel word at the superordinate level when it was taught on multiple exemplars.

What can we conclude from all of this? It appears that there is something special about certain words for things. For example, in many situations we are likely to name things at a particular level of abstraction. This tendency partly reflects patterns of continuity and discontinuity in the environment, and partly reflects linguistic convention. Further, the psychological basic level depends on experience, expertise, and context. Therefore, whereas it is often easy or most natural to generalize things at a particular (intermediate) level of abstraction, this habit is by no means immobile. Unfortunately, there are relatively few studies of conditions that mitigate basic-level effects. This topic deserves more attention: the partial flexibility of basic-level effects (in childhood and adulthood) is extremely important because it constitutes exactly the sort of compromise needed to characterize an organism that is neither rigid nor overreactive, that is, and organism with a stable knowledge-base that can use its knowledge flexibly.

<u>Internal structure of categories</u>

A great deal of research has explored the internal structure of categories, that is, the patterns of similarity and differences among members of a category (see Smith & Medin, 1981). There are currently a number of important and unresolved issues about intra-category structure: features and dimensions, generalization from prototypes vs. exemplars, characteristic and defining features, etc. One issue is particularly central to the flexibility and range of intra-category structure: the nature of similarity. Similarity was long taken for granted, and only recently has it been intensely scrutinized.

Notably, Tversky (1977; Tversky & Gati, 1978; see also Goodman, 1972) revolutionized thinking about similarity by arguing that the similarity between two objects is not static, but dependent on context and task factors. Recently, Smith and her colleagues (Jones & Smith, 1993; Smith & Heise, 1992) have also argued that similarity is flexible. Whereas Smith, like many others, represents similarity as distance in a multidimensional space, she views that space as elastic: certain dimensions may be elongated or foreshortened, depending on the knowledge, goals, and information available to the subject. In addition, Medin, Goldstone & Gentner (1993) argue that similarity is meaningful only when it is highly constrained--that is, similarity is always with respect for some attribute(s) defined by a task. In this sense, similarity is flexible and non-arbitrary. Importantly, Medin et. al. (1993) consider similarity judgments to be the end-product of a decision process. These principles are consistent with nominalism/pragmatism and inconsistent with an objectivist epistemology (see Part I), and consequently are supportive of a functionalist approach to categorization.

A small body of empirical work points to the flexibility of within-category structure. In adults, Barsalou (1989, 1992) has demonstrated that the graded structure (prototypicality) of commonly-tested categories (e.g., birds, furniture) is variable both between individuals (see also Coltheart & Evans, 1981) and, more importantly, within the same individual (see also Bellezza, 1984). Typicality judgment differences between individuals point to the role of experience, and challenge an objectivist epistemology (Barsalou finds strikingly low between-subject correlations--between .2 and .5). Within-subject typicality judgments (from different times) show less variability, but inter-test correlations are between .5 and .8. This finding has significant implications for the mobility of category structure. It implies that category information is not completely static--saituational factors affect judgments about category membership.

The effects of context also have been tested by Barsalou (1989), who found that asking subjects to take different points of view (e.g., that of a Chinese student) led to quite different typicality ratings. This manipulation, however, is a strange operationalization of "context." It may simply show that beliefs about Chinese citizens' experience guide hypothetical judgments. Future studies may manipulate context more effectively by changing the experimental task and situation in principled ways in order to demonstrate responsivity of internal category structure (several studies, e.g., Medin et. al., 1987; Kemler Nelson, 1984, have shown context effects in concept structure acquisition). Other studies (e.g., Roth & Shoben, 1983) have investigated the effects of context on concepts and word meanings; these will be considered in Part IV.

Barsalou's (1989, 1992) demonstrations that typicality judgments are mobile raise questions about local factors affecting category structure. Barsalou (1989) theorizes that category judgments are affected by invariant category knowledge, local context effects, and long-term context effects. However, he does not explain how these factors are combined, nor does he offer empirical justification for positing two kinds of context effects. Barsalou's contention that some information about categories is always evoked or activated, no matter what the context, is tested in Barsalou (1982). In Experiment 1, he had subjects verify ostensibly stable and context-independent property-noun links (e.g., "skunk" and "has a smell") or context dependent links ("roof" and "can be walked on"). The former were verified more quickly than the latter. However, it should be noted that a quicker reaction time in one out of two sentential contexts does not mean the property is <u>always</u> evoked by the concept. In addition, presenting words and phrases might invoke knowledge that is not invoked when the actual referent of the word is encountered. This issue is often overlooked by researchers who treat words as direct instantiations of concepts. In fact, evidence that pictures and words draw upon common, underlying representations is equivocal (Snodgrass, 1984).

An alternative approach to context sensitivity and flexibility is proposed by Medin & Schaffer (1978). In their theory, inducing category membership depends on perceiving clusters of shared attributes among exemplars. An important aspect of this theory is exemplar instantiation. That is, different exemplars may be called up for comparison depending on the context. Since different exemplars have different attributes, membership judgments depend on which exemplar is evoked. This theory is promising for the investigation of flexibility in categorization. However, the original theory proposed a simplistic exemplar retrieval scheme: retrieved exemplars are those most similar to the to-be-categorized exemplar. Based on what we now know about similarity, this proposal needs to be elaborated to accommodate flexibility. It may be fruitful to modify the context theory by integrating aspects of semantic memory theory (e.g., Tulving, 1983): showing, for example, that different exemplars sharing different similarities are evoked in different task situations could justify integrating context theory with a functionalist approach to categorization.

Finally, there is some relevant work on the development of category structure in children. Recently, Keil (1987; Keil & Batterman, 1984) has suggested that the features children use to decide category membership shift from characteristic/probabilistic (during the preschool years) to defining (during middle childhood). A different theory is advanced by Clark (1973): children progressively add features to their semantic concepts, leading to ever-sharper and more constrained concepts. Yet another theory is suggested by Nelson (1974; Lucariello & Nelson, 1985), who argues that early categories are rooted in schemas or events (e.g., "dinner foods," "bathroom items") and only later are taxonomic kinds decontextualized from event frames.

There is probably some truth to all of these suggestions. Keil's (1987) and Clark's (1973) theories, broadly speaking, suggest that as children acquire knowledge it changes their concepts. That can hardly be a controversial point. Nelson's theory suggests that learning about objects occurs in the context of events, and category structure develops as a result of abstraction across numerous experiences and events. Again, this can hardly be controversial. Two question become critical: first, what is the nature of the complementary roles of these two kinds of learning? Second, what are their implications for flexibility? The first question can only be answered empirically, and is beyond the scope of this paper. The second question, however, is conducive to analysis. The role of increasing declarative/factual knowledge may be twofold. On one hand, it suggests static categorization: as a person learns more defining (or highly characteristic) features, these take on greater importance and "disable" alternative (moderately characteristic) sortals. On the other hand, additional facts constitute additional sortals, so that characteristic features, rather than defining features, may be used for categorization. For example, learning that certain kinds of dogs (e.g., Dobermans, Rottweilers) tend to be used as guard and attack dogs may cause me to differentiate these breeds from others (e.g., St. Bernards, Retrievers) when considering which dogs are most appropriate for families with small children. A similar duality arises when considering the proposal that category knowledge is organized around events. To the extent that a particular kind of event is relatively homogeneous from instance to instance, the categories derived from its constituents will be relatively stable. To the extent that instances of an event differ as a result of causally-important variables, the derived categories will be relatively context-dependent. This suggests the intriguing albeit speculative possibility that the adaptive value of an organism's flexibility depends on the overall variability of its environment.

Part III. Range of categorization criteria in children

It is often suggested that young children are limited to categorize according to certain attributes, or prefer to use certain attributes. Such "response hierarchy" hypotheses (so-called because they imply an ordering of criteria from most-useable to least-useable; see Deák & Pick, 1994) have serious implications for the development of flexibility in categorization: if a set of stimuli hypothetically can be sorted according to any one of several criteria, but some of these cannot be apprehended or utilized by the child, flexibility is limited. I will review the evidence for several proposed restrictions or preferences that have received considerable attention.

Preferences among physical attributes

Some physical dimensions may be more salient than or preferred to others. Indeed, many discrimination learning and classification experiments find some effect of stimulus dimension, although such effects are often small and hard to interpret. One of the most sophisticated treatments of "stimulus asymmetries" is found in Garner (1983). Garner draws two broad conclusions. First, it is quite difficult to disentangle the effects of dimensional salience and discriminability. For example, a preference for shape may be eliminated by making color more salient. Second, dimensional asymmetries are somewhat task-specific, making generalization almost impossible.

A related point is raised by more recent research. Landau, Smith & Jones (1987) found a shape bias in preschoolers' label generalization to novel objects varying in shape, size, and texture. In later work (Jones, Smith & Landau, 1991; Jones & Smith, 1993), the researchers found that texture becomes important when eyes are added (presumably preschoolers learn that texture is diagnostic of animal kinds). In addition, reflectance (manipulated by adding "glitter" and a spotlight) becomes more salient when the word is presented as an adjective and objects are presented under a spotlight on a stage. These findings demonstrate that the salience of various dimensions depends not only on their discriminability and the task, but also on correlations among attributes and the context in which objects are presented. The picture begins to look very complicated indeed, and simplicities like "children attend to shape" appear naive. The mere fact that young children can selectively attend indicates the ability to utilize a variety of attributes. However, the fact that preschoolers' selective attention is less effective than older children's (Lane & Pearson, 1982; Smith, 1989) suggests that flexibility of categorization undergoes development.

Perceptual vs. conceptual attributes

Luminaries from every tradition in cognitive development have suggested that preschoolers are "perceptually bound"--that is, they can only use overt physical similarities, not conceptual information, to make categorization judgments (Bruner, Olver & Greenfield, 1967; Inhelder & Piaget, 1964; Kendler & Kendler, 1965; Vygotsky, 1962; Werner, 1957). Agreement among such a diverse group of theorists is noteworthy. It is therefore remarkable that they all were incorrect. To be sure, there is a seed of truth in the observation that preschoolers are less likely to use non-obvious information to make decisions under uncertainty (i.e., when making a difficult choice). After all, preschoolers have less factual knowledge than adults, and their verbal skills are less sophisticated. Furthermore, as we recede along the ontogenic scale, these differences become more acute. Thus, while the three-year-old has factual and lexical knowledge that can be brought to bear in a variety of circumstances, the 12-month-old is probably consigned to rely upon learned expectations about the physical and social world. These expectancies, however, are by no means simple, nor do they regard only static objects and properties.

This qualification aside, it is clear that at a very young age (2-to-3 years) children can make generalizations on the basis of non-immediate, non-physical information. This has been nicely demonstrated in a series of experiments by S. Gelman and her colleagues (e.g., Gelman & Markman, 1986, 1987; Gelman & Coley, 1990). Preschoolers were shown line-drawing "conflict trios" consisting of a standard item, an item that looks like the standard, and an item with the same category name as the standard (e.g., blackbird (standard), bat, and flamingo). Children were told that the flamingo feeds its young mashed-up food and the bat feeds its young milk. They are then asked whether the blackbird feeds its young mashed-up food or milk. Preschoolers draw an inference from the flamingo to the blackbird when the pictures are labelled ("bird," "bird," and "bat"), and from the bat to the blackbird when the pictures are unlabeled. This suggests that preschoolers prefer to make inductive inferences on the basis of category names rather than overall appearance .

We have extended this finding in several ways (Deák & Bauer, in review (a, b)). In several experiments four-year-olds made taxonomic matches (e.g., blackbird & flamingo) more often than chance, even when items were unlabeled. One critical factor was the use of three-dimensional objects instead of line drawings. Apparently preschoolers can ignore salient physical similarities (e.g., overall shape and color) when subtle but diagnostic physical similarities are available. For example, the fact that both blackbirds and flamingos have feathers may override differences in size, shape, and color. Another significant factor is the use of inference questions (e.g., "Which one has cold blood, like this one?") instead of classification questions (e.g., "Which one is the same kind of thing?"). Apparently, inference questions allow the child to select the information that is most important for a particular decision. For example, if asked to make an inference about an animal's diet, the child might look at its teeth (sharp or flat?) and feet (hoofs or claws?). A more neutral sorting question, on the other hand, provides no such clues about what information is important. Nevertheless, Deák & Bauer showed that classification questions can elicit taxonomic-based choices when instructions and training clarify the task for preschoolers. This throws new light on previous studies (e.g., Fenson, Cameron & Kennedy, 1988; Melkman, Tversky & Baratz, 1981; Tversky, 1985) that supported the perceptual boundedness assumption: such studies may have presented tasks that compelled similarity-based responding rather than taxonomic-based responding, even though both modes are available to preschoolers.

These studies point out the difficulties of making simple generalizations about the information children use to make categorization judgments. The fact that at least two response strategies (taxonomic-based and appearance-based) are available to preschoolers and are selected on the basis of task contingencies implies flexibility in preschoolers' categorization abilities.

Taxonomic vs. thematic relations

According to Inhelder & Piaget (1964), preschoolers do not sort taxonomically. Younger preschoolers groups items as various similarities strike their fancy, and older preschoolers put things together on the basis of associations (e.g., dog and bone, spider and web). Since then, several researchers have noted preschoolers' interest in thematic relations (e.g., Denney, 1974; Nelson, 1982; Smiley & Brown, 1979). Denney (1974) suggested that the shift from an emphasis on complementary or thematic relations to taxonomic relations takes place between 6-to-9 years, with variability accounted for by the complexity of the material. Contrary to Denney's claim, younger children appreciate taxonomic relations even when they conflict with thematic relations. Smiley & Brown (1979) found that although young children preferred thematic relations to taxonomic relations, first graders correctly justified the non-preferred taxonomic choice.[10] Moreover, preschoolers could be trained to choose taxonomic over thematic relations in a transfer set. Markman & Hutchinson (1984) reported a preference for taxonomic over thematic relations when the standard item was labelled. This preference obtained at the basic-level for 2-to-3-year-olds and at the superordinate level for 4-to-5-year-olds. More surprisingly, Bauer & Mandler (1989) reported that 1-to-2-year-olds could sort either thematically or taxonomically, based on instructions and social reinforcement of a particular kind of response.

---

10 In many studies no effort is made to unconfound similarity of appearance from taxonomic relatedness. Thus, when children group a cow with a dog rather than milk, it remains to be determined whether the decision was based on appearance or on "animalness." Thus, such choices may be based on appearances.

These findings demonstrate that both thematic and taxonomic relations are available to children as young as 1-to-2 years, but it is possible that thematic relations are preferred at some ages. This hypothesis was tested by Greenfield & Scott (1986) who found a thematic relation preference at every age from 3 to 14 years, and a significant quadratic trend reflecting an increase in thematic choices between 4-to-6 years. The preference hypothesis is nevertheless difficult to accept, since choice preference depends on so many conditions (labelling, reinforcement, taxonomic distance, and probably many more) and has only been shown in experimental contexts. On the other hand, perhaps it could be argued that the experimental tasks are decontextualized and therefore reflect a default bias--one which underscores a response hierarchy and a limitation on flexibility. This argument, however, is fallacious: the triad oddity task is highly artificial and it embodies a very specific and unusual context, not a "default" context. Such tasks should be used to test competence, not performance, since the vagaries of such tasks are significant.

Finally, it is important to note that thematic relations are far from homogeneous. Indeed, they are often poorly defined, since they sometimes they seem to be based on strong associations ("dog-bone"), sometimes on causal relations ("key-door"), and sometimes on place co-occurrence ("lake-beach"). These distinctions are potentially important: Ackerman (1988) tested children's and adults' cued recall for taxonomic, thematic, place co-occurrence, and ad hoc relations. Children's recall was most facilitated by taxonomic and thematic relations, and least facilitated by place co-occurrence relations. Thus, the rubric "thematic" is often applied to a variety of relations that may be more or less effective for children. These relations may not be considered categories per se, although to the extent that they are evoked in particular contexts (e.g., thinking about items needed for a child's birthday party) they may become goal-derived categories.

Shape and function

The suggestion that natural language object classes are defined by shape (Clark, 1973) or function (Nelson, 1974, 1982) has engaged many researchers. Often form and function are pitted against each other, in spite of the fact that function follows from shape, and thus it is somewhat artificial to separate the two.[11] Nevertheless, to the extent that they can be separated, researchers have tested the relative importance of shape and function in a variety of ways.

Several studies have concluded that young children use shape to generalize object labels, whereas older children and adults use function. Gentner (1978) showed subjects two objects with different names, functions, and shapes. She then asked subjects to name a third object having the shape of one and the function of the other. Two- to five-year-olds and adults generalized by shape, whereas older children (five to fifteen years) generalized by function. Unfortunately, there were several methodological problems with the study (e.g., order of exposure was not counterbalanced; stimuli were biased to favor the same-shape object). The experiment may have biased subjects to favor shape, as evinced by adults' unexpected preference for shape-based labeling. Similar problems arise in Tomikawa & Dodd (1980). Two- to three-year-olds explored nine wooden toys comprising a 3 X 3 matrix of shapes and functions. Subjects demonstrated a tendency to sort by shape. However, combining every shape with every function probably biased subjects to treat one aspect as irrelevant. Since objects were presented in their static state during sorting, and function usually involves dynamic properties, subjects were probably biased to sort by shape. A more intriguing consideration arises in Prawat & Wildfong (1980), who varied the height-width ratio of line-drawing vessels (creating a continuum from cups to bowls) and included a variable designed to portray function: the pictures portrayed a hand pouring either coffee or cereal into the vessel. Three- to four-year-olds' naming responses were not affected by context, but older children's were. This was taken to disconfirm Nelson's (1974, 1982) theory. It is not clear, however, that function can be manipulated in pictures. Further, the portrayed functional difference is culturally-defined and arbitrary. That is, cups are equally good at containing cereal and coffee, and so are bowls. It is

---

[11] It is also often assumed that shape is concrete and function is abstract. This assumption makes the form vs. function question an operationalization of the perceptual boundedness question, and naturally leads to the hypothesis that younger children categorize by shape and older children categorize by function. However, this assumption is not valid: just as function can be abstract, shape also can be abstract. Whether either is abstract depends on the way it is varied from item to item in a given task (see Deák & Pick, in preparation, for details).

only tradition that dictates the distinction. Thus, all that can be concluded from Prawat & Wildfong is that in pictures of vessels, shape is used more readily than culturally-conventional typical usage.

Problems also arise in studies purporting to demonstrate functional biases. Nelson (1973) concluded that children between 1-to-2 years learn and extend words on the basis of common function. However, her studies are only suggestive due to methodological problems (e.g., small samples, failure to establish subjects' understanding of various attributes).

Two studies suggest that both shape-based and function-based responding are available to young children. Corrigan & Schommer (1984) tested two-year-olds' sorting of artificial objects varying in shape and function. In Experiment 2 they found a robust crossover in which subjects encouraged to sort either by shape or function did so appropriately. Similarly, Deák & Pick (1993; 1994) presented 4-year-olds with trios of objects in which two objects differed in shape and function, and a third (hybrid) object had the shape of one and the function of the other. When subjects were initially instructed to sort according to one criterion (either shape or function), they did so over 80% of the time (for each criterion) even though they were never reminded of instructions or given any feedback. However, subjects in a control condition who received no specific instruction tended to sort by shape. This tendency, however, was a methodological artifact: in a second experiment the function of the hybrid object was demonstrated while subjects made a sorting decision, and the shape-sorting tendency disappeared. Apparently, presenting objects in their static state (first control condition) made shape information more available than function information, leading to more shape-based sorting. When hybrid objects were presented in a dynamic state which emphasized both shape and function, however, children used the two criteria equally often. However, children were quite systematic in their use of a sorting criteria: in every condition and study, individual children were far more systematic than chance in using one categorization criterion. These findings make a very important point, because they demonstrate flexibility across conditions due to task contingencies (brief training and the mode of presentation of the hybrid object during sorting) while at the same time showing systematicity in the consistency of individual children's responses. Thus, flexibility does not mean unsystematic or unbounded, random responding: in children as young as four, it means sensitivity to task factors.

With regard to shape and function, these studies show that preschoolers as young as two years are capable of using both criteria (also see Kemler-Nelson, 1991). Which criterion they choose depends on contingencies of the task-at-hand, such as experimenter-provided instructions and/or encouragement, the nature of the stimulus sets, and whether objects are presented in static or dynamic states.

Conclusion: Flexibility in criterion use by young children

We may safely conclude that the search for one or two criteria that are overwhelmingly important in children's categorization and generalization has failed. This failure cannot be attributed to limitations on researchers' ingenuity or persistence. Rather it appears to be due to young children's stubborn refusal to ignore any important aspect of to-be-categorized items. Indeed, the best conclusion is that subjects as young as two reliably use whatever information is highly diagnostic, salient, and practical in a particular decision-making situation. No doubt there are limits to this conclusion: we would not expect a two-year-old to use a complex system of abstract conditional contingencies to make decisions. Further, some categorization criteria must require particular knowledge or expertise. It is nevertheless clear that preschoolers can categorize objects according to a wide variety of criteria, including physical dimensions and attributes (e.g., color, shape, texture, the presence of eyes), abstract, "conceptual" relations, knowledge-dependent associations (e.g., thematic and taxonomic connections), and function.

As a final note it should be stressed how difficult it is to predict the categorization criterion a particular subject will use on a given occasion. This point is effectively made by Bjorklund & Zeman (1983), who examined children's clustering and recall strategies in children's cued recall of classmates' names. Although they found a high level of clustering in first-, third-, and fifth-graders, within each group there was considerable variation in the criteria used. Across the groups children were found to use sex, race, seating arrangement, and social organization to organize recall. Furthermore, many children used more than one clustering criterion. Apparently they used one criterion (e.g., social group) until it was exhausted, and then switched to another criterion. Interestingly, subjects could not verbalize their strategies: less than half of the fifth-graders could identify the criteria they used to organize recall. This ability to exploit different criteria for organizational purposes apparently undergoes some development. Ceci & Howe (1978) presented 4-, 7-, and 10-year-olds with lists that could be organized both taxonomically and thematically. In free recall, they found that older children were more likely to switch between the two criteria. For example, a child may recall, "House, tent, igloo..." at which point igloo cues a thematic connection and the child continues, "...Eskimo, sled..." Facility in switching between different criteria may be an important source of development in categorization flexibility, perhaps related to Inhelder & Piaget's (1964) observation that younger subjects have difficulty understanding multi-dimensional matrices. It is also probably related to another probable source of developmental differences in flexibility--selective attention (e.g., Lane & Pearson, 1982). Selective attention is closely related to flexibility because it implies control over criterion choice. Thus, if the assertion that selective attention undergoes substantial development during the preschool years holds up (Smith, 1989), it leads to a highly-plausible hypothesis of parallel development of flexibility in categorization. However, few studies have explored selective attention to aspects of real, interesting stimuli (most studies use monotonous, geometric figures) in motivating tasks that resemble (at least faintly) everyday tasks and situations. Until more such research is undertaken using both selective attention tasks and categorization tasks, it will not be clear whether selective attention and flexibile categorization show parallel development and similar underlying processes.

Part IV. Flexibility and recategorization in children and adults

In this section I will examine how contingencies affect subjects' categorization schemes and criterion choices. Put another way, How do external factors constrain categorization decisions? Although much of the research described in Parts II and III is relevant, here I will only discuss strong evidence of flexibility. Strong evidence consists of demonstrations of recategorization (i.e., the same subject categorizing the same stimulus set in different ways). Because there are few such demonstrations, I will extend the discussion to between-subjects demonstrations, in which different groups from the same population categorize the same stimulus set in different ways. This section is divided into four sections: category learning, category sorting & membership rating, categorical inferences, and memory effects.

Category learning

Adults. Several researchers have examined factors affecting what is learned about a category.[12] This is relevant because if a particular mode of learning affects what is learned, flexibility may be limited during the early stages of learning. For example, if a subject initially learns that categories [A] and [B] can be distinguished by the presence or absence of features [a, b], she may not realize that both categories share [c] and can therefore be categorized together under a superordinate relation. This consideration should become clearer as I evaluate the possibility that different learning conditions lead to acquisition of different information.

Medin & Smith (1981) asked subjects to learn categories of schematic faces varying among four binary dimensions. The categories were probabilistic and non-linearly separable (see Medin, 1983, for a discussion of linear separability). Subjects learned to sort training exemplars under three instruction conditions: standard (neutral) instructions, "rule-plus-exceptions" instructions, or prototype instructions. In learning as well as transfer, the prototype instruction was least effective. Analyses suggested that subjects in the other conditions learned specific features of exemplars, whereas prototype instruction subjects did not. Apparently, then, different information was acquired by subjects in different instruction conditions. However, it could not be determined

[12] I am skeptical of the proposition that categories per se are learned (although some formal categories may be learned through schooling). Most knowledge used in categorization may be acquired in the form of correlated contingencies in the environment, sequences of events and expectancies in various situations, and efficient procedures for achieveinggoals. Factual knowledge may be formally learned but is seldom learned in the form, "Category X has property Y." Instead, categories emerge in particular situations, and most categories are derived to fulfill goals. This means that it may be more appropriate to speak of categorization than categories. Consequently, it is very important to figure out how processes of selective attention, information integration, abstraction, generalization, analogizing and metaphor making and understanding drive situation-specific categorization and category construction. The upshot of all of this is that it may be extremely unnatural to present two categories consisting of well-defined exemplars, and ask subjects to learn the categories. In spite of this, I will review research that has used this type of method.

whether exemplar information was unlearned (i.e., not encoded) by subjects in the prototype instruction condition, or whether that information was simply less available.

Medin, Wattenmaker & Hampson (1987) asked subjects to construct categories from stimuli that could be sorted according to either family resemblance or a single dimension. Although they manipulated several variables, including instruction, subjects overwhelmingly sorted by a single dimension. This indicates that a particular task and stimulus structure can induce systematic use of certain kinds of information, regardless of superficial task factors. What is noteworthy about Medin et. al. is the fact that subjects chose a variety of different dimensions for sorting, suggesting that the one-dimensional strategy was not an artifact of one particularly salient feature. Kemler Nelson (1984) also found that subjects learn criterial attribute categories better than family resemblance categories. However, her subjects did so only under conditions of intentional learning. In contrast, subjects learned family resemblance categories better under incidental learning conditions (see also Brooks, 1978). In a related study, Smith & Kemler Nelson (1984) found that adults use overall similarity (similar to family resemblance) rather than criterial attributes under several conditions: when sorting under time pressure (speeded classification), when simultaneously doing mental subtraction (concurrent task load), or when instructed to base classification on overall impressions of similarity. These results contradict theories that preschoolers categorize by overall similarity and adults categorize by dimensional relations or criterial attributes. Apparently, both learning strategies are available to adults, and learning conditions dictate which strategy is used.

Children. Smith (1979) tested the other half of the generalization that children categorize by overall similarity and adults categorize by criterial relations. She presented kindergartners and second-graders with a sorting task in which stimuli varied on two dimensions, and sets could be sorted according to overall similarity or according to values on one dimension (e.g., all [A] are yellow and all [B] are orange). After subjects learned the two categories, they sorted transfer sets that allowed either an overall-similarity strategy or a dimensional-identity (criterial attribute) strategy. Smith found (Experiment 3) that when children are required to discover a sorting principle for the training stimuli, they are more likely to sort transfer stimuli according to dimensional value. This is an important finding, because it indicates that children as young as five are capable of sorting dimensionally, and tend to do so when training allows them to discover the importance of the relevant dimension. However, if preschoolers have trouble spontaneously sorting dimensionally, it may be a consequence of imperfect selective attention (Smith, 1989). And, as previously suggested, young children's difficulty with selective attention implies a limitation of flexibility.

In a study similar to Kemler Nelson's, Kossan (1981) tested concept learning in second- and fifth-graders. She varied the kind of concept learned (characteristic or necessary features) as well as the learning task (intentional or incidental learning). At both ages, criterial attribute categories were learned more effectively in the intentional condition, and characteristic feature categories were learned better in the incidental learning condition. Apparently, there is some generality across ages in the kinds of conditions best suited for learning certain category structures. However, this does not necessarily mean that subjects in one condition learned nothing about attributes that were more-easily learned in the other condition. For instance, in their rationales many children in the intentional condition (50% second-graders, 87% fifth-graders) cited probabilistic features as bases for characteristic categories. In conparison, 33% and 70% of the incidental children mentioned defining features in the criterial attribute condition. In general, it seems that most of the fifth-graders and some of the second-graders were at least partly aware of non-optimal (as defined by their learning condition) bases for category membership . Apparently the use of particular kinds of information in category learning reflects a preference rather than a constraint, qualifying the view that concepts are rigid or static (e.g., either characteristic or defining, see Keil & Batterman, 1984).

Discrimination learning. Because it is not immediately clear how discrimination learning relates to recategorization and flexibility, I will attempt to clarify the relationship. In these experiments subjects are trained to discriminate among several simple stimuli. Usually subjects are given many presentations of stimuli pairs from which they choose one member and receive feedback. When the subject has learned the choice principle, the experimenter presents a new problem to see how quickly the subject learns the new principle and how many mistakes s/he makes. In the intradimensional shift (IDS) paradigm, stimuli vary on two dimensions (e.g., form and color). Subjects first learn to choose one value (e.g., black) of a dimension (color), and then switch to the other value of that dimension (white). This paradigm is compared to the extradimensional shift (EDS), in which the subject must learn to switch to a value on the previously irrelevant dimension (form, e.g., circles).[13]

It is often found that EDS is easier for preschoolers whereas IDS is easier for older children and adults. Some researchers have taken this to mean that preschoolers do not respond on the basis of dimensionality, but instead respond to particular stimulus combinations (see Zeaman & House, 1974, for alternative hypotheses). Since the EDS involves switching responses to half as many stimuli as does the IDS,[14] EDS is easier if subjects learn particular stimuli rather than a dimensional value. That is, the young child is not constrained by a superordinate criterion (the relevant dimension), but by the difficulties of exemplar relearning. Older subjects, on the other hand, prefer to respond within a dimension: once they learn that a dimension is diagnostic, it is easier for them to switch to a different value on that dimension.

This explanation is very neat, but sadly predictably, the story does not hold up well empirically. Sometimes, for example, preschoolers (Brown & Scott, 1972) learn IDS more easily than EDS. Sometimes both younger and older children learn IDS more easily than EDS (Eimas, 1966). More damning is the finding (Caron, 1970; Tighe, 1973) that for preschoolers, EDS becomes more difficult and IDS easier as additional irrelevant varying dimensions are added to the stimuli. Similarly, reducing irrelevant varying dimensions leads to faster EDS in older children (Tighe & Tighe, 1967). Finally, special procedures such as overtraining (Eimas, 1966; Mumbauer & Odom, 1967) render EDS more difficult for preschoolers. Conclusively, a simple developmental pattern does not stand up to close scrutiny. Instead, a host of contingencies affect the ease of within-dimension response-shifting in both older and younger children.

---

13 There are several other paradigms, notably the reversal shift (actually a special case of IDS) and the optional shift. In the interest of brevity these will not be considered separately. See Kendler & Kendler (1975) and Zeaman & House (1974) for more detailed discussion.

14 This is true only of reversal shift problems.

What does this mean for recategorization and flexibility? Responding to particular stimuli, as young children were believed to do, does not evince recategorization, or categorization of any kind. This is because each instance of a particular kind of stimuli (e.g., black circles) is identical--the subject merely needs to learn to respond to two stimuli. Do either EDS or IDS require recategorization, then? In IDS, subjects partition stimuli according to the relevant dimension (e.g., color), and form a hierarchy containing two subsets (e.g., white things and black things). Switching from white things to black things does not require recategorization, it merely involves switching between subclasses. The greater ease of IDS indicates that recategorization is more difficult, since EDS involves forming two new subsets (defined by a different dimension) that cut across the previously-formed subsets. The finding that EDS is often difficult for older subjects therefore suggests a limitation on flexibility; namely that it is easier to stay within a superordinate dimension than to switch between superordinate dimensions. The fact that preschoolers show an IDS preference when there are more irrelevant dimensions suggests that recategorization is difficult for them as well. Does EDS necessarily evince recategorization, then? Unfortunately, no. It is extremely difficult to determine whether EDS-preferring subjects recategorize into new cross-cutting subsets, or whether they simply extinguish previously-learned responses to exemplars and learn new ones. Subproblem analyses, however, suggest that young children follow the latter strategy (Tighe, 1973; Tighe, Glick & Cole, 1971). Younger children's relative preference for EDS, then, may reflect a learning strategy (i.e., learn particular exemplars) that circumvents the need for recategorization, which is difficult for children.

Sorting and membership rating

This section describes studies of children's sorting and adults' typicality (goodness-of-exemplar, or GOE) ratings. Although these judgments are difficult to compare, they are combined for two reasons. First, the paradigms are relatively age specific: there are few GOE studies with young children, and few sorting studies with adults. Second, both methods ask subjects to judge an item's appropriateness as a representative of a category. In sorting, subjects must decide which of several categories each item belongs to,[15] so the membership judgment is made with regard to particular contrast classes. The GOE/typicality ratings are more precise, because they take for granted that an item is a member of a category and seek to establish how representative or paradigmatic the item is. Thus, the two are not completely comparable since they test category membership at different levels--macro for sorting, and micro for GOE.

Adults. Several studies, mostly conducted by Barsalou (see Barsalou, 1989, 1992), indicate that GOE structure is not static: a given subject produces different GOE judgments for the same set of items on two different occasions. This is true of taxonomic categories (e.g., birds, furniture) as well as ad hoc categories (e.g., ways to escape being killed by the Mafia). Barsalou has also shown that GOE ratings differ systematically when subjects take another person's point of view. For instance, Emory undergraduates taking a faculty viewpoint generated GOE structures identical to faculty members taking their own point of view.

Roth & Shoben (1983) tested the effect of the sentential context of a word on its GOE structure. They used a word verification paradigm and sentence reading time to measure GOE (past studies, for example Rosch, 1975a, have established a reliably high correlation between verification time and GOE ratings). In one experiment subjects read sentences that suggested either a target category member or a non-target category member. For example, one sentence contained the phrase, "...milk the animal..." whereas another contained the phrase, "...ride the animal..." The subsequent sentence contained the target word, "cow." The target sentence was read more quickly following the first sentence than the second sentence. Another experiment showed that the typicality ratings of two category members could be reversed by changing the context sentence. For example, although "crow" is a more typical bird than "duck," the latter is verified more quickly following the sentence, "The hunter shot at the bird..." These findings demonstrate that typicality structure is not "unstable," as Barsalou suggests. Rather, it is flexibly stable--it changes in

---

[15] I will include studies using the triad oddity paradigm, since it involves minimal sorting--namely, choosing which of two objects goes better with a standard.

predictable ways with contextual contingencies. In fact, Roth & Shoben contend that these contingencies can actually restructure the internal structure of a category.

This demonstration is buttressed and expanded by research on conceptual combination (Medin & Shoben, 1988; Osherson & Smith, 1982). These studies show that typicality ratings of nouns are changed by modifiers such as adjectives. For example, Medin & Shoben (1988) found that large wooden spoons are judged more typical than small wooden spoons, but the size effect is reversed for metal spoons. Another study pointed out the complexity and sophistication of conceptual combination: Medin & Shoben found that, for example, brass railings and gold railings are rated more similar to each other than to silver railings. In contrast, gold coins and silver coins are more similar to each other than to brass coins. Apparently, the color of metals is important in the context of railings, whereas the value of metals is important in the context of coins. This finding exemplifies flexibility: different aspects of a set change their importance in different situations. Substance, for example, determines objects' color, hardness, density, malleability, rigidity, monetary value, etc. Which of these properties is important depends on the context or task. Conceptual combination, then, emphasizes the dynamic nature of semantic knowledge.

Children. Relatively few studies have examined children's ability to sort the same set of items according to different criteria. Sharp, Cole & Lave (1979) presented geometric figures varying in number, color, and size or shape to schooled Mexican children and unschooled Mexican adults. They asked subjects to sort the cards more than one way and found that the probability of correct reclassification increased with years of schooling rather than age. Fewer than 10% of first-graders reclassified, over 90% of sixth-graders reclassified, and only 50% of adults with fewer than three years of education reclassified. Kalish & Gelman (1992) asked 3-year-olds to sort both usual (e.g., wooden spoon) and unusual objects (e.g., fur bowl, wooden pillow) by texture and by the room they belong in. Subjects correctly sorted along both dimensions. This is one of very few demonstrations of within-subjects recategorization in children as young as three. In other studies, Kalish & Gelman found that 4-year-olds could sort unusual objects according to material when attributes such as fragility and hardness were relevant criteria.

Several studies have shown that preschoolers prefer to classify one way but are capable of using another criterion. For example, Miller (1973) presented first-graders and adults with sets of four objects that could be grouped into two overlapping sets of three items. One triad within a set was unified by a "concrete" relation (e.g., orange, ball, and plum: round), the other by an "abstract" relation (e.g., banana, orange, and plum: edible). Subjects were given two opportunities to sort each quartet. Children initially formed abstract sets on 74% of the trials; adults on 83%. Although Miller does not provide complete data about recategorization, it appears that children recategorized less frequently than adults on only three out of eight sets. Thus, the developmental trend is inconclusive (since Miller made no attempt to equate knowledge or familiarity of the relations unifying the various subsets), but does not appear to demonstrate a qualitative difference in flexibility between older and younger children. In a triad oddity task, Deák & Bauer (in review (a)) demonstrated that training and instructions affected preschoolers' tendency to sort according to either overall appearance or taxonomic relatedness. Deák & Bauer (in review (b)) showed that stimulus type (line drawings or objects) and labeling also affected the proportion of taxonomic vs. appearance-based sorting.

Preschoolers' ability to sort by thematic and taxonomic relations has been demonstrated repeatedly. Smiley & Brown (1979) showed that young children prefer thematic relations, but can be trained to sort a set of stimuli by either thematic or taxonomic relations. Markman & Hutchinson (1984) showed that 2-to-3-year-olds tend to sort by thematic relations when items are unlabeled, but by taxonomic relations when items are labeled. Finally, Bauer & Mandler (1989) found that different instructions and selective social reinforcement compelled 1-to-2-year-olds to sort a set of stimuli either by taxonomic or thematic relations. Other studies have shown that children can sort by both shape and function. Corrigan & Schommer (1984), for instance, showed that 2-year-olds could sort novel objects either by shape or function, and Deák & Pick (1994) and Kemler Nelson (1993) report similar results with older children. In one study, Deák & Pick instructed children to sort according to one criterion, and then, after sorting half of the test trios, re-instructed children to sort using the other criterion, after which children sorted the remaining test trios. Children overwhelmingly sorted according to the trained, appropriate criterion on both the first and second half of the test trios. Switching from one criterion to the other did not appear to detrimentally affect performance on the latter criterion.

The sum of these findings indicates that young children are aware of a number of alternative relations and criteria (as concluded in Part III). More importantly, they have some facility for recategorization and criterion switching. It is not clear whether this ability increases with age, as Ceci & Howe's (1978) results suggest (see Part III), with schooling, as Sharp, Cole & Lave's (1979) finding suggests, or both. Furthermore, it is not clear how sensitive or complex this ability may be. It would be useful to adapt context studies with adults (e.g., Medin & Shoben, 1988) to younger subjects, although it is doubtful that such methods could work with children under four.

Categorical inferences

Inductive inference is an important function of categorical knowledge, and it is one of the most promising means of studying categorical flexibility because it can be used to select specific similarities between disparate items, and because solving difficult inferences sometimes depends on focusing upon subtle or non-obvious information about items. This selecting and focusing is at the heart of recategorization, since recategorizing means changing the relative importance of different sortals. For example, imagine pointing to a sunbathing lizard and telling a child, "That lizard is cold-blooded." Later the child sees a snake sunning itself and tries to decide whether it is also cold-blooded. What information might she use to decide? The answer probably depends on her world-knowledge and beliefs about animal kinds, her memory of the lizard, her preconceptions of similarities between snakes and lizards, her observations of snake and lizard behavior, her trust in your authority, etc. Ultimately, though, to reach a decision she must select some combination of information specific to the question at hand. She might reach a very different decision if asked whether snakes, like (some) lizards, are good climbers.

Several studies have looked at young children's inference-making. Gelman & Markman (1986) asked preschoolers to make inferences that implicate either taxonomic relations (e.g., What do birds feed their young?) or overt physical properties (e.g., What is the texture of cotton?). In a triad oddity task in which subjects extend an inference to either a similar-looking or a taxonomically-related item, some children chose the taxonomic item for the former type of question and the appearance item for the latter type of question. Thus, at least some preschoolers are selective in their categorical inferences: they do not extend all properties to all category members. Massey & Gelman (1988) asked whether preschoolers could correctly make an inference (Can this move up and down a hill by itself?) about a number of different items (e.g., animals, statues of animals, wheeled vehicles, a robot). They noted that out of several different strategies preschoolers might use (e.g., physical similarity, similarity to people, etc), children's answers were based on animacy. Finally, Kalish & Gelman (1992) explored whether three- and four-year-olds make inferences about unusual objects (e.g., plastic light bulbs, fur doors) on the basis of nominal category or on the basis of material. Inferences implying material (e.g., "Will get sodden if put in water") were extended on the basis of material by 4-year-olds, but not by 3-year-olds. Note that preschoolers do not know what "sodden" means, so they are using the semantic content of the sentence to "guess" that the answer has to do with material.

The small literature on inductive inference strongly suggests that young children can recategorize according to meaningful dimensions or attributes. This is most aptly pointed out by Kalish & Gelman (1992). Although there are too few studies to specify the conditions that facilitate inferential recategorization, Deák & Bauer (in review (b)) suggest one possibility. They found that preschoolers made more taxonomic-based inferences when presented with three-dimensional objects than with line drawings. They speculated that these objects contain more physical information, allowing selection of subtle information from which to make inferences. The same study (as well as Gelman & Markman, 1987) showed a significant effect of labelling. Apparently labels, as well as detailed stimuli, can be used to base inferences. Perhaps when more information is available to preschoolers, they are more likely to make adult-like inferences.

Memory research

Adults. There are two sets of findings pointing to categorical flexibility in adult memory. The first shows a recall advantage and recall clustering of categorically-related word lists (see Part II).

Most studies use taxonomically-organized lists, but two reports, Rabinowitz & Mandler (1983) and Khan & Paivio (1988), found that schematically-organized word lists have similar effects. Similarly, Broadbent et. al. (1978) found a recall advantage for matrix-organized word sets. These studies are important because they used the same sets of words in two conditions. The only difference, in each study, was the organization in which the words were presented: subsets based on explicit superordinates (either schematic, matrix or hierarchical) provide a structure that is comprehensible to adults, and one that facilitates retrieval.

The second set of findings demonstrates context effects on cue effectiveness in cued recall. Many studies use some variant of the following procedure: subjects read several sentences after which they are given a list of cues (one for each sentence) and asked to recall the sentence. The critical feature of these experiments is the relationship between the sentence and cue. Subjects hear one of two sentences, either (for example) "The man played the piano" or "The man lifted the piano." During recall the cue is either "Music" or "Heavy." Several studies using this paradigm (Anderson & Ortony, 1975; Anderson, Pichert, Goetz, Schallert, Stevens & Trollip, 1976; Barclay, Bransford, Franks, McCarrell & Nitch, 1974; Greenspan, 1986) have found that the more appropriate cue for each sentence is more effective: "Music" is a better cue for the first sentence, whereas "Heavy" is a better cue for the second. The effect is fairly robust, and Anderson et. al. (1976) showed that the effect does not depend on word associations between the cue and the critical noun (in fact, context can make low-associates more effective cues than high-associates), nor does it depend on associations between the cue and other words in the sentence.

These effects must be considered in discussions of flexibility. The reason is this: theories of semantic knowledge or word meaning often assume static representation (see Part I): words (or nodes) evoke a certain set of information. Contrary to this assumption, these studies demonstrate that in the context of a sentence some connotations of a word are activated more strongly than others. This implies that word meaning is not static. Instead, words consist of a number of connotations and elements, the relative strengths of which depend on context. In some contexts, one element may be highly activated almost to the exclusion of most others. This implication needs to be qualified: recall Barsalou's (1982, 1989) contention that some aspects of word meaning are context independent (always evoked) whereas others are context dependent (conditionally evoked). Greenspan (1986) tested the possibility that cues corresponding to central properties of a word are effective regardless of context, whereas cues corresponding to peripheral properties are effective only in specific contexts. For example, "water" will always be an effective cue for "lake," whereas "reflection" will only be effective in certain contexts. Greenspan's hypothesis was confirmed, but his data do not indicate that central properties are always evoked--only that central properties are effective cues even when they are not especially relevant to the sentential context.

At first glance, Barsalou's (1982) hypothesis about context-independent information appears to be supported by the semantic ambiguity literature. Several studies imply that all meanings of an ambiguous word are initially retrieved or activated, regardless of the word's context (e.g., Conrad, 1974; Foss & Jenkins, 1973; Seidenberg, Tanenhaus, Leiman & Bienkowsky, 1982; Tanenhaus, Leiman & Seidenberg, 1979). For example, "watch" primes "look," even if it embedded in the sentence "I bought the watch." Most studies find this effect early in processing (within 200-400 msec), whereas later in processing context has an profound effect.[16] However, most of these studies use homonyms (e.g., "rose," "club") rather than polynyms (e.g., "bite," "plant," "throw"). This is an important distinction: homonyms possess several distinct and unrelated meanings, whereas polynyms possess strongly related meanings. In fact, there is a continuum ranging from words with several unrelated meanings ("rose") through words with marginally-related meanings ("straw") and words with strongly-related meanings ("plant") to words with such closely-related meanings that it is difficult to specify the differences between meanings (e.g., "there"). It is the latter two regions of the continuum that address flexibility. The fact that both meanings of a homonym are primed within half a second regardless of context has little bearing on the question of whether core connotations of a word are activated regardless of context.

---

[16] Some studies find marginally significant or interactive effects of context early in processing (e.g., Conrad, 1974; Tanenhaus et. al., 1979).

Context effects in semantic processing may have far-reaching implications for flexibility. This is because every word (or phrase) refers to a category. It may refer to a set of concrete objects (object nouns), actions (action verbs), relations (prepositions), etc.[17] A great many words do not always refer to exactly the same set of things (polysemy); furthermore, the referents of a word often are complex and heterogeneous. The denotations of such words are often ambiguous in decision-making: that is, there are several possible aspects of a word's meaning (or connotation) that may be "activated" in order to make a decision. This ambiguity may be resolved by using sentential and conversational context to limit the relevant connotations of a word, and therefore its denotation. For example, a word that refers to a natural kind such as "dog" has many connotations, as determined by our knowledge and associations of the concept. For example, we have experienced many friendly dogs and heard the idiom, "Man's best friend." Yet we also know that dogs are carnivores, are closely related to wolves, and can be trained to guard property and attack strangers. When deciding whether to pet a strange Cocker Spaniel puppy, the "friendly" connotation may be weighed heavily, whereas when deciding whether to pet a strange growling Doberman, the "ferocious" connotation may be weighed heavily. Although this example appears far removed from categorization, there is a connection: recategorization involves exactly the same kind of weighing and re-weighting of different attributes of a set of items.

<u>Children</u>. Although little semantic context research has been done with children, several studies do address categorical flexibility. The Bjorkland & Zeman (1983) study described in Part III demonstrated that children 7 to 11 years spontaneously use various criteria (e.g., age, gender) to organize recall of their classmates' names. Furthermore, most of their subjects switched from one criterion to another at least once during recall. A similar result is reported by Ceci & Howe (1978), who found that 4-, 7-, and 10-year-olds could sort a set of stimuli into both taxonomic and thematic sets, and that recall was facilitated by switching between the two kinds of relations. Significantly, older children demonstrated more switching than younger children. This finding is provocative

[17] Proper names are an obvious exception (although many proper names contain categorical information, such as gender and ethnicity, which may count as connotations). It is unclear whether some parts of speech, such as articles, refer to categories.

because it is perhaps the only demonstration of developmental differences in categorical flexibility, as well as the practical benefits that such flexibility may impart. Sadly, the finding is only suggestive because older children are generally more facile at spontaneously organizing recall (Lange, 1978).

The relative effectiveness of several organizing principles on cued recall was explored by Ackerman (1988; see Part III). Taxonomic and thematic relations were more effective than place co-occurrence and ad hoc relations, although the latter relations also facilitated recall. However, examination of sample stimuli suggests that some organizing principles were not well instantiated. For example, "lily" is cued by "bunny" and "jelly bean" in the thematic category "Things having to do with Easter," and by "candy" and "book" in the ad hoc category "Things to take to the hospital as gifts." I doubt these connections are compelling to second- and fourth-graders. Lucariello & Nelson (1985) compared a word list organized into either taxonomic categories or slot-filler categories (i.e., sets that fulfill the same function in a given script or situation) and found better recall, with more clustering, under the slot-filler organization. Furthermore, the slot name (e.g., "Things you could eat for lunch") was a more effective cue than the taxonomic category name. Finally, Worden (1976) trained groups of second- and fifth-graders to sort a set of pictures either taxonomically or thematically, and then tested recall. Both groups exhibited significant clustering, although clustering and recall was higher in the thematic group. Interestingly, a free-sorting group (allowed to develop their own organization) displayed recall and clustering slightly higher than the thematic group. Analysis revealed that most of these children spontaneously sorted thematically. This study suggests alimitation on flexibility: although subjects may be able to organize a set in several different ways, one way may be preferred, rendering alternative methods less effective. However, it is not known how much, and in what ways, such preferences limit flexibility and impede performance. In fact, it is not clear how such effects could be measured. We are then left with a rather uncomfortable conclusion: preschoolers can categorize sets of items in different ways although their preferences may render some categorization schemes less effective, but we do not know whether the preferences are attributable to the possession and overlearning of specific declarative knowledge, or implicit learning, or both--or, as some researchers might suggest, more or less determinate differences in sensitivity to different kinds of information. For example, if children are more sensitive to certain thematic relations (say, those involving animals) than to taxonomic relations involving the same entities because exposure to certain materials--for example, nursery rhymes, children's songs, and picture books--has established and primed the former relations, the sensitivity (and consequent preference) is not necessarily a theoretically interesting or generalizable observation about children's thought.

Conclusions

I have reviewed theory and research pertaining to flexibility of categorization in human development. Most of the tenable empirically-based conclusions were mentioned in the body of the paper, but I will briefly restate them for the sake of emphasis. I will also interject some broader and more speculative conclusions, and discuss important directions for future research.

A review of the philosophical and theoretical literature suggests two major points. First, traditional theories of cognition labor under the weight of an anachronistic assumption that representations are static and well-defined. According to this "stamp collector" view of cognition and concept acquisition, we acquire ideas as discrete wholes with neatly perforated edges (although children might start out with the perforations in the wrong places) and systematically file them. When judging a new stamp, we compare it to old ones and decide which provides the best match.

This assumption is not right or wrong a priori. Its status must be established empirically, and empirical evidence is the fuel for my skepticism. Prior to revisiting that evidence, however, I will stress the second point: This is not the only plausible view of knowledge. Theorists such as Edelman (1992), Lakoff (1987), Olson (1970) and Putnam (1988) propose serious challenges to the traditional view. However, none of these theorists emphasize the potentially critical point that if categorization occurs in the process of planning and achieving goals, it is necessary to take a functional approach to categorization. Of course, it is difficult to determine how much categorizing takes place in the process of goal-directed activity, and how much takes place by force of habit or automatization (imagine an automatic coin sorter). Yet this question is crucial, because the former process implies flexibility whereas the latter implies static rigidity. That is, insofar as we engage in a great deal of planning and goal-directed activity, different aspects of entities may be important at different times. If I want to hang a picture but have no hammer, I will look for something hard, flat, and heavy to drive the nail--perhaps the heel of a shoe or a paperweight. On the other hand, if static rigidity is the rule, then without a hammer my goal will be frustrated.

Perhaps Duncker's (1945) findings of "functional fixedness" address this question. Duncker's subjects had difficulty using objects for unintended purposes. Does this imply that concepts are static and rigid rather than flexible? Unfortunately Duncker's methods allow no conclusions, since subjects had to solve a difficult problem as well as use an object for a non-traditional purpose. For example, in one problem subjects had to tie two strings hanging from the ceiling. The strings were so far apart that subjects could not grasp both at one time, and the correct solution was to tie a pair of pliers to one string and swing it like a pendulum. This is a difficult problem, and perhaps no object would have improved problem solving. More importantly, once subjects understood the solution to a problem, would they demonstrate further functional fixedness? After seeing the solution to the pendumum problem subjects could no doubt solve the problem effortlessly using a rock, a shoe, or whatever.

Logically, a functionalist approach to categorization is at least as tenable as the traditional view, given the unfounded assumptions (e.g., objectivist epistemology) grounding the latter. If anything, functionalism requires fewer assumptions: its only premises are that the goal or purpose for a categorization decision strongly and significantly affects the outcome of that decision, and that people are able to entertain a number of possible categorization schemes, as defined by a plausible range of tasks with definable goals. Functionalism is not incompatible with evidence that the same set of items will always be subdivided into the same subsets: as long as the goal of the task remains substantially the same, and/or there is only one plausible way to subdivide a set of items, the outcome of a categorization decision will not change. For example, a set consisting of two normal, edible red delicious apples and two normal, edible navel oranges will always be divided into apples and oranges. By the same token, a functionalist approach need not, and should not, contend that task variables are the only factor affecting categorization decisions. Subject variables (e.g., knowledge) and stimulus variables (e.g., the internal and external structure of a set of items) also play a role, and in fact it is the interaction between task, subject, and stimulus variables that determines categorization decisions. The functionalist approach differs from more traditional approaches to categorization by suggesting that task factors typically play a strong and significant role in categorization decisions.

This being said, what is the evidence for flexibility in categorization decisions--evidence that supports a functionalist approach to categorization, as opposed to the traditional view that representations of categories are static and well-defined? First I will review the indirect evidence. Adults organize information into a wide variety of category system structures, not just taxonomic hierarchies. Since these findings range across diverse stimulus sets, we cannot rule out the possibility that a particular stimulus set compels a particular organization. However, the few studies that have varied organization within a stimulus set (e.g., Rabinowitz & Mandler, 1983) evince flexibility. A similar conclusion holds for children: those as young as three to four years use one-, two-, and three-dimensional sorts, complex associative complexes, and taxonomic hierarchies. Unfortunately there are even fewer developmental studies that vary organization within a stimulus set, but these suggest some flexibility in preschoolers (e.g., Deák & Pick, 1994). One finding (Ceci & Howe, 1978) suggests a developmental trend towards greater spontaneous flexibility in memory organization through middle childhood. These literatures indicate the need for clarification of interactions between stimulus set and task that lead to particular system structures. They also suggest the need to explore developmental trends in flexibility of categorization. One promising clue to the latter: Sharp, Cole & Lave (1979) found that the probability of recategorization increased with education rather than age. However, becuase their geometric stimuli varied along several dimensions (e.g., shape, color) that are probably taught and reinforced in the first few years , early education may accelerate selective attention among these dimensions. In more complex stimuli (e.g., biological kinds), recategorization might actually be <u>reduced</u> with elementary education, because Western education teaches us to parse things into neat- mutually-exclusive packages--for example, consider the fallacious elementary school chestnut that "a noun is a person, place, or thing" (nouns can also be events, actions, attributes, etc.). Clearly the relationship between flexibility and formal education needs explication.

Regarding the internal structure of categories, context effects and instability of GOE indicate that adults' internal structure is flexible rather than fixed, and variability from instance to instance may be rule-governed and predictable. Very little (if any) evidence for or against the flexibility of internal structure exists for children. These two issues press for future research: clarification of the source(s) of internal structure variability, and determining whether children's internal category structures are fixed or flexible, and why.

The range of categorization criteria used by children (Part III) provides indirect evidence for flexible internal structure. Contrary to numerous hypotheses and assertions, the data demonstrate that preschoolers are capable of using a wide variety of physical, functional, and taxonomic criteria. Children may be instructed to use these criteria, but they can be used spontaneously as well. Although there is some evidence about conditions that compel the use of one criterion over another, it is difficult to reach firm conclusions due to incomparability of the data. It would be useful to conduct a large-scale study to investigate the relations among these conditions.

Finally, direct evidence of flexibility (Part IV) implies three generalizations: Flexibility is not difficult to observe through simple experimental manipulations, it is often statistically robust, and it can be observed in several different kinds of tasks. Few other conclusions can be drawn. Two kinds of studies (at least) are desperately needed in this area.

The first kind is ecologically valid, quasi-experimental studies with children and adults. Adaptations of tasks like Duncker's (1945), in which problem-solving difficulty is either mitigated or systematically varied, is a promising place to start. Direct tests of flexibility may be easier to construct for children than adults: if indeed flexibility is tied into goal-directed and pragmatic contexts, it may be difficult to establish experimental situations in which adults' goals are systematically varied. Young children, however, are accustomed to role- and pretend-play, and may readily accept experimenter-provided premises for "pretend" or character-specific goals. Although this method would only work with a limited age range, it suggests a starting point.

The second kind of study would establish the relationship between verbal/semantic measures of flexibility, and non-verbal behavioral measures of flexibility. The reader may recall that all of the adult evidence cited in Part IV came from verbal tasks, allowing only the conclusion that context and world knowledge are implicated in categorization tasks relying on semantic memory. Because most studies assess the role of verbal context and knowledge within verbal tasks, it remains to be established whether such effects can be found in non-verbal or partially-verbal tasks.

There is, then, considerable evidence that categorical knowledge is flexible and context-sensitive. What is the evidence that it is well-defined and static? First, it is simply incontrovertible that some knowledge relevant to categorization-related judgments is well-established and overlearned. Furthermore, many results implying organized knowledge (e.g., recall advantage for words organized in taxonomic hierarchies) are no doubt robust and replicable. Yet these facts have little bearing on flexibility--remember the assumption that an organism must have stabile knowledge in order to mnegociate its environment effectively. Stability, however, does not mean immobility. An organism may make different generalizations across the same set of stimuli, but if the generalizations are predictable on the basis of outstanding contingencies, then that organism is flexible yet systematic. This is the source of my contention with Barsalou's (e.g., 1989) use of the word "instability." To be sure, there is probably some stochastic variation, but far more variation is likely to be causally determined. A goal of research in flexibility should be prediction of changes in internal and external category structure, rather than conclusions that category structure is either inflexible or unstable.

Three strands of research suggest static category knowledge. First, basic-level effects are extremely robust, replicable, and valid. Basic-level categories may be relatively fixed "reference points" in semantic and non-verbal categorization. However, there are systematic qualifications to basic-level effects (see Part II). Second, Barsalou (1982, 1983, 1989) and Greenspan (1986) argue that some semantic knowledge of taxonomic categories is stable, and is always evoked upon hearing a category name. Although this point is difficult to prove (how do you prove that a connotation is always evoked?), the weight of evidence lends some credibility to this position. It would be worthwhile to attempt to generate non-verbal analogies to these findings. Third, there is some neuropsychological evidence that categorical distinctions are represented in the anatomical substrate. Unfortunately, much of this evidence consists of case studies of individuals with seemingly singular deficits (e.g., Hart, Berndt & Caramazza, 1985). More common, however, are demonstrations of a dissociation of knowledge of artificial and natural kinds following neurological insult (e.g., Warrington & Shallice, 1984). It may be that a very few global distinctions, such as natural kinds vs. artifacts, are represented neurologically, and perhaps processed in distinct ways. This pattern fits theories such as Keil's (1989) that a natural kind vs. artifact distinction is learned relatively early, as well as evidence (Mandler & Bauer, 1988) that very young children can discriminate "global" categories of animals vs. vehicles.

All of these strands of evidence suggest an important point about flexibility. From an evolutionary standpoint, the most adaptive condition is a compromise between flexibility and stable knowledge in which knowledge can be "weighted" and selected so that the most relevant knowledge for a task-at-hand is given greatest consideration. Of course, there is no guarantee that humans are optimally adapted, but a reasonable starting hypothesis is limited flexibility. All of the evidence cited suggests just such a compromise. Although it would be absurd to attempt to establish the exact balance between flexibility and stability, a case-by-case treatment of particular categorization problems--particularly contextualized, applied problem--may allow investigation of sources of variability and stability and thus establish predictive validity.

Finally, a potentially interesting area of flexibility research is individual differences. It may be that more creative individuals (Kogan, 1983) have a greater capacity for recategorization and criterion shifting. It is also possible that great expertise in a particular area facilitates flexibility. Both questions are worth investigating: intelligence testing, and tests of cognitive styles and creativity (especially creativity and metaphor), may correlate with individual differences in flexibility. The role of expertise is easy to investigate by comparing recategorization and task sensitivity in two domains, using experts in each domains. Subjects may exhibit relatively more flexibility in their own domain. Examination of individual differences might provide useful evidence for theory-building, a necessary next step in the study of flexibility in categorization. More importantly, it may shed some much-needed light onto the processes by which people recategorize and flexibly access their knowledge about the world.

Appendix

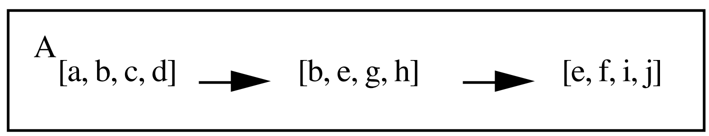
I. Simple systems

A. Closed sort: A (a) vs. B (â)

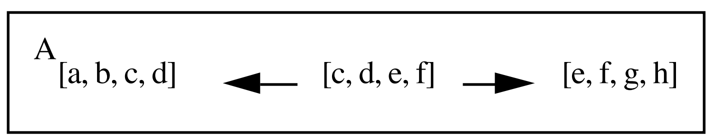
B. Open sort: A (a) vs. B (b)

II. Associative systems (arrows denote specific inter-item connections uniting category)

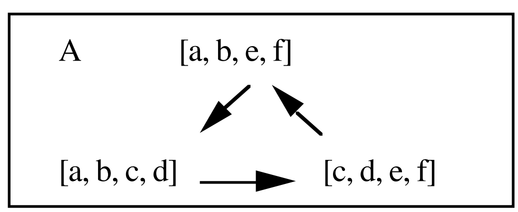
A. Chain system:



B. Pivot system:



C. Closed system:



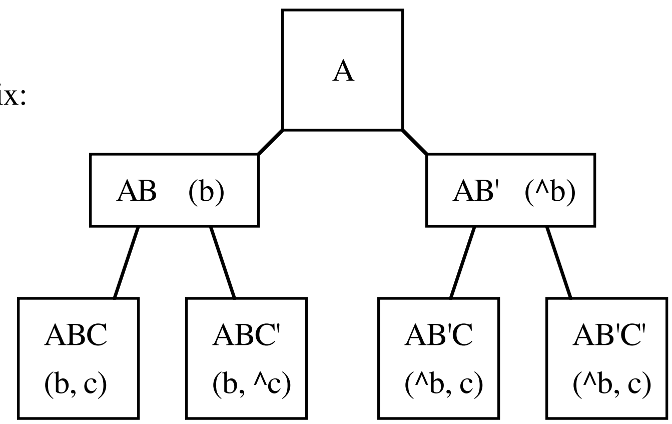
Note: Capital letters denote categories, lower-case letters denote features or sorting criteria. Letters with ^ indicate "not" relations (e.g., "â" means "does not have feature a"). Feature lists in brackets denote specific exemplars. Feature(s) in parentheses denote criteria posessed by all members of a class.

III. Matrix systems

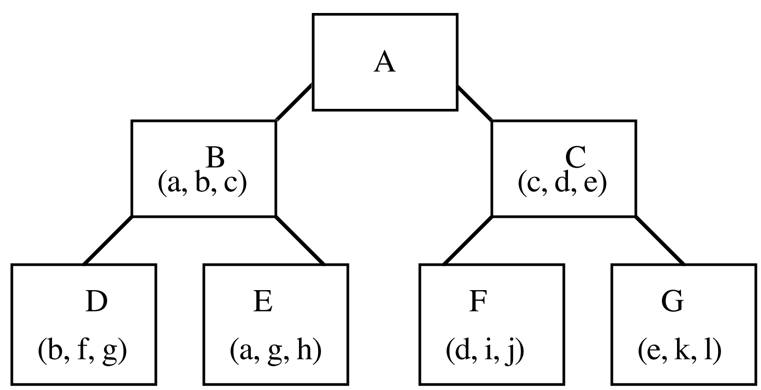
A. Symmetrical matrix:

	b	^b
a	W (a, b)	X (a, ^b)
â	Y (â, b)	Z (â, ^b)

B. Prioritized matrix:



IV. Taxonomic hierarchy



References

Ackerman, B. P. (1988). Cued recall for category, thematic, and ad hoc classified events in children and adults. Journal of Experimental Child Psychology, 45, 88-118.

Alba, J. W., Chromiak, W., Hasher, L. & Attig, M. S. (1980). Automatic encoding of category size information. Journal of Experimental Psychology: Human Learning and Memory, 6(4), 370-378.

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, D. L., 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., Gleitman, L & 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, L. W. (1982). Context-independent and context-dependent information in concepts. Memory and Cognition, 10(1), 82-93.

Barsalou, L. W. (1983). Ad hoc categories. Memory and Cognition, 11(3), 211-227.

Barsalou, L. W. (1989). Intraconcept similarity and its implications for interconcept similarity. In Vosniadou, S. & Ortony, A. (Eds.), Similarity and analogical reasoning. Cambridge: Cambridge University Press.

Barsalou, L. W. (1991). Deriving categories to achieve goals. In Bower, G. H. (Ed.), Psychology of learning and motivation: Advances in research and theory. Vol. 27. NY: Academic Press.

Bauer, P. J. & Mandler, J. M. (1989). Taxonomies and triads: Conceptual organization in one- to two-year-olds. Cognitive Psychology, 21, 156-184.

Bellezza, F. S. (1984). Reliability of retrieval from semantic memory: Noun meanings. Bulletin of the Psychonomic Society, 22(5), 377-380.

Birch, H. G. & Bortner, M. (1966). Stimulus competition and category usage in normal children. The Journal of Genetic Psychology, 109, 195-204.

Bjorklund, D. F. & Zeman, B. R. (1983). The development of organizational strategies in children's recall of familiar information: Using social organization to recall the names of classmates. International Journal of Behavioral Development, 6, 341-353.

Bower, G. H., Clark, M. C., Lesgold, A. M. & Winzenz, D. (1969). Hierarchical retrieval schemes in recall of categorized word lists. Journal of Verbal Learning and Verbal Behavior, 8, 323-343.

Bower, G. H. & Clark-Meyers, G. (1980). Memory for scripts with organized vs. randomized presentations. British Journal of Psychology, 71, 369-377.

Broadbent, D. E., Cooper, P. J. & Broadbent, M. H. P. (1978). A comparison of hierarchical and matrix retrieval schemes in recall. Journal of Experimental Psychology: Human Learning and Memory, 4, 486-497.

Brooks, L. (1978). Nonanalytic concept formation and memory for instances. In Rosch, E. & Lloyd, B. B. (Eds.), Cognition and categorization. Hillsdale, NJ: Erlbaum.

Brown, A. L. & Scott, M. S. (1972). Transfer between the oddity and relative size concepts: Reversal and extradimensional shifts. Journal of Experimental Child Psychology, 13, 350-367.

Brown, R. (1958). How shall a thing be called? Psychological Review, 65, 14-21.

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

Callanan, M. A. (1985). How parents label objects for young children: The role of input in the acquisition of category hierarchies. Child Development, 56, 508-523.

Callanan, M. A. (1989). Development of object categories and inclusion relations: Preschoolers' hypotheses about word meaning. Developmental Psychology, 25(2), 207-216.

Callanan, M. A. & Markman, E. M. (1982). Principles of organization in young children's natural language hierarchies. Child Development, 53, 1093-1101.

Caron, A. J. (1970). Discrimination shifts in three-year-olds as a function of shift procedure. Developmental Psychology, 3(2), 236-241,

Ceci, S. J. & Howe, M. J. A. (1978). Age-related differences in free recall as a function of retrieval flexibility. *Journal of Experimental Child Psychology, 26*, 432-442.

Clark, E. V. (1973). What's in a word? On the child's acquisition of semantics in his first language. In Moore, T. E. (Ed.), Cognitive development and the acquisition of language. NY: Academic Press.

Collins, A. M. & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. Psychological Review, 82(6), 407-428.

Collins, A. M. & Quillian, M. R. (1969). Retrieval time from semantic memory. Journal of Verbal Learning and Verbal Behavior, 8, 240-247.

Coltheart, V. & Evans, J. St. B. T. (1981). An investigation of semantic memory in individuals. Memory and Cognition, 9(5), 524-532.

Conrad, C. (1974). Context effects in sentence comprehension: A study of the subjective lexicon. Memory and Cognition, 2(1), 130-138.

Corrigan, R. & Schommer, M. (1984). Form versus function revisited: The role of social input and memory factors. Child Development, 55, 1721-1726.

Deák, G. & Bauer, P. J. The effects of task comprehension on preschoolers' and adults' categorization choices. Manuscript in review (a).

Deák, G. & Bauer, P. J. The dynamics of preschoolers' categorization choices. Manuscript in review (b).

Deák, G.O. & Pick, A. D. (1993). Against perceptual boundedness: Flexibility and systematicity in preschoolers' categorization choices. Poster presented at the Meetings of the Society for Research in Child Development, New Orleans, LA

Deák, G. & Pick, A. D. (1994). What is a football-telephone? Flexibility and systematicity in preschoolers' categorization decisions. Paper presented at the Annual Convention of the Americal Psychological Society, Washington, D. C.

Denney, N. W. (1974). Evidence for developmental changes in categorization criteria for children and adults. Human Development, 17, 41-53.

Denney, N. W. & Acito, M. A. (1974). Classification training in two- and three-year-old children. Journal of Experimental Child Psychology, 17, 37-48.

Dewey, J. (1920/1948). Reconstruction in philosophy. Boston: Beacon Press.

Duncker, K. (1945). On problem solving. Psychological Monographs, 58 (No. 270).

Edelman, G. M. (1992). Bright air, brilliant fire. NY: Basic Books.

Eimas, P. D. (1966). Effects of overtraining and age on intradimensional and extradimensional shifts in children. Journal of Experimental Child Psychology, 3, 348-355.

Fenson, L., Cameron, M. S., & Kennedy, M. (1988). Role of perceptual and conceptual similarity in category matching at age two years. Child Development, 59, 897-907.

Fodor, J. A. (1975). The language of thought. Cambridge, MA: Harvard University Press.

Foss, D. J. & Jenkins, C. M. (1973). Some effects of context on the comprehension of ambiguous sentences. Journal of Verbal Learning and Verbal Behavior, 12, 577-589.

Garner, W. R. (1983). Asymmetric interactions of stimulus dimensions in perceptual information processing. In Tighe, T. J. & Shepp, B. E. (Eds.), Perception, cognition, and development: Interactional analyses. Hillsdale, NJ: Erlbaum.

Gelman, S. A. & Coley, J. D. (1990). The importance of knowing a Dodo is a bird: Categories and inferences in 2-year-old children. Developmental Psychology, 26(5), 796-804.

Gelman, S. A. & Markman, E. M. (1986). Categories and induction in young children. Cognition, 23, 183-209.

Gelman, S. A. & Markman, E. M. (1987). Young children’s inductions from natural kinds: The role of categories and appearances. Child Development, 58, 1532-1541.

Gelman, S. A. & Medin, D. L. (1993). What's so essential about essentialism? A different perspective on the interaction of perception, language, and conceptual knowledge. Cognitive Development, 8(2), 157-167.

Gentner, D. (1978). A study of early word meaning using artificial objects: What looks like a jiggy but acts like a zimbo? Papers and Reports on Child Language Development, 15, 1-6.

Goodman, N. (1972). Problems and projects. Indianapolis: Bobbs-Merrill.

Greenfield, D. B. & Scott, M. S. (1986). Young children's preference for complementary pairs: Evidence against a shift to taxonomic preference. Developmental Psychology, 22(1), 19-21.

Greenspan, S. L. (1986). Semantic flexibility and referential specificity of concrete nouns. Journal of Memory and Language, 25, 539-557.

Halford, G. S. (1980). A learning-set approach to multiple classification: Evidence for a theory of cognitive levels. International Journal of Behavioral Development, 3, 409-422.

Harnad, S. (1987). Category induction and representation. In Harnad, S. (Ed.), Categorical perception: The groundwork of cognition. Cambridge: Cambridge University Press.

Hart, J. Jr., Berndt, R. S. & Caramazza, A. (1985). Category-specific naming deficit following cerebral infarction. Nature, 316, 439-440.

Heider, E. & Olivier, D. (1972). The structure of the color space in naming and memory for two languages. Cognitive Psychology, 3, 337-354.

Inhelder, B. & Piaget, J. (1964). The early growth of logic in the child. NY: Harper & Row.

Jolicoeur, P., Gluck, M. A. & Kosslyn, S. M. (1984). Pictures and names: Making the connection. Cognitive Psychology, 16, 243-275.

Jones, S. S. & Smith, L. B. (1993). The place of perception in children's concepts. Cognitive Development, 8, 113-139.

Jones, S. S., Smith, L. B. & Landau, B. (1991). Object properties and knowledge in early lexical learning. Child Development, 62, 499-516.

Kalish, C. W. & Gelman, S. A. (1992). On wooden pillows: Multiple classification and children's category-based inductions. Child Development, 63, 1536-1557.

Katz, J. J & Fodor, J. A. (1963). The structure of a semantic theory. Language, 39, 170-210.

Keil, F. C. (1979). Semantic and conceptual development: An ontological perspective. Cambridge, MA: Harvard University Press.

Keil, F. C. (1987). Conceptual development and category structure. In Neisser, U. (Ed.), Concepts and conceptual development: Ecological and intellectual factors in categorization. Cambridge: Cambridge University Press.

Keil, F. C. (1989). Concepts, kinds, and cognitive development. Cambridge: MIT Press.

Keil, F. C. & Batterman, N. (1984). A characteristic-to-defining shift in the development of word meaning. Journal of Verbal Learning and Verbal Behavior, 23, 221-236.

Kemler, D. G. (1983). Holistic and analytic modes in perceptual and cognitive development. In Tighe, T. J. & Shepp, B. E. (Eds.), Perception, cognition and development: Interactional analyses. Hillsdale, NJ: Erlbaum.

Kemler Nelson, D. G. (1984). The effects of intension on what concepts are acquired. Journal of Verbal Learning and Verbal Behavior, 23, 734-759.

Kemler-Nelson, D. G. (1991). Principle-based inferences in preschoolers' categorization of novel artifacts. Poster presented at the Meetings of the Society for Research in Child Development, Seattle, WA.

Kendler, H. H. & Kendler, T. S. (1975). From discrimination learning to cognitive development: A neobehavioristic odyssey. In Estes, W. K. (Ed.), Handbook of learning and cognitive processes, Vol. 1. Hillsdale, NJ: Erlbaum.

Khan, M. & Paivio, A. (1988). Memory for schematic and categorical information: A replication and extension of Rabinowitz and Mandler (1983). Journal of Experimental Psychology: Learning, Memory, and Cognition, 14(3), 558-561.

Kogan, N. (1983). Stylistic variation in childhood and adolescence: Creativity, metaphor, and cognitive styles. In Flavell, J. & Markman, E. M. (Eds.), Handbook of child psychology, Vol. 3. NY: Wiley.

Kossan, N. E. (1981). Developmental differences in concept acquisition strategies. Child Development, 52, 290-298.

Kripke, S. (1972). Naming and necessity. In Davidson, D. & Harmon, G. (Eds.), Semantics of natural language. Dordrecht, Holland: Reidel.

Lakoff, G. (1987). Women, fire, and dangerous things: What categories reveal about the mind. Chicago: University of Chicago Press.

Landau, B., Smith, L. B. & Jones, S. S. (1987). The importance of shape in early lexical learning. Cognitive Development, 3, 299-321.

Lane, D. M. & Pearson, D. A. (1982). The development of selective attention. Merrill-Palmer Quarterly, 28(3), 317-337.

Lange, G. (1978). Organization-related processes in children's recall. In Ornstein, P. A. (Ed.), Memory development in children. Hillsdale, NJ: Erlbaum.

Loftus, E. F. (1975). Spreading activation within semantic categories: Comments on Rosch's "Cognitive representations of semantic categories." Journal of Experimental Psychology: General, 104(3), 234-240.

Lucariello, J., Kyratzis, A. & Nelson, K. (1992). Taxonomic knowledge: What kind and when? Child Development, 63, 978-998.

Lucariello, J. & Nelson, K. (1985). Slot-filler categories as memory organizers for young children. Developmental Psychology, 21(2), 272-282.

Malt, B. C. (1988). Three kinds of concepts? Proceedings of the Tenth Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Erlbaum.

Mandler, J. M. & Bauer, P. J. (1988). The cradle of categorization: Is the basic level basic? Cognitive Development, 3, 247-264.

Mandler, J. M., Bauer, P. J. & McDonough, L. (1991). Separating the sheep from the goats: Differentiating global categories. Cognitive Psychology, 23, 263-298.

Mandler, J. M. & McDonough, L. (1993). Concept formation in infancy. Cognitive Development, 8, 291-318.

Markman, E. M. (1984). The acquisition and hierarchical organization of categories by children. In Sophian, C. (Ed.), Origins of cognitive skills. Hillsdale, NJ: Erlbaum.

Markman, E. M., Horton, M. S. & McLanahan, A. G. (1980). Classes and collections: Principles of organization in the learning of hierarchical relations. Cognition, 8, 227-241.

Markman, E. M. & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. Cognitive Psychology, 16, 1-27.

Massey, C. M. & Gelman, R. (1988). Preschoolers' ability to decide whether a photographed unfamiliar object can move itself. Developmental Psychology, 24(3), 307-317.

McClelland, J. L. & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. Journal of Experimental Psychology: General, 114(2), 159-188.

McCloskey, M. E. & Glucksberg, S. (1978). Natural categories: Well-defined or fuzzy sets. Memory and Cognition, 6(4), 462-472.

Medin, D. L. (1983). Structural principles in categorization. In Tighe, T. J. & Shepp, B. E. (Eds.), Perception, cognition and development: Interactional analyses. Hillsdale, NJ: Erlbaum.

Medin, D. L. (1989). Concepts and conceptual structure. American Psychologist, 44(12), 1469-1481.

Medin, D. L., Goldstone, R. L. & Gentner, D. (1993). Respects for similarity. Psychological Review, 100(2), 254-278.

Medin, D. & Ortony, A. (1989). Psychological essentialism. In Vosniadou, S. & Ortony, A. (Eds.), Similarity and analogical reasoning. Cambridge: Cambridge University Press.

Medin, D. L. & Schaffer, M. M. (1978). Context theory of classification learning. Psychological Review, 85, 207-238.

Medin, D. L. & Shoben, E. J. (1988). Context and structure in conceptual combination. Cognitive Psychology, 20, 158-190.

Medin, D. L. & Smith, E. E. (1981). Strategies and classification learning. Journal of Experimental Psychology: Human Learning and Memory, 7(4), 241-253.

Medin, D. L., Wattenmaker, W. D. & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. Cognitive Psychology, 19, 242-279.

Melkman, R., Tversky, B. & Baratz, D. (1981). Developmental trends in the use of perceptual and conceptual attributes in grouping clustering, and retrieval. Journal of Experimental Child Psychology, 31, 470-486.

Mervis, C. B. & Crisafi, M. A. (1982). Order of acquisition of subordinate-, basic-, and superordinate-level categories. Child Development, 53, 258-266.

Mervis, C. B. & Mervis, C. A. (1982). Leopards are kitty-cats: Object labelling by mothers for their thirteen-month-olds. Child Development, 53, 267-273.

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

Miller, R. (1973). The use of concrete and abstract concepts by children and adults. Cognition, 2, 49-58.

Mumbauer, C. C. & Odom, R. D. (1967). Variables affecting the performance of preschool children in intradimensional, reversal, and extradimensional shifts. Journal of Experimental Psychology, 75(2), 180-187.

Murphy, G. L. (1982). Cue validity and levels of categorization. Psychological Bulletin, 91(1), 174-177.

Murphy, G. L. & Smith, E. E. (1982). Basic-level superiority in picture categorization. Journal of Verbal Learning and Verbal Behavior, 21, 1-20.

Murphy, G. L. & Wisniewski, E. J. (1989). Categorizing objects in isolation and in scenes: What a superordinate is good for. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15(4), 572-586.

Nelson, K. (1973). Some evidence for the cognitive primacy of function and its functional basis. Merrill-Palmer Quarterly, 19, 21-39.

Nelson, K. (1974). Concept, word, and sentence: Interrelations in acquisition and development. Psychological Review, 81(4), 267-285.

Nelson, K. (1982). The syntagmatics and paradigmatics of conceptual development. In Kuczaj, S. (Ed.), Language development, Vol. 2. Hillsdale, NJ: Erlbaum.

Nelson, K. (1986). Event knowledge: Structure and function in development. Hillsdale, NJ: Erlbaum.

Nowak, R. M. & Paradiso, J. L. (1983). Walker's mammals of the world, 4th Ed. Baltimore: Johns Hopkins University Press.

Oldroyd, D. (1986). The arch of knowledge. Australia: New South Wales University Press.

Olson, D. (1970). Language and thought: Aspects of a cognitive theory of semantics. Psychological Review, 77(4), 257-273.

Osherson, D. N. & Smith, E. E. (1982). Gradedness and conceptual combination. Cognition, 12, 299-318.

Palmer, C. F., Jones, R. K., Hennessy, B. L., Unze, M. G. & Pick, A. D. (1989). How is a trumpet known? The "basic object level" concept and perception of musical instruments. American Journal of Psychology, 102(1), 17-37.

Palmer, D. C. & Donahue, J. W. (1992). Essentialism and selectionism in cognitive science and behavior analysis. American Psychologist, 47(11), 1344-1358.

Poincaré, H. (1905). Science and hypothesis. New York: The Science Press.

Prawat, R. S. & Wildfong, S. (1980). The influence of functional context on children's labeling responses. Child Development, 51, 1057-1060.

Putnam, H. (1975). The meaning of meaning. In Putnam, H. (Ed.), Mind, language and reality. Vol. 2. London: Cambridge University Press.

Putnam, H. (1988). Representation and reality. Cambridge, MA: MIT Press.

Quine, W. V. O. (1977). Natural kinds. In Schwartz, S.P. (Ed.), Naming, necessity, and natural kinds. Ithaca, NY: Cornell University Press.

Rabinowitz, M. & Mandler, J. M. (1983). Organization and information retrieval. Journal of Experimental Psychology: Learning, Memory, and Cognition, 9(3), 430-439.

Ratcliff, R. & McKoon, G. (1981). Does activation really spread? Psychological Review, 88(5), 454-462.

Rips, L. J. (1989). Similarity, typicality, and categorization. In Vosniadou, S. & Ortony, A. (Eds.), Similarity and analogical reasoning. Cambridge: Cambridge University Press.

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

Rosch, E. (1975a). Cognitive representations of semantic categories. Journal of Experimental Psychology: General, 104(3), 192-233.

Rosch, E. (1975b). Reply to Loftus. Journal of Experimental Psychology: General, 104(3), 241-243.

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

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.

Roth, E. M. & Shoben, E. J. (1983). The effect of context on the structure of categories. Cognitive Psychology, 15, 346-378.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M. & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. Cognitive Psychology, 14, 489-537.

Sharp, D., Cole, M. & Lave, C. (1979). Education and cognitive development: The evidence from experimental research. Monographs of the Society for Research in Child Development, 44(1-2, Serial No. 178).

Shepp, B. E., Burns, B. & McDonough, D. (1980). The relation of stimulus structure to perceptual and cognitive development: Further tests of a separability hypothesis. In Wilkening, F., Becker, J. & Trabasso, T. (Eds.), Information integration by children. Hillsdale, NJ: Erlbaum.

Smiley, S. S. & Brown, A. L. (1979). Conceptual preferences for thematic or taxonomic relations: A nonmonotonic age trend from preschool to old age. Journal of Experimental Child Psychology, 28, 249-257.

Smith, C. L. (1979). Children's understanding of natural language hierarchies. Journal of Experimental Child Psychology, 27, 437-458.

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

Smith, J. D. & Kemler Nelson, D. G. (1984). Overall similarity in adults' classification: The child in all of us. Journal of Experimental Psychology: General, 113(1), 137-159.

Smith, L. B. (1979). Perceptual development and category generalization. Child Development, 50, 705-715.

Smith, L. B. (1984). Young children's understanding of attributes and dimensions: A comparison of conceptual and linguistic measures. Child Development, 55, 363-380.

Smith, L. B. (1989). A model of perceptual classification in children and adults. Psychological Review, 96, 125-144.

Smith, L. B. & Heise, D. (1992). Perceptual similarity and conceptual structure. In Burns, B. (Ed.), Advances in Psychology-percepts, concepts, and categories: The representation and processing of information. Amsterdam: Elsevier.

Snodgrass, J. G. (1984). Concepts and their surface representations. Journal of Verbal Learning and Verbal Behavior, 23, 3-22.

Sober, E. (1985). The nature of selection: Evolutionary theory in philosophical focus. Cambridge, MA: MIT Press.

Steinberg, E. R. & Anderson, R. C. (1975). Hierarchical semantic organization in 6-year-olds. Journal of Experimental Child Psychology, 19, 544-553.

Tanenhaus, M. K., Leiman, J. M. & Seidenberg, M. S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. Journal of Verbal Learning and Verbal Behavior, 18, 427-440.

Tighe, T. (1973). Subproblem analysis of discrimination learning. In Bower, G. H. (Ed.), The psychology of learning and motivation, Vol. 7. NY: Academic Press.

Tighe, T. J., Glick, J. & Cole, M. (1971). Subproblem analysis of discrimination shift learning. Psychonomic Science, 24, 159-160.

Tighe, T. J. & Tighe, L. S. (1967). Discrimination shift performance of children as a function of age and shift procedure. Journal of Experimental Psychology, 74, 466-470.

Tomikawa, S. A. & Dodd, D. H. (1980). Early word meanings: Perceptually or functionally based? Child Development, 51, 1103-1109.

Trabasso, T., Isen, A. M., Dolecki, P., McLanahan, A. G., Riley, C. A. & Tucker, T. (1978). How do children solve class-inclusion problems? In Siegler, R. S. (Ed.), Children's thinking: What develops? Hillsdale, NJ: Erlbaum.

Tulving, E. (1983). Elements of episodic memory. NY: Oxford University Press.

Tulving, E. & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. Journal of Verbal Learning and Verbal Behavior, 5, 381-391.

Tversky, A. (1977). Features of similarity. Psychological Review, 84, 327-352.

Tversky, A. & Gati, I. (1978). Studies of similarity. In Rosch, E. & Lloyd, B. B. (Eds.), Cognition and categorization. Hillsdale, NJ: Erlbaum.

Tversky, B. (1985). Development of taxonomic organization of named and pictured categories. Developmental Psychology, 21(6), 1111-1119.

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

Warren, R. E. (1972). Stimulus encoding and memory. Journal of Experimental Psychology, 94(1), 90-100.

Warrington, E. K. & Shallice, T. (1984). Category specific semantic impairments. Brain, 107, 829-854.

Werner, H. (1957). The concept of development from a comparative and organismic point of view. In Harris, D. B. (Ed.), The concept of development.. Minneapolis: University of Minnesota Press.

White, T. G. (1982). Naming practices, typicality, and underextension in child language. Journal of Experimental Child Psychology, 33, 324-346.

Winer, G. A. (1980). Class-inclusion reasoning in children: A review of the empirical literature. Child Development, 51, 309-328.

Wittgenstein, L. (1953). Philosophical investigations. Ascombe, G. E. M. (Ed.). Oxford: Blackwell.

Zeaman, D. & House, B. J. (1974). Interpretations of developmental trends in discriminative transfer effects. In Pick, A. D. (Ed.), Minnesota symposia on child psychology, Vol. 8. Minneapolis: University of Minnesota Press.