Journal of Experimental Psychology: General
2001, Vol. 130, No. 1, 3–28

0096-3445/01/$5.00 DOI: 10.1037//0096-3445.130.1.3

Thematic Relations in Adults' Concepts

Emilie L. Lin and Gregory L. Murphy
University of Illinois at Urbana-Champaign

Concepts can be organized by their members' similarities, forming a kind (e.g., animal), or by their external relations within scenes or events (e.g., cake and candles). This latter type of relation, known as the *thematic relation,* is frequently found to be the basis of children's but not adults' classification. However, 10 experiments found that when thematic relations are meaningful and salient, they have significant influence on adults' category construction (sorting), inductive reasoning, and verification of category membership. The authors conclude that concepts function closely with knowledge of scenes and events and that this knowledge has a role in adults' conceptual representations.

The conventional view of a category is a set of objects or entities that share an essential core or are similar in some perceptual, biological, or functional properties (Medin & Ortony, 1989; Medin & Smith, 1984; Smith & Medin, 1981). It is designated by a name (e.g., *dog, animal*) that refers to entities that are in many respects equivalent (E. M. Markman, 1989; Rosch, 1978). Most research on adults' concepts has focused on categories that can be represented in *taxonomies*—a hierarchical system in which concepts are differentiated into levels of varying specificity (e.g., animal, dog, collie) related by class inclusion (e.g., a collie is a dog, a dog is an animal, a collie is an animal; see E. M. Markman & Callanan, 1983; Murphy & Lassaline, 1997, for reviews). Such taxonomic categories include natural kinds (e.g., dog, flower), artifacts (e.g., chair, car), and artificially constructed stimuli (e.g., groups of geometric shapes, dot patterns). In each case, the category includes items of the same kind.

Despite the psychological reality of concepts as kinds interrelated by class inclusion, a taxonomic perspective does not capture the entire spectrum and richness of human concepts and categorization. Many human concepts also include knowledge about the thematic relations among taxonomically unrelated concepts. As this article will show, these relations can be just as powerful as, or in some cases more powerful than, class inclusion relations in influencing adults' categorization-related behaviors. In the following discussion, we will define *thematic relations* and will present studies that examine the psychological reality of thematic categories.

Emilie L. Lin and Gregory L. Murphy, Department of Psychology and Beckman Institute, University of Illinois at Urbana-Champaign.

This research was supported by National Institute of Mental Health Grant MH41704. This research was part of Emilie Lin's doctoral dissertation. She would like to thank the members of her committee for their guidance and many valuable comments: Stephanie Doane, Gregory Murphy, Karl Rosengren, Brian Ross, and Edward Shoben. We thank Eric Peng for help in running experiments.

Correspondence concerning this article should be addressed to Gregory L. Murphy, Beckman Institute, University of Illinois, 405 North Mathews Avenue, Urbana, Illinois 61801. Electronic mail may be sent to gmurphy@s.psych.uiuc.edu.

Thematic Relations

Thematic relations are the external or complementary relations among objects, events, people, and other entities that co-occur or interact together in space and time (D. R. Denney, 1975; D. R. Denney & Moulton, 1976; Lucariello, Kyratzis, & Nelson, 1992; Lucariello & Rifkin, 1986; E. M. Markman, 1981, 1989). Some examples of thematic relations are spatial (e.g., a roof is on top of a house), functional (e.g., a piece of chalk is used to write on a blackboard), causal (e.g., electricity makes a light bulb glow), and temporal (e.g., bills typically come after meals in restaurants). Note that certain concepts can share more than one type of thematic relation; the functionally related chalk and blackboard are also spatially proximate. A thematic category, then, can be formed when thematically related items are grouped together in a conceptual sorting situation. Thus, unlike members of a taxonomic category, constituents of a thematic category may bear little if any physical or internal resemblance to each other. In fact, as the just mentioned examples show, many thematically related instances are not even the same kinds of entities; a blackboard and piece of chalk are different shapes and sizes, are made out of different materials, and have different functions.

It is instructive to contrast thematic groupings with the related notion of *ad hoc categories* (Barsalou, 1982, 1983). Categories such as plunder taken by conquerors or things to carry out of a burning house are different from taxonomic categories in a number of respects; the most important difference is that ad hoc categories do not have a strong internal structure. Members of ad hoc categories are usually not very similar to one another (e.g., slaves and jewels might be good examples of plunder taken by conquerors but are not otherwise similar). However, in ad hoc categories, but not in thematic categories, members share a common property or function. That is, although slaves and jewels are not generally similar, they are things that have typically been taken by conquerors and so they share that property. In contrast, consider the thematic relation of a dog and its leash. These items do not share the same function, but instead, one item fits the function required by the other. Similarly, in temporal relations, the restaurant bill comes after the meal, rather than the bill and meal occurring at the same time or sharing a common function. Thus, the different items in a thematic relation have different and often complementary roles in the same event or scene. Bees, honey, and a hive go together thematically not because of a shared physical property or function, but because bees live in the hive where they make honey. When people group such items together, it is because their constituents "belong together" as a unified scene or event (Inhelder & Piaget, 1964; E. M. Markman, 1981), not because they function similarly for a particular goal or purpose. We will use the term *thematic categories* to refer to such groupings as a contrast to *taxonomic categories,* although it should be remembered that both terms refer to a rather broad range of categories.

Thematic relations certainly are not unheard of in theories of concepts. For example, various network models of semantic memory do incorporate these relations (e.g., Anderson, 1976; Collins & Loftus, 1975). However, within concept research, such relations have been overshadowed by all the attention given to taxonomic categories, including almost all research performed on category learning. The only area that has generated extensive work on thematic relations is children's conceptual development. Much work from this research suggests that young children use thematic relations as the basis for category construction (i.e., sorting). In contrast, older children and young adults predominantly use similarity but not thematic relations to categorize. Thus, there has been an assumption that thematic relations are not an important part of adults' concepts. However, after reviewing the developmental work in some detail, we will argue that this conclusion may be premature.

Thematic Relations in Conceptual Development

Inhelder and Piaget (1964) used many object-sorting tasks to assess children's conceptual organization. Given the instructions to "put together things that are alike" or "things that go together," children under age 5 primarily constructed what Inhelder and Piaget called *graphic collections,* which were spatial configurations of objects. For example, given an array of blocks that varied in shape and color, these young children constructed multiple collections, such as a triangle block on top of a square block, a tower of four circles decreasing in size, and another collection of circles and squares (Inhelder & Piaget, 1964, p. 31). Similarly, given an array of small toys, these young children often put objects that "belonged" together to build a scene and then told a story about it. Likewise, when children were instructed to find all the blocks with the same name, Vygotsky (1962) found that they initially did not group the same-shaped blocks together, but instead sorted them on the basis of contiguity, or relations "observed in practical experience, in which collections of complementary things often form a set or a whole" (Vygotsky, 1962, p. 63). According to Inhelder and Piaget, it was not until age 8 or older that children could begin to sort by similarity alone.

Many other studies, using a variety of stimuli and methodologies, have also reported a complementary–similarity shift in conceptual development between ages 4 to 8, finding that beyond age 8 and throughout much adulthood, similarity or taxonomic category membership remains the primary basis for sorting (E. M. Markman, 1981, 1989). Appendix A presents some of these studies, along with their stimuli and a short description of the task used. The stimuli in these studies were pictures or physical objects. The critical tasks were of two types. One was category construction, in which subjects sorted together the stimuli that were the same (Olver & Hornsby, 1967, Experiment 2) or that go together (e.g., Annett, 1959; Smiley & Brown, 1979) and then justified their response. The other type required subjects to provide reasons for why a given group of items belonged together or state what these items had in common (e.g., Goldman & Levine, 1963; Olver & Hornsby, 1967, Experiment 1).

Given these developmental classification studies and similar findings in memory recall (Ceci & Howe, 1978; N. W. Denney & Ziobrowski, 1972; see E. M. Markman, 1981, for a summary) and word-association responses (Lucariello et al., 1992; Nelson, 1977), most researchers draw the general conclusion that thematic relations play little if any role in adults' categorization. Especially if taxonomic and thematic relations are pitted against one another, current theories would predict that adults would prefer the taxonomic structure. However, there are reasons to reevaluate the role of thematic relations in adults' concepts, as we will discuss shortly.

The traditional literature on the thematic-to-taxonomic shift has argued that it represents a radical change in children's cognitive abilities. From the Piagetian or Vygotskyan perspective, young children are unable to form logical classes, and so they rely on thematic relations as a more primitive approximation to real categories. More contemporary approaches do not take such a global view, recognizing that children may for many years make a mixture of taxonomic and thematic responses. Furthermore, it is now widely recognized that even young children are able to make taxonomic classes at the basic level and that their language use is consistent with taxonomic categories (Huttenlocher & Smiley, 1987). Nonetheless, the more recent literature often suggests that taxonomic categories are a major accomplishment of cognitive development, in some cases questioning whether children truly have a full taxonomic category system, including an understanding of class inclusion and inference (E. M. Markman & Callanan, 1983). Clearly, young children do tend to form thematic groupings rather than superordinate categories, such as animals or vehicles, which suggests that they may rely on thematic relations in everyday life more than adults do. At the very least, there is a suggestion that adults strongly prefer taxonomic responding, probably because they realize that such categories are the most useful. Thus, one goal of the present research is to examine the assumption that adult cognition is generally oriented toward taxonomic categories.

If it is found that adults do form thematic categories, this would be a significant finding for two main reasons. First, as already discussed, much literature has found (or has assumed) that adults strongly prefer taxonomic categories. Indeed, virtually all the literature on adult category learning is on taxonomic categories (e.g., see Lamberts & Shanks, 1997, for reviews). Second, there are good theoretical reasons for thinking that taxonomic categories should be strongly preferred. As summarized by E. M. Markman (1989; and see Fodor, 1972), thematic categories are not useful for much learning and induction about kinds of things. That is, if one learns that a dog has a liver, one might infer that other dogs have livers as well. But thematic groupings like dog-and-leash do not allow such inferences. Although the dog has a liver, the leash probably does not, and neither does it have a tongue, a reproductive system, fur, and so on. If one relied primarily on thematic categories, one would not be able to predict properties of new objects very well. For example, suppose you were told that there was an example of dog-and-leash in your back yard. You wouldn't know whether it was likely to be digging up the garden or was simply lying there. You wouldn't know whether to feed it or hang it up in the closet, because members of this grouping are very different, with entirely different physical and functional properties. Thematic groupings are very important, as E. M. Markman emphasizes, for understanding typical events and activities in our culture, but they do not provide a basis for categorization, which involves grouping items that are of the same kind rather than items that are in complementary relations. Although we agree with much of Markman's argument, it is still possible that the shortcomings of thematic categories have been overstated and that there are unrecognized benefits of grouping items thematically. We explore this possibility next.

Why Adults Might Use Thematic Relations to Perform Categorization-Related Tasks

Thematic Relations Could Provide Conceptual Coherence

Previous research has suggested that children may use external relations to link category members when taxonomic relations are weak. For example, E. M. Markman (1981, p. 202) contrasted classes (i.e., taxonomic categories) with collections (e.g., family, forest), arguing that collections have an internal organization that makes them coherently structured (e.g., the trees in the forest are close together, members of a family are related). Even though thematic categories are unlike collections due to the dissimilarities among their constituents, the part–whole relations that make an object coherent also characterize the interrelations among constituents of thematic categories (E. M. Markman, 1981, p. 230). Thus, it is possible that thematic relations will influence categorization decisions as well.

The idea that thematic relations can provide conceptual coherence is analogous to conclusions of studies that have examined the role of background knowledge in concept formation (e.g., Kaplan & Murphy, 2000; Murphy & Allopenna, 1994; Pazzani, 1991; Wattenmaker, Dewey, Murphy, & Medin, 1986). Murphy and Medin (1985), for example, argued that a (taxonomic) concept would lose much meaning and coherence if there were no background knowledge that relates features of the concept to each other and to category membership. This similarity to thematic categories is only an analogy, however, because it is features rather than category members that are related. For example, Murphy and Allopenna examined relations among features of a concept that either could be readily explained by prior knowledge (e.g., a vehicle that goes on glaciers and has treads) or could not (e.g., a vehicle with radial tires and four doors). When prior knowledge could explain the relations among features, they found category learning to be much faster than if the features were not obviously related (see also Kaplan & Murphy, 2000). If one extends the notion of background knowledge to include knowledge about thematic relations, then it seems likely that knowledge about thematic relations can make a category coherent. Specifically, if background knowledge makes thematic relations particularly salient or relevant, then people might even use the relations to construct categories. Here, however, the constituents of the knowledge are not features of the category (like flying and having wings) but are the category members themselves (like dog and leash).

In short, thematically related concepts could be more tightly bound and strongly connected to each other than some taxonomically related concepts are. For example, chalk and a blackboard seem to be a more integrated unit as props for the classroom scene than chalk and a marker are as types of writing implements. Consequently, it is possible that even adults would prefer to construct thematic over taxonomic categories in some circumstances. After all, people probably spend little time in "cataloguing objects, in trying to generate the taxonomies to which objects belong" (E. M. Markman, 1981, p. 203). People may spend more time in organizing their experiences by identifying the temporal, functional, or spatial relations that cause entities to form unified wholes, such as looking for chalk near a blackboard or expecting a bill after a meal.

Critical Evaluations of the Developmental Classification Studies

If thematic relations can make a set of items more coherent, why have adults in previous developmental studies not formed thematic categories? One likely reason is the very weak thematic relations among the stimuli. For example, Appendix A shows that very few meaningful thematic relations were embedded in Annett's (1959) and Olver and Hornsby's (1967) stimuli. Similarly, it is almost impossible to derive sensible, thematic explanations to account for the groupings other than the two smoking clusters in Goldman and Levine (1963). This potential stimulus bias can also be found in the studies of Inhelder and Piaget (1964), Vygotsky (1962), N. W. Denney (1972), and perhaps several others, where geometric blocks were used. The interrelations among geometric blocks are mainly defined by perceptual similarities rather than by meanings or thematic contents, and thus it is not surprising that older children and adults sorted them according to shape, since there was no strong thematic organization. Also, the "taxonomic" sorting done in studies with geometric stimuli was sometimes the separation of items based on a single defining feature, such as shape, rather than the formation of a family resemblance category (Rosch, 1978). Thus, the taxonomic sorting found in such experiments may not reflect real taxonomic categories. Even among the toy objects used in Inhelder and Piaget (1964; e.g., "7 people, 8 houses, 9 animals, 4 fir trees, 7 fences, benches, fountains, motor cars, 2 babies and 2 cradles, etc.," p. 37) or the pictures used in Olver and Hornsby (1967, Experiment 2), the items that were related by similarity completely outnumbered the very few objects that were related by salient themes. When a large number of objects must be sorted, it is usually not possible to use thematic relations to sort all the items into a reasonable number of groups.

In short, thematic relations in the majority of the previous stimuli were not very salient at all. (One exception to this observation is the stimuli of Smiley & Brown, 1979, which will be discussed at some length later.) Adults may not have used weak thematic relations to construct categories because the members of the categories would have been rather loosely related. It thus remains an empirical question whether more meaningful thematic relations might encourage thematic categorization by adults.

Finally, it should be noted that some studies have failed to find the expected thematic bias for young children. Most notably, Waxman and Namy (1997) did not find a general tendency for 2- to 4-year-olds to make thematic choices, nor did they find an increase in taxonomic responding over these ages. Instead, they found that there were considerable individual differences and that the exact wording of the instructions had significant effects. Waxman and Namy suggest that children do not have an overwhelming bias for one or the other kind of categorization and are somewhat flexible in the kind of relation they choose to emphasize. We will return to this possibility in the General Discussion section.

Effects of Interactive Experiences With Instances

The hypothesis that thematic relations may influence adults' classification is consistent with findings that interactions with categories affect classification to the extent that the resulting categorical structure violates the principle of biological taxonomy. For example, Boster and Johnson (1989) asked fishermen and undergraduates to sort line drawings of fishes based on their similarities. Undergraduates, who were novices in the fish domain, primarily used the morphology (the physical shapes) of the depicted fish to perform the task. In contrast, the fishermen used both morphology and functional characteristics (e.g., edibility, fighting ability, habitat—characteristics that concerned their uses of fish), even though such criteria crosscut biological categories. Thus, people's concepts are not completely determined by "objective natural discontinuities" in the environment (Boster & Johnson, 1989; p. 867). The functional roles or the utilities that different instances serve in people's activities (e.g., eating and fishing) also influence the way people categorize the instances (see also Malt, 1995).

The effects of category use (i.e., the functional roles that categories serve in people's activities) have been seen in a number of recent studies of natural concepts, including Medin, Lynch, Coley, and Atran's (1997) study of tree categories and Ross and Murphy's (1999) exploration of food categories. Medin et al. found that taxonomists and maintenance workers mainly used morphological features to sort the trees, but landscape workers sorted them according to their roles and values in landscape design (e.g., whether the trees were ornamentals or street trees). Furthermore, maintenance workers constructed a utilitarian category, weed trees, for fast-growing, weak-wooded trees that were a nuisance in their job. Ross and Murphy found that people categorize foods both by taxonomic grouping (e.g., meats) and by their place in eating scripts (e.g., main dishes, breakfast foods). They also showed that both kinds of categories could be used to draw inferences and to prime category members. Ross (1996, 1997) has shown the importance of such functions experimentally, for artificial categories.

The studies just described did not actually provide evidence for thematic sorting per se, because they did not provide thematically related items for subjects to choose (e.g., Boster & Johnson, 1989, gave their subjects only fish; Ross & Murphy, 1999, only foods). However, by showing considerable influence of functional relations, they raise the possibility that subjects would have used thematic relations if given the opportunity. The present experiments will make it possible for subjects to use either taxonomic or thematic relations to categorize items.

Effects of Age and Education

Another finding relevant to the current study is that elderly and uneducated people are more drawn to thematic relations than the younger and the educated. Smiley and Brown (1979) found that the majority of their healthy, educated elderly subjects preferred thematic over taxonomic matches (using a procedure like that of the present experiments, described below). Annett (1959) also found that educated adults over age 40 grouped pictures and justified their sorting according to thematic relations more often than those under 40. N. W. Denney (1974, p. 49) suggested that it is the absence of educational settings and occupational pressures that induce older adults to perform thematic categorizations, because, after all, thematic relations are a more "natural" and "obvious" basis for organizing one's experiences.

Anthropological evidence suggests that Western education is also a factor in increasing taxonomic categorization. Luria (1976—but carried out in the 1930s) tested adults in a rather primitive part of Uzbekistan and found that they often insisted on grouping items thematically. For example, they would insist that an ax be grouped with wood, so that you could cut it, or that a boy be grouped with adults, so that he could run errands for them. Indeed, they often characterized taxonomic choices (placing the ax with a saw) as "stupid" when asked about them (p. 54). However, there is no evidence that in real life these people did not understand taxonomic categories or that they had any temptation to confuse axes with wood, or children with adults. Their use of language was apparently normal. Similarly, Sharp, Cole, and Lave (1979) found that uneducated Mayan adults tended to make more thematic groupings than did children in the sixth grade or secondary school. That is, education appeared to matter more than age in thematic categorization (see also Scribner, 1974). Nonetheless, uneducated adults could use taxonomic relations when there were no competing thematic relations, as in a memory task. Thus, it seems likely that the preference for taxonomic categories in such tasks is to some degree a result of Western education's inculcation of analytic thinking skills. Indeed, the many examples Luria gives suggest more that the subjects did not understand the nature of the question and task rather than not understanding taxonomic categories. If these people did not understand taxonomic organization, it is unclear how they could use words like *animal, tool, ax,* and *boy* correctly, because such words, like most nouns, pick out items that are taxonomically related (E. M. Markman, 1989).

These considerations lead to somewhat opposite predictions for whether one should find taxonomic sorting among American college students. On the one hand, there is some suggestion that taxonomic sorting is not a necessary consequent of adult conceptual structure and that many people simply prefer to sort thematically, even though they categorize and name objects perfectly normally. On the other hand, research from the developmental and anthropological literatures suggests that young-adult, educated, American subjects would be among those most likely to emphasize taxonomic grouping. Thus, it would certainly be a surprise to find thematic grouping in a college-student population.

Theoretical Goals and Overview of Current Experiments

The present study had two related goals. The first, pursued in Experiments 1–8, was to discover whether thematic categorization can be found in American adult college students. As just noted, this group has the benefit of having reached their mature conceptual organization, is literate, has been educated for at least 12 years, and is not elderly. Thus, it provides the strongest test for the notion that thematic categorization may be a natural way for people to group items. If thematic grouping could be found in such a population, it would contradict the claims of those who argue that taxonomic categories are the major achievement of mature conceptual structure (i.e., Inhelder & Piaget, 1964; Vygotsky, 1962). It would also appear to contradict the arguments against thematically organized concepts put forward by researchers such as E. M. Markman (1989).

The second goal, carried out in Experiments 9 and 10, was to investigate whether thematic groupings have significant conceptual content. This question arises from the observation that taxonomic groupings depend on task and population differences, as reviewed earlier. It is possible that such differences do not directly reflect conceptual content but rather reflect simple preferences or biases about what kind of relations are more salient or important (this is the interpretation given by Skwarchuk & Clark, 1996). The traditional position has been to accept thematic sorting as reflecting a conceptual structure that is in some large part organized thematically. From that perspective, one might be able to argue that young children group thematically because they are not yet aware of taxonomic categories. However, it is stretching credulity to make the same argument for illiterate adults or American educated senior citizens. There is no reason to believe that Luria's subjects did not view different axes as being the same kind of thing or did not realize the similarities between dogs, cows, and sheep. Indeed, it would be hard to imagine how they could have survived, much less used language appropriately, if they truly lacked such taxonomic categories. Would they really treat the horse and its pen as the same kind of thing, and would they not use different words to refer to the ax and the wood it cuts? It seems manifestly obvious that such subjects were aware of taxonomic groupings but simply found them uninteresting or not salient in such a task. (American elderly readily acknowledge the validity of taxonomic groupings even when they do not create such categories; see Smiley & Brown, 1979.)

One way to make sense of such observations, then, is to argue that categorization in sorting tasks does not necessarily reveal conceptual content. Perhaps people are generally aware of both thematic or taxonomic relations, but which one they attend to is simply a function of what is salient to them and how they perceive the task. If this is the case, then it is necessary to show that thematic groupings do in fact have some content, rather than simply reflecting an individual's choice to attend to that relation in a particular task. There is already massive evidence that taxonomic categories are used in everyday thought and behavior, through studies of category learning, stereotypes, language, induction, and knowledge acquisition. Thus, there is no need to show that such groupings are informative. However, there is little such evidence showing that thematic categorization has a conceptual reality beyond the grouping of related items. Could such categories be used in reasoning or provide information about category members? Answering that question was the second goal of this research.

Experiment 1

This experiment used a forced-choice category construction task commonly used in developmental classification studies (e.g., Greenfield & Scott, 1986; Lucariello et al., 1992; E. M. Markman & Hutchinson, 1984; Smiley & Brown, 1979; Waxman & Namy, 1997) to determine whether young adults would frequently construct thematic categories in the presence of taxonomic relations. Subjects saw names of three items that referred to objects, events, people, and so on arranged in a triangle on each trial. The item on top of the triangle was the target, and the other two items were the target's taxonomic and thematic matches. The subjects' task was to decide which of the two matches "goes best with the target to form a category."

Unlike in most of the developmental studies, instructions in Experiment 1 explicitly directed subjects to form a category in the task and provided a similarity-based definition of category in the written instructions. The instructions defined a category as "a set of things or people that share some commonalities—be it genetic makeup, functions, purposes, physical and perceptual characteristics, or behavioral predispositions." Note that this definition is consistent with taxonomic categories rather than thematic categories, because it specifies categories as items that have properties in common. As discussed in the introduction, members of taxonomic categories generally share parts, perceptual properties, and functions, whereas members of thematic categories typically do not, since they play different roles in some setting or activity. For example, dogs share many characteristics, as do meals, but dog-and-leash or dinner-and-bill do not. In fact, it is exactly this lack of common characteristics that is the basis of the claim that thematic groupings are uninformative. Thus, if our subjects do make thematic groupings, it will be against the background of instructions that describe typical taxonomic categories. Most studies have not provided such a definition, and this makes the interpretation of their results somewhat difficult when the subjects are adults. For example, Skwarchuk and Clark (1996) found fairly high levels of thematic responding for subjects who were asked to choose items that "went together" or were "most related." Under these circumstances, however, it is unclear whether thematic selections indicate subjects' belief that the items form a category or have some other kind of relation, since items can go together and be related for a wide variety of reasons. One could well believe that a hand and a ring "go together" without believing that they are in the same category. Given our goal, then (which was not Skwarchuk & Clark's goal), it was most appropriate to make it very clear to the subjects on what basis they were to respond. Naturally, most developmental studies have not used such definitions, which would not be readily understood by children. Thus, the present method is stricter than that used in many past studies, and any evidence of thematic categorization will be all the more impressive.

Furthermore, to ensure that thematic relations were salient in the stimuli, we selected (a) thematic matches that we believed were integral to people's concepts of the targets or were meaningfully and coherently related to the targets and (b) taxonomic matches that were related to the targets at the superordinate level (e.g., animal). Note that many of the taxonomic pairings in previous studies were also related at the superordinate level (e.g., Annett, 1959; Greenfield & Scott, 1986; Lucariello et al., 1992; see Appendix A). Thus, the current results could be directly compared to these previous findings.

If adults almost always prefer taxonomic over thematic categories, as the literature often implies, then young adults should select thematic choices very infrequently in the category construction task. However, if thematic relations can be a sensible, coherent basis for category construction, then thematic choices might be selected just as often as or even more than taxonomic choices.

Method

Pretest. The pretest was a relation verification task whose goal was to select stimuli in which the grouping choices of each triad clearly shared a taxonomic and thematic relation with the target. Forty-eight triads of items were verified. Two verification questions from each triad were constructed, one for the presumably taxonomic pair (e.g., cat and lion) and the other for the presumably thematic pair (e.g., cat and litter box). The questions were in the form of "What makes *X* and *Y* go together to form a category?", with *X* and *Y* being the target and one of the choice items from a triad (e.g., "What makes *cat* and *lion* go together to form a category?"). All the questions were then divided into two versions of a paper questionnaire, so that each target appeared only once in a version. Half of the questions in each version involved taxonomic and half involved thematic pairs. Each version presented the questions in different randomized orders. Within each version, half of the questions mentioned the target before the choice item (*X* and *Y*) and half after (*Y* and *X*).

Thirty-two subjects from the University of Illinois participated in the pretest for either pay or to fulfill a course requirement. Half of them were randomly assigned to complete one version of the questionnaire and half to the other version. Written instructions defined a category as "a set of things or people that share some commonalities—be it genetic makeup, functions, purposes, physical and perceptual characteristics, or behavioral predispositions." Subjects were instructed to write no more than a sentence for each question. They were told that there were no absolute right or wrong answers and that answers like "they are both pencils" for mechanical pencil and drawing pencil and "libraries store books" for library and books were acceptable.

Two graduate students in cognitive psychology who were unaware of the purpose of the experiment classified each explanation as taxonomic or thematic for each question. Written instructions informed the judges to code an explanation as "taxonomic" if it mentioned class inclusion (e.g., daisy and rose are flowers) or a shared category, as "thematic" if it showed how the items shared an external relation (e.g., spatial, temporal, causal) within a scene or an event, and as "other" if the answer belonged to neither type (e.g., swan is graceful and turkey is not). The judges first coded the results independently and then discussed any ambiguous answers with each other. Overall, the judges agreed with each other on 97% of the responses. Their agreement was even higher (99%) for the 27 triads selected for the experiment. These 27 triads were selected with the following criterion: More than twice as many subjects provided the expected explanation (e.g., a taxonomic explanation for a taxonomic pair) as provided the other explanation (e.g., a thematic explanation for a taxonomic pair) for both kinds of pairs in a given triad. As described later, further items were added to this initial set, resulting in 38 total triads. The mean proportions of taxonomic and thematic explanations for all 38 items were 96% and 2% for taxonomic pairs and 12% and 85% for thematic pairs, respectively.

Subjects. Thirty-two native speakers of English from the University of Illinois volunteered to participate for pay. None of them participated in the pretest. All but 1 subject was below age 40 (the oldest was 44 years old).

Materials. Table 1 shows the stimuli of the forced-choice category construction task. The Micro-Experimental Laboratory (MEL) software was used to present the stimuli and record the data on PCs. The first 27 triads in Table 1 were those selected from the pretest, but the remaining triads were not previously verified. A posttest verification task was therefore constructed for these items, as described later.

Procedure and design. Before subjects performed the category construction task, they completed a profile sheet that requested information about their gender, year of birth, and occupational and educational background. Since subjects could be nonstudent members from the university community, the profile questionnaires served to ensure that they all had at least a high school education and were not elderly.

After subjects completed the profile sheet, they read the instructions for the category construction task. The instructions informed them that they would see short instructions along with three items arranged in a triangle

Table 1
Stimuli in Experiments 1, 2, 6, 7, and 8

Triad no.	Target	Taxonomic	Thematic
1	cat	lion	litter box
2	spider	wasp	spider web
3	French fries	baked potato	ketchup
4	panda bear	grizzly bear	bamboo
5	chalk	marker	blackboard
6	king	president	crown jewels
7	organ	accordion	church
8	Tortilla chips	potato chips	salsa
9	pepperoni	pork chops	pizza
10	bees	flies	honey
11	camel	antelope	desert
12	crib	water bed	baby
13	police car	sedan	police officer
14	pencil	pen	eraser
15	Hollywood	Chicago	movie stars
16	monastery	synagogue	monk
17	can opener	bottle opener	can
18	diamond ring	bracelet	engagement
19	Michael Jordan	Babe Ruth	basketball
20	robbery	treason	bank
21	beer	juice	party
22	airplane	car	pilot
23	swimming	golf	swimming suit
24	Hawaii	Missouri	beach
25	milk	soda	calcium
26	saxophone	harp	jazz
27	turkey	swan	Thanksgiving
28	waitress	stewardess	restaurant
29	igloo	cabin	Eskimo
30	hot dog	steak	mustard
31	cow	buffalo	farm
32	pig	dog	barn
33	toothbrush	hairbrush	teeth
34	coconut	pineapple	palm tree
35	movie theatre	opera house	popcorn
36	penguin	goose	The Antarctic
37	cactus	willow	dry climate
38	ambulance	fire truck	stretcher

Note. The first 27 items were those selected from the pretest.

presented on the computer. Their task was to think about the item on top of the triangle, then decide which of the two remaining items "goes best" with the first one to form a category. The same definition of category that was given in the instructions of the pretest was provided. However, unlike the pretest subjects, these subjects were not told that thematic relations could be a basis for forming a category. Subjects were told that there were no right or wrong answers, so they could select whatever choice seemed most sensible to them. They were not told to do the problems as quickly as possible. Subjects were given the following example to understand what each trial would look like and were told that the *X*, *Y*, and *Z* would be replaced by meaningful words that referred to objects, events, people, and so on.

Consider X.

Pick one of the choices that goes best with *X* to form a category.

X

1) Y 2) Z

Subjects typed a number that corresponded to their choice on each trial. They typed the number 0 if they did not know an item in the triad. As soon as they made a response, the message "Press the space bar to begin a trial" appeared. The computer presented the 38 triads of stimuli in an random order, using the format shown above. For each subject, half of the trials presented the taxonomic choice first and half presented it second, as was randomly determined by the computer. After subjects finished the task, they completed the relation verification task to verify the triads that had not been pretested. The procedure in which the questionnaires were distributed to the subjects was the same as in the pretest. Two new naive judges (both psychology graduate students) coded the subjects' explanations. The judges agreed on 97% of the responses. All the new triads met the criterion that was set in the pretest.

Results

Only one subject did not know the meaning of some item on one trial, and that trial was excluded from analysis. Averaged across subjects and items, thematic categorizations occurred 62% of the time. Thematic categorizations occurred at least as often as taxonomic categorizations for 34 out of 38 triads. This occurred even though the pretest found that the taxonomic relations were easier to identify correctly. We classified subjects as either predominantly taxonomic or predominantly thematic, if they selected a particular type of choice for at least 26 out of 38 (68%) items. Binomial probabilities determined this criterion, which was reliably different from random responding or equal preference, $z = 2.27, p < .025$, two-tailed. This criterion was used in later experiments that used the same triads as well. Table 2 presents the average percentage of triads for which each group selected the taxonomic or thematic choice. Surprisingly, there were almost twice as many subjects (66%) who were predominantly thematic than those (34%) who were predominantly taxonomic. In light of the developmental studies, these results are quite interesting because they indicate that, if anything, young adults tend to prefer thematic categories with the current stimuli.

One potential concern with these results was that the physical similarities of the names in some triads could have influenced the subjects' responses. For example, subjects might have selected their choices because the names of the choices (e.g., "spider web") shared a word with the names of the targets (e.g., "spider") or because the names of the choices and the targets were both proper nouns. The results of item analyses showed that only name overlapping possibly influenced the task: There were 4% more thematic categories when the targets and their thematic matches shared a word (thematic name overlap) than when they did not, and there were 8% more taxonomic categories when the targets and their taxonomic matches shared a word (taxonomic name overlap) than when they did not. However, the inclusion of these name-biasing triads did not affect the overall results, because the mean rate of thematic categorizations for the remaining triads was identical to the grand mean (i.e., 62%).

A related concern about the current results was that if name overlapping occurred on the first trial, subjects might then adopt the strategy to respond consistently throughout the task according to this first-trial response. To determine whether this occurred, the first triad presented to each subject was examined. The results showed that only 1 out of 4 subjects who had a thematic name overlap triad on the first trial was predominantly thematic, and only half of the 6 subjects who had a taxonomic name overlap triad were predominantly taxonomic. Thus, none of the evidence suggests that the name-biasing stimuli on the first trial determined subjects' dominant response in the category construction task. (For further analyses of possible influences of the first trial, see Experiment 2.)

Table 2
Experiments 1 and 2: Mean Percentages of Taxonomic and Thematic Categories Formed by Each Subject Group

Subject group	Taxonomic	Thematic
Experiment 1		
Predominantly taxonomic (n = 11)	92	8
Predominantly thematic (n = 21)	10	90
Overall M	38	62
Experiment 2		
Predominantly taxonomic (n = 14)	87	12
Predominantly thematic (n = 15)	15	83
No preference (n = 3)	51	48
Overall M	50	49

Discussion

In contrast to many developmental studies, Experiment 1 showed that (a) 62% of the categories that young adults constructed were thematic, (b) almost twice as many subjects consistently preferred thematic than preferred taxonomic categories, and (c) individual subjects clearly showed one or the other performance quite strongly. Even though some of the stimuli might have biased category preference slightly due to name sharing between the targets and their two matches, 62% of the categories constructed for the nonbiasing stimuli were still thematic. As far as we know, such a high rate of thematic categorization has not been demonstrated among young, educated adults. The evidence therefore supports the hypothesis that salient and meaningful thematic relations can be just as sensible as taxonomic relations for category construction.

One might argue that the high occurrence of thematic categorizations in Experiment 1 was a result of stronger word-association strengths between the targets and their thematic matches. It is certainly possible that associations had an influence on subjects' responses. What is unexpected is the finding that associations or thematic relations had an effect even though subjects were asked about categories and not associations. Word associations should have little if any effect on people's ability to identify common taxonomic concepts such as animals and beverages. If taxonomic concepts are the basis for adults' category structure, as has been argued throughout the literature, it is very surprising that so many subjects should ignore the taxonomic relations when asked to put together items that are in the same category, especially after reading a definition of taxonomic categories. One might argue that the associations between the thematic items were so strong that they overwhelmed perception of the taxonomic relations. However, we will show in later experiments that subjects respond taxonomically to the same stimuli under different circumstances. Ultimately, a claim about associations is not very different from the main hypothesis we are investigating, because the premise of our study is that concepts that occur together within scenes, events, and everyday activities are seen as related and that these relations can influence a simple categorization task that does not require speeded performance. Note that the claim is not that any kind of idiosyncratic association would have the same effect. The argument is that when associations are meaningful and are constrained by knowledge of common scenes and events, such as thematic relations, they can influence adults' categorization. The results suggest that thematic relations are probably more useful to adults in categorization-related situations than the literature has previously assumed. The question of whether these thematic relations are informative and useful, or only idiosyncratic, is addressed in Experiments 9 and 10.

Experiment 2

Experiment 2 examined other possibilities for the high level of thematic categorizations. First, the instruction to find the item that "goes with" the target might have induced people to form thematic categories (D. R. Denney, 1975; D. R. Denney & Moulton, 1976; Inhelder & Piaget, 1964; Skwarchuk & Clark, 1996), so Experiment 2 eliminated the "goes with" instructions. However, it should be noted that many developmental studies have used this wording and still documented the developmental progression of thematic to taxonomic classifications across different ages (Annett, 1959; D. R. Denney, 1975; D. R. Denney & Moulton, 1976; N. W. Denney, 1972; Lucariello et al., 1992; Smiley & Brown, 1979). Another modification in Experiment 2 was that subjects were explicitly told about the unspeeded nature of the task. Even though Experiment 1 never emphasized speed, some subjects might have perceived the task as being timed because it was presented on a computer. Speeded performance might favor thematic categorization, because concepts that are highly associated can be quickly activated, whereas concepts that are taxonomically related (especially at the superordinate level, like the current stimuli) might share a more abstract relation such that their similarities can be detected only through a deeper level of analysis (J. M. Mandler, 1983). In sum, Experiment 2 eliminated two factors that could potentially bias against taxonomic categorization.

Method

Thirty-two native speakers of English from the University of Illinois participated for either pay or to fulfill a course requirement. All were younger than 40 years of age with at least a high school education. No subject participated in more than one experiment reported in this article. The procedure was the same as in Experiment 1 except that MEL presented the same triads in the following format, where the X, Y, and Z refer to the items in a triad:

1) X

2) Y 3) Z

For each subject, the six different orderings of the target and its thematic and taxonomic matches occurred at least six times, with two of the orderings occurring seven times. The computer randomly determined which triad was to be presented in a particular order. Subjects received similar written instructions as in Experiment 1, with the following changes: Subjects were told to decide which two of the three items in a triad "best form a category," using the definition of category from Experiment 1. They were told to respond at their own pace, because their speed was not recorded, by typing the two numbers that corresponded to the items of their choice (e.g., "13" or "31"). An unknown item in the triad was indicated by typing two zeros.

Results

Two trials from 2 subjects were excluded from analysis because of a "00" response. Subjects formed a taxonomic category if they selected the target and its taxonomic match and a thematic category if they selected the target and its thematic match. Selections of the two matches (omitting the target; e.g., water bed and baby) occurred 11 times total (0.9%).

The average percentage of thematic categorizations was 49%. It is interesting to compare this level of responding to that in Experiment 1. Of course, such cross-experimental comparisons are at best suggestive, because subjects are not randomly assigned to conditions and the experiments are not run simultaneously (although in fact the comparisons reported in this article do generally reflect experiments run on the same subject population at about the same time). Nonetheless, we believe that these comparisons can be useful in understanding the possible determinants of taxonomic and thematic responding, especially for variables that are not the main focus of this inquiry and therefore are not the topic of a separate experiment.

The level of thematic responding in Experiment 2 was 13% lower than that in Experiment 1, a difference that was reliable by items, $t(37) = 10.28, p < .0001$, but not by subjects, $t(62) = 1.35$, $p > .15$. (This pattern is not surprising given that the comparison was within items and between subjects.) Furthermore, in comparison to Experiment 1, fewer subjects (47% total) were predominantly thematic and more subjects (44%) were predominantly taxonomic or had no preference (9%) for the two category types (see Table 2). The average percentage of the dominant response in each group was also lower than that in Experiment 1. Thus, changes in the instructions and procedure may have induced people to form more taxonomic categories. However, these results are still quite a contrast to the developmental studies, because there were just as many people who were predominantly thematic as people who were predominantly taxonomic, and thematic categorization still occurred about half of the time. Thus, there is no clear sign of a preference for taxonomic categories among young, educated adults with the current stimuli.

Similar to Experiment 1, name-biasing triads appeared to have some influence in the task, but thematic categorizations still occurred quite often in the nonbiasing triads ($M = 48\%$). Furthermore, the name-biasing stimuli presented on the first trial did not determine subjects' dominant response in the task, since only half of the 6 subjects who saw a thematic name overlap triad on the first trial were predominantly thematic and 2 of the 5 subjects who saw a taxonomic name overlap were predominantly taxonomic. In short, the name-biasing stimuli did not change the overall response pattern, just as in Experiment 1.

A slightly different hypothesis (suggested by a reviewer) about subjects' consistent preferences is that they may have been based on the first item encountered. For example, if a subject came across a strongly thematic item first, he or she may have made a thematic response and then stuck with thematic choices throughout most of the subsequent trials. Obviously, this possibility would be greatest if the first item were overwhelmingly preferred as thematic or taxonomic. An examination of the item results for Experiments 1 and 2 showed that there were no overwhelmingly strong items that would be likely to determine subjects' responses.

The strongest taxonomic preference was 75%, and the strongest thematic preference was 78%.

It may be that overwhelming preferences are not necessary to induce subjects' strategies. For example, if the first item had a more salient thematic than taxonomic choice, that could cause subjects to respond thematically and then to continue in that vein. However, the same item encountered by a subject who was already making taxonomic responses could well receive a taxonomic response. Thus, even if the initial item did not have an overwhelming tendency toward one or the other response, it could have influenced a subject's strategy if it had a bias toward one response. We examined this possibility by analyzing the overall response rates of the items seen *first* by thematic and taxonomic subjects. If the first item a subject saw greatly determined responding, then its overall preference should predict the subject's preference. That is, taxonomic subjects should have seen predominantly taxonomic items first, and thematic subjects should have received thematic items first. However, we found no such relationship. In Experiment 1, the taxonomic subjects first received items that had an overall level of 37% taxonomic responses, compared to 36% for the thematic subjects. In Experiment 2, taxonomic subjects first received items that were 48% taxonomic, compared to 49% for the thematic subjects. Obviously, the first item encountered could not be responsible for subjects' preferences, as their response rates were virtually identical for subjects who had opposite strategies.

Discussion

The elimination of the "goes with" instructions and the explicit information about the unspeeded nature of the task apparently did encourage people to construct more taxonomic categories. Nevertheless, the 49% rate of thematic categorizations is contrary to the previous findings that adults are predominantly taxonomic. Smiley and Brown (1979), for example, used a task very similar to Experiment 1 and found that 75% of their college students constructed taxonomic categories for at least 83% of their triads. In contrast, only 34% (Experiment 1) and 44% (Experiment 2) of subjects in the present studies constructed taxonomic categories to a much weaker criterion (at least 68% of the triads). Furthermore, all the present subjects received a definition of a category that emphasized similarity (taxonomic) relations. Thus, the level of observed thematic sorting is quite striking.

Before exploring the nature of thematic categories more fully, we attempted to resolve the discrepancies between the current and previous findings. One possible cause for the discrepancies is item differences. In comparison to previous stimuli, thematic relations among the current stimuli were much more salient (compare Appendix A and Table 1). Another cause may be differences in stimulus modality. The current experiments presented the stimuli in words, whereas many previous developmental studies presented them in pictures with or without words (Annett, 1959; Lucariello et al., 1992; Smiley & Brown, 1979) or used actual physical objects, usually toys, as the stimuli (e.g., N. W. Denney, 1972; Goldman & Levine, 1963; Inhelder & Piaget, 1964; Scribner, 1974; Vygotsky, 1962; Waxman & Namy, 1997; see D. R. Denney, 1975, for a review). Assuming that the pictorial stimuli used in the previous studies depicted individual objects in isolation (i.e., without background scenes), it is possible that such pictorial stimuli or the actual physical objects would induce people to group items according to taxonomic relations. Visual depictions of isolated objects might highlight the physical similarities among taxonomically related items and the physical dissimilarities among thematically related items. For example, one can see that the cat and dog both have four legs, a nose, two eyes, and so on but that the dog and leash do not share such properties. Hence, visual stimuli could make taxonomic categories more salient than thematic ones if individual, isolated objects are presented. Indeed, some evidence does suggest that visual and verbal stimuli elicit different types of sorting and sorting justification responses (Davidon, 1952; Olver & Hornsby, 1967). However, this would not explain thematic sorting by nonliterate adults (e.g., Luria, 1976), who of course saw pictures rather than words as stimuli.

The next three experiments were therefore conducted to determine whether item differences and stimulus modality were responsible for the discrepancies between the current and previous findings. These experiments all used the sample stimuli published in Smiley and Brown (1979), because their stimuli were very similar to the current ones (in fact, some of the triads were identical), and yet their undergraduate subjects showed strong preferences for taxonomic categories, unlike our subjects. Furthermore, their study is one of the few that actually tested adults in the triad task.

Experiment 3

Smiley and Brown (1979) used a target matching task with the "goes with" instructions to compare conceptual preferences across different ages. Experiment 3 therefore used that language, with their sample stimuli. However, in Experiment 3, the stimuli were presented in words, not in combinations of pictures and words as in Smiley and Brown's study. This way, the independent contribution of item differences and wording of the task to the replication failure could be assessed.

Method

Subjects. Eighteen native speakers of English from the University of Illinois participated to fulfill a course requirement.

Stimuli and procedure. The category construction task of Experiment 1 was used with the 10 triads published by Smiley and Brown (1979). The first letters of the item names were capitalized, because that was how the stimuli were printed in their article (see Appendix A). The instructions and procedure were the same as in Experiment 1, except that the experimenter orally told the last 10 subjects that the task was not timed and that they should respond at their own pace.

Results and Discussion

Because the response patterns were very similar between subjects who were told about the unspeeded nature of the task and those who were not, the data of the two groups were combined for analysis. Surprisingly, even with Smiley and Brown's (1979) sample stimuli, thematic categorizations occurred 73% of the time overall. Subjects were classified as predominantly taxonomic or predominantly thematic if they selected a particular type of choice for at least 9 out of 10 triads, which differed reliably from random responding in a sign test, $p < .025$, two-tailed. Table 3 shows that three times more subjects (33%) were predominantly thematic than were predominantly taxonomic (11%), and that the subjects who had no significant preference tended to prefer thematic categories

Table 3
Experiments 3–5: Mean Percentages of Taxonomic and Thematic Categories Formed by Each Subject Group

Subject group	Taxonomic	Thematic
Experiment 3		
Predominantly taxonomic (n = 2)	95	5
Predominantly thematic (n = 6)	5	95
No preference (n = 10)	27	73
Overall M	27	73
Experiment 4		
Predominantly taxonomic (n = 6)	97	3
Predominantly thematic (n = 6)	7	93
No preference (n = 6)	28	72
Overall M	44	56
Experiment 5		
Predominantly taxonomic (n = 1)	90	10
Predominantly thematic (n = 6)	10	90
No preference (n = 13)	34	66
Overall M	30	70

(selecting them 73% of the time). Thus, undergraduates' preferences for thematic categories still occurred even with Smiley and Brown's (1979) sample stimuli.

Experiment 4

Since Smiley and Brown (1979) used pictures in their study, Experiment 4 added visual stimuli to the task in Experiment 3. Perhaps the visual representation of the items is critical to obtaining taxonomic sorts, allowing subjects to notice the similarity between the two taxonomic items (but cf. Waxman & Namy, 1997). It would certainly be interesting if visual presentation were necessary to obtain taxonomic selections, but it would hardly be consistent with the idea that taxonomic categories are the primary way that adults categorize and relate objects in the world. If there is a developmental progression from a thematic to a taxonomic way of thinking, it is unclear why familiar items have to be presented pictorially to access taxonomic relations.

Method

Subjects. Eighteen students from the University of Illinois who were native speakers of English volunteered to participate for pay.

Stimuli. Smiley and Brown (1979) used color pictures from children's books to present their stimuli. Thus, color pictures of the stimuli in Experiment 3 were also found from children's books and dictionaries. All of the pictures were color drawings except for {Sheep, Goat, Wool}, which were color photos. All the pictures were then color photocopied, and the pictures from the photocopies were cut and each was pasted onto a piece of black construction paper (4.5 × 6 in. [11.43 × 15.24 cm]). The names of the pictures were also pasted beneath the pictures.

Procedure. Written instructions were very similar to those in Experiment 3, except that subjects were told to respond based on the presented items in general, not the specific items depicted on the pictures, to avoid responses based on idiosyncratic features of the pictures. The instructions provided the example that if subjects saw pictures of a library, a church, and a book, they should respond based on all libraries, all churches, and all books in general, not on the particular items illustrated.

To facilitate the process of data entry and randomization of the choice items' positions (i.e., left or right beneath the target), the experimenter used the computer version of the category construction task in Experiment 3 to administer the task. That is, the experimenter used the program from Experiment 1 to determine the order of triads and the positions of the stimuli on each trial. Subjects saw only the pictorial stimuli labeled with their names and not the computer screen. The experimenter read aloud the short instructions on the screen (i.e., "Consider [target's name]. Pick one of these choices that goes best with [target's name] to form a category") on each trial and typed in the number that corresponded to the subject's response.

Results and Discussion

Once again, thematic categorizations occurred quite often (M = 56%). This is somewhat (17%) less than the rate found in Experiment 3, a difference that was reliable by items (a within factor), $t(9) = 6.05$, $p < .0002$, but not by subjects (a between factor), $t(34) = 1.38$, $p > .15$. Although this is a cross-experiment comparison, it provides some evidence that visual stimuli encourage adults to perform taxonomic categorizations. Nevertheless, many people still favored thematic categories despite the visual stimuli. As Table 3 shows, there were equal numbers of subjects in each preference group, and the no-preference group tended to favor thematic categories. In particular, whereas only 5% of Smiley and Brown's (1979) undergraduates responded thematically at least 83% of the time, 33% of the current subjects were above this response rate.

In sum, the results show considerable thematic responding even when pictures are used. There is a suggestion that pictorial stimuli encourage a bit more taxonomic responding, but it is not a very large effect. Since our goals did not include issues of stimulus format, we did not follow this up further (i.e., with a specific experiment comparing different kinds of stimuli). One reason we did not is that there have been recent demonstrations that children may respond taxonomically even when stimuli are toy objects or photographs (Waxman & Namy, 1997), which casts doubt on the idea that stimulus format is an overriding factor.

Experiment 5

Close inspection of Smiley and Brown's (1979) methodology suggests one procedural step that might explain why we are not finding similar results. Smiley and Brown asked their subjects to justify their response after they selected a choice in each triad, but Experiment 5 did not, in part because such justification is not typical in categorization and sorting tasks. There is reason to suspect that response justification may be more likely to prompt taxonomic than thematic responses. Smith and Sloman (1994) presented sparse descriptions of an object (e.g., "circular object with a 3-inch diameter") and two categories (e.g., a pizza or a quarter), and asked their subjects whether the object was more likely to be in one or the other category. The results showed that subjects' responses differed depending on whether they were thinking out loud while making their decisions. Subjects who thought out loud were less likely to choose the superficially similar item (e.g., classifying the 3-inch diameter object as a quarter). Presumably, thinking out loud encouraged the subjects to analyze stimuli beyond surface features, such that the critical and necessary properties of the stimuli were considered (e.g., a quarter by definition cannot have a 3-inch diameter). Parallel to this finding, asking people to justify their matching choice in a category construction task may encourage them to construct more taxonomic categories, because members of the taxonomic categories in the current stimuli share more inner, core features than do the thematically related items. However, there is evidence that contradicts this prediction. Greenfield and Scott (1986) failed to find preferences for taxonomic groupings among 10- to 14-year-olds when these older children justified their responses. Experiment 5 examined the effects of response justification on category construction.

Experiment 5 was very similar to Experiment 4. The main modification was that subjects were asked to justify their decision after they selected a choice to match the target.

Method

Twenty students from the University of Illinois volunteered to participate for pay or to fulfill a course requirement. The stimuli and procedure were the same as those in Experiment 4 except for the following minor changes. First, the instructions read to the subjects on each trial were "Consider (target's name). Which one of the following goes best with the (target's name)?" The addition of the article "the" before the target's name mentioned for the second time and the elimination of the phrase "to form a category" were changes to match Smiley and Brown's (1979) instructions. Second, after subjects made a response, the experimenter asked the subjects to explain why they made their choice and recorded their justification on a piece of paper.

Results and Discussion

Experiment 5's procedure matched Smiley and Brown's (1979) most closely, and yet the mean rate of thematic responses was 70%. Table 3 shows that only 1 subject (5%) was predominantly taxonomic, whereas 6 (30%) were predominantly thematic. The subjects with no significant preference again tended to prefer thematic responses (M = 66%). The somewhat higher rate of thematic responding in Experiment 5 than in Experiment 4 may be due to the deletion of the phrase "to form a category" in the question. In short, Smiley and Brown's (1979) results were not replicated. Because the thematic relations in their sample stimuli are quite meaningful, familiar, and comparable to the stimuli used in Experiments 1 and 2, the failure to replicate is actually consistent with the idea that adults readily use meaningful thematic relations to construct categories.

Why, then, did Smiley and Brown (1979) find very different results? The answer may lie in the other two thirds of their stimuli not published in their article. It is possible that the thematic relations in these stimuli were very weak, leading to a bias for taxonomic choices throughout the task. The hypothesis that the relative salience of taxonomic and thematic relations affects category construction will be tested more directly in Experiments 7 and 8. Another possibility is uncontrolled differences in subject populations. However, it should be pointed out that both of these possible explanations are inconsistent with the notion that adults overwhelmingly prefer taxonomic categorizations as part and parcel of mature conceptual development.

Experiment 6

Thematic responses occurred 62%, 49%, 73%, 56%, and 70% of the time in Experiments 1–5, respectively. One might worry that this pattern reveals a problem with the taxonomic categories used. Perhaps these categories are so weak or unfamiliar that they do not have categorical cohesion. Experiment 6 investigated whether the taxonomic concepts in the current stimuli are conventional and well-known to adults in our subject population. Specifically, Experiment 6 required subjects to select two items from each triad that could be called by the same name. People should respond taxonomically in this task, because members of a taxonomic category are designated by the same name (e.g., *dog, furniture*). Furthermore, Markman and Hutchinson (1984) found that even 2- to 3-year-olds were more likely to select a taxonomic than a thematic match when the task was to select an item that was called by the same name as the target (although cf. Waxman & Namy, 1997). Note that the taxonomic pairs used in these studies (see Table 1) would require a superordinate name like *musical instrument* or *insect* rather than a more common basic-level name like *saxophone* or *bee,* as the category members were always in different basic categories. Thus, the name that would include these items would tend not to be the most common and preferred name to categorize objects (see Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976, Experiments 10 and 11). Nonetheless, we predicted that responding would be overwhelmingly taxonomic, because there is generally no name that includes thematically related items. Such a result would rule out the possibility that the taxonomic relations used in these studies were unfamiliar or too weak to act as a basis for grouping.

Method

Eighteen students from the University of Illinois participated to fulfill a course requirement. The stimuli were triads from Experiments 1 and 2. The computer program from Experiment 2 presented the stimuli in the following format on each trial:

Which two of the three items can be called by the same name?

1) X

2) Y 3) Z

The procedure to perform this task was the same as in Experiment 2. Written instructions informed the subjects that if all three items in a triad could be called by a common name, they should select the two that could be called by the most specific name. Subjects were also informed that their speed would not be measured and that they should make whatever responses were most sensible to them at their own pace.

Results and Discussion

The same analysis was conducted on the present data as was conducted in Experiment 2. As expected, the majority of the responses were taxonomic (M = 85%). Table 4 shows that most of the subjects (83%) were predominantly taxonomic, and their average rate of taxonomic responses was quite high (M = 95%). A small number of subjects still expressed a thematic preference. Overall, Experiment 6 reinforced the relation verification pretests of Experiment 1: It showed that most people do have these taxonomic concepts and are willing to sort together the specific items tested. The relatively low frequency of taxonomic categorizations in Experiments 1–5 is therefore unlikely to be caused by any peculiarity or unconventionality in the stimuli, procedure, or equipment used. That said, it is quite striking that 3 subjects (see

Table 4
Experiment 6: Mean Percentages of Taxonomic and Thematic Pairs Selected by Each Subject Group in the Same-Name Task

Subject group	Taxonomic	Thematic
Predominantly taxonomic (n = 15)	95	3
Predominantly thematic (n = 2)	17	82
No preference (n = 1)	61	39
Overall *M*	85	14

Table 4) strongly preferred thematic pairings in spite of the instructions, which were incompatible with thematic responses. For these subjects, there is something so compelling about thematic relations that the instructions and task were overridden.

Experiment 7

Taxonomic categories are usually said to be similarity-based. Members of the same category share shapes, parts, properties, and functions and, as a result, are more similar than members of different categories. In contrast, members of thematic categories play complementary roles in a setting or activity but are generally not particularly similar. The previous experiments have shown that subjects can access these complementary relations and in fact on average prefer to do so. However, it seems unlikely that these college student subjects do not know the taxonomic categories (and see Experiment 6) or do not find them useful. Instead, it seems more likely that they find the thematic relations more salient, cohesive, or interesting. If that is correct, one way to promote taxonomic categorization with the current stimuli might be to use an orienting task that would highlight the structure underlying taxonomic relations. If preferences for taxonomic over thematic categories do depend on the relative salience of the two kinds of relations, it should be possible to alter response preferences by emphasizing one relation. If such an effect were found, it would be further evidence against the idea that taxonomic responding is a result of a general conceptual shift toward taxonomic organization, since such shifts should be immune to minor contextual or task effects.

One way to highlight taxonomic relations is to have subjects perform similarity judgments. According to recent research on similarity comparison, judging the commonality of a pair involves the alignment of structure and feature compositions between the pair (Gentner & Markman, 1994; A. B. Markman & Gentner, 1993; Medin, Goldstone, & Gentner, 1993). For example, in judging the commonalities of a cat and its taxonomic match lion, people might align the cat's legs with the lion's legs, the cat's head with the lion's head, the cat's body size with the lion's body size, and so on. This structural alignment process, in turn, would reveal the degree to which two comparison items have (a) matching features (e.g., four legs), (b) alignable differences (e.g., smaller body size for cats and larger body size for lions), and (c) nonalignable differences (e.g., lions have manes but cats don't). Alignable differences result from finding the corresponding parts of the two concepts and then noting that the values of those parts are different. These differences seem to be particularly important in similarity comparisons and category learning (Goldstone, 1994; Kaplan, 1999; Lassaline & Murphy, 1998). Nonalignable differences occur when one item has a dimension or feature that has no corresponding dimension in the other item.

Research of Gentner, A. B. Markman, and their colleagues (especially Gentner & Markman, 1994; A. B. Markman & Gentner, 1993) shows that when people make similarity judgments, they attempt to find the corresponding parts of the things being compared and use this as a scaffolding to judge similarity and to find common and different features. Consequently, similarity judgments should highlight the fact that the taxonomically related items have much more in common than do the thematically related items. When the similarity of thematic items is judged, it will become apparent that it is very difficult to align items such as a cat and a litter box or a bee and honey. Note that we are making this prediction even though our taxonomically related items are not in the same basic categories and so are not highly similar (Mervis & Crisafi, 1982). In fact, members of superordinates share relatively few features (Rosch et al., 1976). However, members of the same superordinates are still alignable, because they generally have the same dimensional structure (A. B. Markman & Wisniewski, 1997). For example, even if bees and flies differ in many respects, comparing them reveals that they are similar in terms of having common dimensions such as legs, a head, habitats, behaviors, and so on, which would not be true of bees and honey.

We explored this possibility by providing all the subjects in Experiment 7 with a paper version of Experiment 2's triad task. Prior to selecting the category members from each triad, half of the subjects (the prior similarity group) wrote down the most important commonality between the target and the taxonomically related item and between the target and the thematically related item pair from the triad. The other half of the subjects (the control group) did not perform this task. The prediction was that subjects in the prior similarity condition would construct more taxonomic categories than those in the control condition.

We did not use a simple similarity rating as the orienting task because people could construct a taxonomic category simply because they had given a higher rating to the taxonomic than the thematic pair. The current procedure did not have this demand characteristic, because subjects were told to write down the most important commonality for each pair. There should be little difficulty in finding a single commonality for any pair. The alignment view addresses comparisons of objects in general, not only similarity judgments, which are one particular kind of comparison. Since finding a commonality in two items requires a comparison, any effect of alignability should be found in this task.

Method

Subjects. Thirty-two native speakers of English from the University of Illinois community volunteered to participate for pay. All were under age 30 and had at least a high school education.

Materials. The 38 triads from Table 1 were used to construct two types of questionnaire, one for the control and one for the prior similarity condition. For the control condition, each triad was presented as three numbered items arranged in a triangle and preceded by the question "Which two of the three items best form a category?" The six possible orderings of the items in the triangle were randomly assigned to the triads, with two of the orderings occurring seven times. The item ordering for a given triad remained the same across all questionnaires. Half of the questionnaires presented the triads in one randomized order and half presented them in the reverse order.

Questionnaires for the prior similarity condition were constructed from those in the control condition with the addition of two similarity questions prior to each categorization question. Specifically, each similarity question presented the target and one of its matches in brackets (e.g., {CAT, LION}), along with the following instructions: "Think about the commonalities between the two items. Write down the most important one." The two orders of the similarity questions (i.e., taxonomic pair first or second) were randomly assigned to the triads equally often, and the order assigned to each triad remained the same across all the questionnaires. For each similarity question, half of the time the target was presented first in the brackets and half second. All the questionnaires constructed for both conditions printed the names of the test items in capital letters.

Procedure. All the subjects completed the subject profile sheet used in Experiments 1 and 2 before they filled out a questionnaire. Half of them were randomly assigned to the prior similarity condition and half to the control. All of them received written instructions that asked them to answer the questions in the order in which they were printed. The instructions also asked the subjects to circle any items that they did not know and to skip questions that had such items. For the categorization questions, all the subjects were told to select items that seemed most sensible by writing down their corresponding numbers. The definition of a category provided in Experiments 1 and 2 was provided again, except that the word "predispositions" was eliminated. The prior similarity subjects received additional instructions for the similarity questions. These instructions gave the example that for a pair such as {LIBRARY, BOOKS}, answers like "they can both be found in schools" would be acceptable for the commonality between the two items. However, answers like "libraries store books" would be unacceptable because the answer specifies a relation rather than a property that is true for each of the items separately. Prior similarity subjects completed two sets of practice questions, each consisting of two similarity questions followed by a categorization question. The practice questions used items other than the experimental ones.

Results and Discussion

One categorization question was excluded from 10 subjects (5 in each condition) because of an error in the question. Two other triads in which 2 subjects indicated their unfamiliarity with an item and three triads in which 1 subject left the similarity questions for the thematic pairs blank were excluded from the analysis. Categories consisting of the two matches occurred 20 times overall in 16 different triads.

As Table 5 shows, results from the control condition were very similar to those from Experiment 2, where subjects performed the same task on the computer. In Experiment 2, taxonomic and thematic response rates were 50% and 49%, respectively; whereas in Experiment 7, they were 46% and 52%, respectively. Hence, the finding that adults frequently preferred thematic over taxonomic categories with the current stimuli was replicated. The main question was whether the frequency of taxonomic categorizations would increase as a result of the prior similarity judgment. The answer is yes: The prior similarity group on average constructed 33% more taxonomic categories than the control group, which was reliable by subjects and items, $t(30) = 2.72$, $p < .02$; $t(37) = 18.80$, $p < .0001$. Furthermore, only 1 subject (6%) was predominantly thematic in the prior similarity condition, whereas 9 control subjects (56%) were; 13 subjects (81%) were predominantly taxonomic in the prior similarity condition, whereas 7 (44%) of the controls were. In sum, performing a similarity comparison task prior to each category construction did encourage more people to construct taxonomic categories. The hypothesis that the relative salience between taxonomic and thematic relations affects people's category construction is therefore supported. Furthermore, the more specific prediction that comparison processes lead to the recognition of common dimensions was also supported.

Table 5
Experiments 7 and 8: Mean Percentages of Taxonomic and Thematic Categorizations as a Function of Experimental Condition and Subject Group

Experimental condition and subject group	Taxonomic	Thematic
Experiment 7		
Control (categorization only)		
Predominantly taxonomic ($n = 7$)	89	9
Predominantly thematic ($n = 9$)	9	86
M	46	52
Prior similarity		
Predominantly taxonomic ($n = 13$)	90	6
Predominantly thematic ($n = 1$)	5	92
No preference ($n = 2$)	43	57
M	79	18
Experiment 8		
Control (categorization only)		
Predominantly taxonomic ($n = 7$)	89	9
Predominantly thematic ($n = 6$)	7	92
No preference ($n = 3$)	39	61
M	49	50
Prior difference		
Predominantly taxonomic ($n = 7$)	88	11
Predominantly thematic ($n = 3$)	12	87
No preference ($n = 6$)	46	52
M	58	41

Experiment 8

The goal of Experiment 8 was the same as Experiment 7, but Experiment 8 instituted a prior difference judgment task instead. According to the structural alignment view, both similarities and differences are derived from the structural correspondence between a pair of items in an alignment process (Gentner & Markman, 1994; A. B. Markman & Gentner, 1993). Specifically, people often conceptually link alignable differences (e.g., buses have more wheels than do bicycles) with commonalities (e.g., both buses and bicycles have wheels), but they do not do so for nonalignable differences (A. B. Markman & Gentner, 1993). In addition, the number of alignable differences between two items increases as the number of their commonalities increases (A. B. Markman & Gentner, 1993). This latter finding is counterintuitive, because the similarity of a pair generally increases with its commonalties and decreases with its differences (Tversky, 1977).

Because alignable differences and similarities are positively correlated and conceptually linked, and because the taxonomic pairs should have more alignable differences than the thematic pairs do, judging the differences between the targets and their two matches might therefore also highlight the many properties that the taxonomic pairs have in common. This, in turn, could encourage taxonomic categorizations in the subsequent category constructions. Thus, even though this is in some sense the opposite judgment from that made in Experiment 7, we predicted that the alignment process involved in listing differences would also lead to taxonomic responses.

Method

Subjects. Thirty-two native speakers of English from the University of Illinois community who had at least a high school education volunteered to participate for pay. All but 3 subjects were under age 30. The oldest of these 3 was age 50, and the other 2 were in their late 30s.

Materials and procedure. The materials and procedure were the same as those in Experiment 7 with the following exceptions. The word "commonalities" was replaced by "differences" for the difference judgment questions. A different randomized order and its reverse order were used to present the categorization questions. The instructions for the difference judgment questions gave the example that for a pair like {LIBRARY, BOOKS}, an answer such as "library is a building and books are reading materials" would be acceptable, but an answer such as "libraries store books" would be unacceptable because it does not specify how the items differ.

Results and Discussion

Six questions in which subjects circled an item as unfamiliar and one which a subject left blank were excluded from the analysis. Categories consisting of the two matches occurred 14 times overall in nine different triads.

Table 5 shows that the results in the control condition were again very similar to those of the same condition in Experiment 7 and to the results of Experiment 2, that is, there was a significant amount of thematic categorizing. The main question was whether prior difference judgments induced more taxonomic categorizations. The answer is again yes, but by only 9%. This increase was highly significant by items (a within factor), $t(37) = 5.16$, $p < .0001$, but was not reliable by subjects (a between factor), $t(30) < 1$. Thus, not all subjects were influenced by the prior difference judgment task. One reason that the results of the subject analyses are not as strong as those of the item analyses in Experiments 7 and 8 is that subjects often have very different strategies, as shown in the earlier experiments, thus leading to quite high variance in any between-subject comparison. But items do not differ as much, because the variance is primarily due to the strategy subjects choose rather than properties of individual items. The difference judgment manipulation was indeed weaker than the commonality judgment: The 21% reduction in taxonomic categorizations between the prior difference and prior similarity conditions (Experiments 7 and 8) was marginally significant by subjects, $t(30) = 1.96$, $p < .06$, and was highly significant by items, $t(37) = 12.63$, $p < .0001$. Again, such cross-experimental comparisons are only suggestive, but it seems likely that there is a real difference between commonality and difference judgments in this task.

It is not all that surprising that the prior difference manipulation was weaker than the prior similarity manipulation. After all, subjects were not asked to directly focus on a pair's commonalities in Experiment 8, and hence the similarities of a taxonomic pair might not have been as apparent as in the prior similarity condition of Experiment 7. In addition, subjects might have listed many nonalignable differences for the taxonomic pairs, since there is no reason why the most important difference between a pair has to be alignable. Since alignable but not nonalignable differences are linked to commonalities, a significant number of nonalignable differences could cancel some effects from the alignable differences. To examine this possibility, we classified the differences that subjects produced for the taxonomic and thematic pairs according to whether they were alignable or nonalignable (as in A. B. Markman & Gentner, 1993). The results of this analysis showed that subjects did list many nonalignable differences. For the thematic pairs, the mean percentage of alignable differences (M = 12%) was much lower than the mean percentage of nonalignable differences (M = 88%); but for the taxonomic pairs, the percentage of alignable differences (M = 56%) was slightly higher than that of nonalignable differences (M = 44%). (This pattern is consistent with that found by A. B. Markman & Wisniewski, 1997.) We looked for a relation between the type of difference listed and the type of category constructed, but the results showed no such relation. At this point, the only explanation for the weaker manipulation is that difference judgment does not directly focus on the commonalities and so is not at powerful at promoting taxonomic categorizations.

Together, the results of Experiments 7 and 8 are counterintuitive. The two seemingly opposite kinds of judgments both led to the same effect, albeit to different degrees. Although not central to the current thesis, it should be pointed out that these results are consistent with the structural alignment view of comparison processes: Both similarity and difference judgments can highlight the commonalities of a pair, which promote taxonomic categorizations.

A related experiment (published after the present one was completed) by Wisniewski and Bassok (1999) asked subjects to list commonalities and differences of pairs of items. The items could be related by being alignable, by a thematic relation, by both, or by neither. Although this task did not require category formation, the results are nonetheless of interest, because the authors expected subjects to integrate the two items into a common theme in spite of instructions to list commonalities and differences. Wisniewski and Bassok did find that subjects integrated items, especially in cases in which the items shared a thematic relation but were not alignable (such as milk and cow); they virtually never integrated items that were alignable but had no obvious thematic relation (such as milk and lemonade). These integration responses correspond to the thematic relations of the present study (e.g., "a mechanic works on cars"). However, the absolute number of integrations listed by subjects was rather low—about .48 on average, compared to about 7.0 commonalities and differences. Thus, their results are consistent with the present findings in that comparisons generally led to taxonomically relevant responses. Wisniewski and Bassok focused on the fact that subjects integrated items even when they were instructed not to do so, but Experiment 8 focuses on the finding that similarity comparisons actually reduce thematic responses.

Experiment 9

Experiments 1–8 showed that adult subjects will often select thematic responses in categorization tasks, even when there is a taxonomic response available. One question about this finding is whether it is a strategy or preference that applies in a forced-choice selection task but that does not truly reflect categorization and concepts more generally. Perhaps the attractiveness of the thematic relation overwhelms subjects in the selection task; such a result would not reflect true concepts that are used in thought, language, reasoning, and so on. The final two experiments are directed toward investigating whether this thematic tendency has consequences for important functions of categories that have been documented in the literature.

One major function of categories (arguably the most important function) is that of induction. When one knows properties of a category, one can use this knowledge to make inferences about novel objects. For example, if a friend told us "I got a new dog yesterday," we would understand this to mean that our friend had obtained a pet, which likely has four legs, barks, will be a playmate and perhaps guard, will require feeding and veterinary care, and so on. None of this has been actually said, but these obvious inferences are drawn when the category (dog) is recognized. Without being able to make such inductions from a category, there is little point in labeling an item with a category name, because no information would thereby be provided (see Lassaline & Murphy, 1998; Murphy & Ross, 1994; E. M. Markman, 1989; Smith & Medin, 1981, for discussions). Thus, Experiment 9 focused on whether thematic categories could be the basis for induction. If they cannot, then their utility must not be very great even if subjects often select them in the triad or other categorization tasks.

Past research on category-based induction has focused on taxonomic categories (Carey, 1985; Gelman, Collman, & Maccoby, 1986; Gelman & Markman, 1986, 1987; Jones & Smith, 1993; Keil, 1987; Medin et al., 1997; Osherson, Smith, Wilkie, Lopez, & Shafir, 1990; Rips, 1975; Sloman, 1993). For example, if robins are susceptible to a certain disease, subjects would generally claim that sparrows would be more susceptible to the disease than chickens would, because robins are more similar to sparrows than they are to chickens (Rips, 1975). The most popular account of category-based induction relies on the similarity of the categories involved (Osherson et al., 1990), and so it is unclear whether thematic categories could influence inductive judgments. As pointed out earlier, members of thematic categories do not generally share properties and so are not very similar.

Even though much inductive reasoning does involve inferences within the same taxonomic categories, there are other kinds of inferences that rely minimally on taxonomic category membership or on similarity comparisons. For example, Smith, Shafir, and Osherson (1993) found that people rated the argument "Poodles can bite through wire, [therefore] German shepherds can bite through wire" to be stronger than the argument "Dobermans can bite through wire, [therefore] German shepherds can bite through wire." Subjects may have thought "If a poodle can do it, clearly a German shepherd can do it, since it is stronger." Here, the particular feature of relative strength is critical, rather than overall similarity or category membership. It seems possible, then, that thematic relations might be able to justify inductive judgments, if the inductive property is related to the particular relation.

Indeed, several studies have shown that induction depends on the relation between the property being induced and the kind of relationship between the items involved even for taxonomic categories. Given that a metal bow is used for "commemorating," Kalish and Gelman (1992) found that preschoolers inferred that a cotton bow was more likely than a metal garbage can to have the same attribute. However, when the property was "a metal bow will get corroded if put in water," the pattern of induction was reversed. Heit and Rubinstein (1994) found a similar property effect on adults' inductive reasoning. Subjects in this study were asked to estimate the probability of a type of animal having some property given that another type of animal also has that property. The results showed that the probability estimate was higher if the particular similarity relation between the premise and conclusion categories matched the type of the inferred property. For example, bears and whales are similar anatomically but dissimilar behaviorally, whereas the opposite is true with worms and snakes. Correspondingly, subjects were more confident in drawing an inference from bears to whales when the feature was anatomical, and they were more confident in inferences from worms to snakes for behavioral properties. For other examples, see Ross and Murphy (1999).

These results demonstrate that category-based induction is not an invariant function of the similarity of the two categories. Although similarity of concepts is one basis for induction, more specific relations between the concepts, such as their biological or functional relatedness, may be more important for some properties. Perhaps if the information being inferred is determined by co-occurrence in scenes or events, thematic relations could guide induction. Consider bacteria, which are transmitted through contact and spatial proximity. When predicting whether an item has germs, people may be more likely to use thematic or, specifically in this case, spatial relations rather than taxonomic relations to make their inferences. For example, knowing that a certain kind of bacteria exists in airplanes, people may infer that pilots are more likely than trucks to have the same kind of bacteria, even though trucks are much more similar to airplanes (and are taxonomically related) than pilots are.

To determine whether thematic relations can indeed promote inductive reasoning even in the presence of taxonomic relations, Experiment 9 used an inductive-reasoning task similar to the one in Medin et al. (1997). Specifically, the task presented triads like those in Experiments 1–8. In each triad, the target item had some kind of bacteria, and subjects decided whether the taxonomic or thematic match would be more likely to have the same bacteria. Note that this property is very similar to the property of disease susceptibility that Medin et al. (1997) and Rips (1975) used to examine inferences across taxonomically related instances. The property *bacteria* was chosen partly because many of our stimuli were inanimate and so could not have a disease, but also because the possession of this property may rely on external contacts among items that co-occur in space and time. If thematic relations can be coherent and meaningful, then subjects might be willing to use them as the basis for inductive inferences.

Method

Subjects. Twenty-four native speakers of English from the University of Illinois participated to fulfill a course requirement.

Materials and procedure. Experiment 9 used 42 triads, 25 of them plural versions of items from the previous experiments. Not all the previous stimuli were included, because it would make no sense if some of the items (e.g., robbery, Thanksgiving) possessed bacteria. Also, care was taken to make sure that the taxonomic matches were items that do not frequently co-occur with their corresponding targets. The computer program for the category construction task from Experiment 1 presented the induction task in the following format:

Suppose scientists discovered a new kind of bacteria in X.

Which one of the following would be more likely to have the same bacteria?

1) Y 2) Z

Subjects were informed that the task was unspeeded.

Results

One trial from a subject was eliminated due to technical error. Another trial from another subject was eliminated due to unfamiliarity with some item. Subjects selected the thematic choice (M = 81%) over four times as often as they selected the taxonomic choice (M = 19%). Subjects who selected a particular type of choice for at least 28 out of 42 triads (67%), $z = 2.16$, $p < .04$ two-tailed, were categorized as being predominantly taxonomic or predominantly thematic. Table 6 shows that only 2 out of 24 subjects (8%) were predominantly taxonomic or had no preference for either choice, whereas the rest of the subjects (92%) were predominantly thematic. The average rate of thematic responses among the predominantly thematic subjects was also reasonably high (M = 85%), and the no-preference subjects again tended to prefer thematic choices (M = 62%).

One concern about this finding is that the wording "having the same bacteria" is potentially ambiguous. The wording was intended to direct subjects to select a match that would be more likely to *carry* the same bacteria as the target. However, the wording could have been interpreted to mean to select a match that would be more likely to *suffer from* the same bacteria. Thus, subjects who favored the second interpretation might have selected animate concepts as their answers, whether these concepts were taxonomically or thematically related to the targets. However, further analysis showed that this factor did not affect the results of Experiment 9. There were 18 items that had an inanimate thematic choice and an animate taxonomic choice. If animacy were the driving factor, there would have been more taxonomic choices in these items. The thematic response rate for these 18 triads was 82%, which is very close to the response rate (81%) averaged across all the triads.

Discussion

Experiment 9 clearly showed that people predominantly generalized "having the same bacteria" from the targets to their thematic matches. This reveals that the thematic relations were not idiosyncratic associations, because most people consistently used them to infer information. It may make more sense to infer some properties, especially those possessed through contacts with surrounding entities, based on the external, spatial relations rather than internal, taxonomic relations. No doubt, functional properties could also be inferred for items that have a functional relation, and temporal properties could be inferred for items that have a temporal relation.

Note that the current results do not contradict those of Medin et al. (1997) and Rips (1975), in which people used taxonomic relations to infer the very similar property "have the same disease." Those studies did not manipulate spatial proximity or other relations of the categories involved. The most salient difference between the choice items was the degree of their similarities to the target. Thus, it was natural for people to make inferences or evaluate arguments based on the degree of similarity between the premise and the conclusion categories. However, such a strategy seems less appropriate when one of the choices frequently co-occurs with the target. For example, if cats have a certain disease, we suspect that many people would infer that dogs are more likely to have the disease than lions, given that physical proximity is a medium for the spread of disease. Thus, dogs should become a very likely candidate even though they are biologically less similar to cats than lions are. Future studies could explore this issue further by using all animate stimuli and pitting their biological relatedness against their physical, spatial proximities (since this article was written, a study of this sort has been done; Proffitt, Coley, & Medin, 2000).

The current results are also relevant to some findings from Lopez, Atran, Coley, Medin, and Smith (1997). These researchers examined how Americans and Itzaj-Mayans organized various mammals into taxonomies. They also determined whether these two groups' inductive reasoning corresponded to their folkbiological taxonomies. In diversity-based inductions (e.g., "Rats and pocket mice have a disease. Tapirs and squirrels have another disease. Do you think all other mammals on this island have the disease of rats and pocket mice or the disease of tapirs and squirrels?"), only Americans' responses corresponded to their folkbiological taxonomy. Itzaj-Mayans, in contrast, incorporated ecological considerations into their reasoning. For example, rather than relying on the biological relatedness or similarities among the species, Itzaj-Mayans were more likely to consider factors that induce disease to spread, such as the physical proximity and the likelihood of co-occurrence between the two premise animals, how far the animals can travel into other parts of the island to infect other animals, and so on. Hence, the current subjects behaved more like the Itzaj-Mayans than the Americans in Lopez et al.'s study. Our results suggest that Americans might also use ecological knowledge rather than taxonomic categories and similarity in the diversity task if they had detailed knowledge of the animals (and see Proffitt et al., 2000). In Experiment 9, the thematic relations were very obvious ones, well known to all our subjects, whereas Americans may not be as familiar with the behaviors and ecology of many wild animals. The current results confirm previous findings that inductive reasoning does not necessarily correspond to the structure of the folkbiological or scientific taxonomy. One must consider the types of property inferred and the particular relations between the items used in order to predict the reasoning process.

One direction for future research on induction is to identify the types of properties and inductive tasks in which thematic relations matter more than taxonomic relations. We suspect that such properties are limited, however. Since thematic relations are externally defined, the properties that can be inferred through the relations are restricted to only a few types—primarily those that are constrained

Table 6
Experiment 9: Mean Percentage of Stimulus Items as a Function of Subject Group and Selected Choice in the Induction Task

Subject group	Taxonomic	Thematic
Predominantly taxonomic (n = 1)	98	2
Predominantly thematic (n = 22)	15	85
No preference (n = 1)	38	62
M	19	81

by contiguity and passed on by external contacts. In contrast, taxonomic relations can be used to infer a variety of properties that are not limited by contact, such as functional, perceptual, behavioral, physiological, or internal attributes. The fact that thematic relations were used more often than taxonomic relations in Experiment 9 in no way implies that thematic relations are more effective in all cases of induction. The point of the experiment was to demonstrate that thematic relations can be meaningful and useful to the extent that they can even serve as mediators for some inductive reasoning.

Experiment 10

The goal of the final experiment was to determine whether thematic relations might also affect the verification of taxonomic category membership. Many thematic relations seem so integral to one's conceptual representation (e.g., bee and honey, cow and farm) that they might influence people's judgment of taxonomic category membership. Indeed, past demonstrations of context effects on categorization seem to be precisely the manifestations of the influence of thematic relations (e.g., Biederman, Mezzanotte, & Rabinowitz, 1982; Palmer, 1975; Roth & Shoben, 1983; Murphy & Wisniewski, 1989). For example, Roth and Shoben (1983) found that people were faster to accept "cow" as a possible referent for "animal" in reading "Stacey volunteered to milk the animal whenever she visited the farm" than in reading "Fran pleaded with her father to let her ride the animal." Roth and Shoben described this result as a *context effect,* but such an effect is more specifically due to the well-established knowledge about the thematic relations among cow, milk, and farm. Similarly, Biederman et al. (1982), Murphy and Wisniewski (1989), and Palmer (1975) all found that visual object recognition is affected by the spatial and other thematic relations that an object has with the surrounding objects in a scene.

Experiment 10 extended these previous studies by using a different paradigm. On every trial, a category name (e.g., ANIMAL) was presented, followed by names of two items one above the other (all in capital letters). Subjects then determined as quickly and as accurately as possible whether any item in the pair belonged to the category just named. In the critical true trials, only one item in a pair was the true (target) item (e.g., DOG), which appeared beneath the false item that could either be thematically related or unrelated to it (e.g., LEASH or NEST). If thematic relations are indeed integral to concepts, then categorization in the true trials might be faster and more accurate for the related than for the unrelated pairs. The source of this priming could be just the presemantic, associative strength between the prime and target or both the association bias and the activation of the target's taxonomic category membership. That is, in the latter case, subjects would actively use taxonomic relations to respond positively to the target's category membership. As will be seen next, this is likely the case, because reliance on association alone would hurt performance in the false trials.

In the critical false trials, both items of a pair were nonmembers of a specified category. The top item of a pair was again either thematically related or unrelated to the bottom item. As mentioned earlier, subjects might use a strategy in which they respond according to the association bias. For example, when the category is OCCUPATION, subjects might still make a positive (incorrect) response to related pairs like LEASH–DOG because they are thematically related. However, such a strategy would devastate accuracy. Because feedback was provided after each response, subjects should be very likely to use the stimuli's conceptual, core properties to perform the task rather than relying on the unreliable cue of thematic relation. In the DOG example, the core properties of DOG should indicate to the subjects that the correct response should be negative when the category is OCCUPATION. If subjects use thematic relations only as a strategic cue to category membership, this would be seen in much lower accuracy in the false trials, where the cue is inaccurate. On the other hand, if thematic relations help to activate taxonomic knowledge, then they could aid categorization without causing problems on false trials.

In many ways, Experiment 10 is like the many semantic priming studies widely known in the literature (e.g., Meyer & Schvaneveldt, 1976). However, most of the semantic priming phenomena are observed in naming or lexical decision tasks. In these tasks, the relations between the primes and targets usually are not controlled—they could be taxonomic (e.g., CAT and DOG), thematic (e.g., FARM and COW), or some other form of association (e.g., OPEN and CLOSE). Hence, whether thematic relations alone would produce priming is an empirical question. Furthermore, the current task was categorization, and hence the current results have more direct implications for theories of concepts.

Method

Subjects. Thirty-two native speakers of English from the University of Illinois participated to fulfill a course requirement. Half were randomly assigned to Group 1 and half to Group 2.

Materials and design. A total of 96 pairs of items belonging to 14 different categories were constructed. Each item occurred only once. For each category (e.g., ANIMAL), half of the pairs were critical (e.g., LEASH–DOG), because they were used to create the two experimental conditions (thematically related and unrelated); the other half were fillers (e.g., LEOPARD–JADE). There were no pairs of items belonging to the same category, like COW–DOG. The filler pairs consisted of unrelated items. For true trials, both the critical and filler pairs consisted of one true and one false item. The true items always appeared below the false items in the critical pairs on the computer screen, but this was reversed in the filler pairs. (Because subjects responded as soon as they identified a category member, we could not place the category member first in the critical trials, because the thematic associate beneath it would generally not be read. The filler served the function of having some category members appear first, so that subjects could not ignore the first word.) Appendix B shows all the critical pairs in the related condition for the true trials. We created pairs in the unrelated condition by pairing the true item of a related pair with the false item of another related pair of the same category (e.g., NEST–DOG). We created the false trials by presenting the critical and filler pairs in the incorrect category (e.g., presenting NEST–DOG and LEOPARD–JADE in the OCCUPATION category). Hence, both items of a pair in the false trials were not category members, even though they might be thematically related. As a result of this method of stimulus construction, subjects could not use thematic relation as a cue to the correct response.

Half of the categories presented in Appendix B (ANIMAL, HUMAN DWELLING, BUILDING, COUNTRY, SPORT, FURNITURE, and INSECT) were assigned to Group 1 and half to Group 2. Pairs that did not belong to the categories assigned to the subjects were used in the false trials (e.g., stimuli from the OCCUPATION category were paired with the ANIMAL category for Group 1). Thus, stimuli were counterbalanced across true and false trials. Within each subject group, the critical pairs were equally divided into and counterbalanced across the related and unrelated conditions for both true and false trials. As a result of this counterbalancing scheme, each critical pair appeared in both related and unrelated conditions equally often across true and false trials.

Procedure. The MEL program presented all the stimuli and category names in capital letters on a PC. On each trial, subjects first saw a prompt screen that said "Press the space bar to begin a trial." After they pressed the space bar using their thumb, the prompt disappeared, and a category name appeared beneath the prompt line. They pressed the space bar again when they understood the category name. (Category exposure time was therefore controlled by the subjects. A constant exposure time for all the category names was not used because of the differences in length and familiarity of category names.) When the space bar was pressed, the category name disappeared, and a pair of items one above the other appeared beneath the line where the category name had been shown earlier. Subjects were instructed to decide, as quickly and as accurately as possible, whether any one item in the pair belonged to the category just named. They used the index finger of their dominant hand to press a key labeled *TRUE* and the index finger of their other hand for a key labeled *FALSE*. The two response keys were / and *z*. If subjects were unfamiliar with any one item on a trial, they were told to press the number *0*. As soon as subjects made a response, the items disappeared, and the computer presented either a 400-ms beep and the message "Wrong Response!" simultaneously or just the message "Correct Response!" for 1 s. After the feedback, the prompt screen for the next trial appeared. Subjects were told that their fingers should remain on the response keys at all times during the trial and that they could take a break only during the prompt screen.

Each subject group received a different order of stimulus presentation. The order was random for each group except that no category name occurred consecutively. Within each subject group, half received the stimuli in one order and half in the reverse order.

Results and Discussion

The main question was whether category verification would differ between the related and the unrelated conditions. In the true trials, if categorization of a target (true) item can be facilitated by the activation of a previous (false) item that is thematically related to it, then categorization should be faster and more accurate in the related condition. Table 7 shows that subjects were 57 ms faster in the related condition, but the priming was significant by subjects only, $F_1(1, 31) = 4.33, p < .05$; $F_2(1, 47) = 1.62, p > .2$. The 8% accuracy increase due to priming was, however, significant in both analyses, $F_1(1, 31) = 36.59, p < .0001$; $F_2(1, 47) = 17.20, p < .0001$. Thus, the results showed that thematic relations facilitated the affirmation of taxonomic category membership.

The source of thematic priming in the true trials did not come from a bias to respond positively on related items, because this would have led to low accuracy and a fast mean response time (RT) in the related condition of the false trials. Instead, Table 7 shows that subjects' mean accuracy in this condition was quite high (94%) and that mean false RTs did not differ between the two conditions (Fs < 1). The 3% difference in accuracy was marginal, $F_1(1, 31) = 3.92, p < .06$; $F_2(1, 47) = 3.65, p < .07$. In sum, subjects did not use the presence of the relation as a cue to quickly guess "true" but instead used their categorical knowledge about the items to perform the task. The results show that thematic relations are integral to this categorical knowledge—they help people to retrieve the core information of concepts.

The effects demonstrated in Experiment 10 were apparently an automatic component of the retrieval process in that thematic relations had no explicit relation to the task and yet they influenced performance under speeded conditions. In contrast, category construction and induction involves more conscious decision making and is probably subject to strategic responding (e.g., making similar categories throughout the task). Thus, the current study has demonstrated that thematic relations can influence both relatively fast and automatic as well as relatively slow and conscious categorization processes.

Table 7
Experiment 10: Mean Response Times (RTs; in Milliseconds) and Accuracy of Category Verification as a Function of Relatedness and Trial Type

	Trial			
	True		False	
Condition	RT	%	RT	%
Related	1,098	96	1,127	94
Unrelated	1,155	88	1,109	97

General Discussion

Summary of Results

Experiments 1 and 2 showed that young, educated adults performed thematic categorizations very often (62% and 49% of the time, respectively). Experiments 3–5 attempted to replicate Smiley and Brown's (1979) finding that college students seldom perform thematic categorization, using the stimuli published in their article, but without success; thematic categorizations occurred 73%, 56%, and 70% of the time, respectively. Experiment 6 showed that people did respond taxonomically with the current stimuli (85% of the time) when they were instructed to select items that could be called by the same name. Experiments 7 and 8 showed that the performance of similarity judgments—and, to a lesser degree, difference judgments—increased the amount of taxonomic responding, probably reflecting the process of alignment that emphasized the structural similarity of the taxonomic pair. The final two experiments revealed that thematic categories influence processing beyond category formation. Experiment 9 showed that thematic relations can guide category-based induction even in the presence of taxonomic relations. Experiment 10 showed that thematic relations can also affect decisions about taxonomic category membership. Together, these experiments suggest that thematic relations in adults' concepts play a more important role than past research has typically concluded.

Since our experiments were completed, Bassok and Medin (1997) and Wisniewski and Bassok (1999) have reported related findings for similarity judgments. That is, subjects claim that pairs such as cup and tea or peanut butter and knife are similar, apparently because of thematic relations that connect the two. These results occurred in spite of instructions that emphasized featural similarity (Wisniewski & Bassok, Experiment 2). Wisniewski and Bassok argued that the items themselves elicit the most appropriate form of processing. That is, pairs of items that are easily alignable will evoke the usual similarity comparison, whereas pairs of items that are involved in the same events will evoke the event or relation that connects them, and this tendency is resistant to instructions to attend to only one of these relations. Thus, it appears that use of thematic relations is not restricted to the category construction task but may be a broader cognitive strategy. In the following discussion, we will discuss the origins and likely function of the use of thematic relations, but we first address the more specific methodological and empirical issues arising from our results.

Issues Concerning Category Construction

As far as we know, no studies have demonstrated that adults not only use thematic relations to construct categories but in many circumstances prefer them to similarity relations (though see Skwarchuck & Clark, 1996, for a related finding in a task that did not emphasize categorization). This finding thus contradicts the general assumption that educated adults prefer taxonomic over thematic categories and particularly contradicts the strong claim that there is a developmental shift from thematic to taxonomic thinking. In the following discussion, we explore potential reasons for individual differences in category construction, point out response flexibility in the task, address the generality of the current results, and speculate on the mechanisms underlying the results.

Individual differences. One of the most surprising findings of the present study was the very strong individual differences in categorization. The results of Experiments 1–5 clearly showed that subjects tended to respond predominantly taxonomically or thematically rather than to mix their responses. Specifically, averaged across those five experiments and the two control conditions in Experiments 7 and 8, 45% of the subjects predominantly constructed thematic categories, whereas 32% predominantly constructed taxonomic categories. Only 23% did not have a reliable preference for one form of categorization. What, then, led to these individual differences? We can only speculate about the causes, especially given that the experiments were not designed to investigate individual differences.

One possible explanation is that the predominantly taxonomic subjects might have viewed the task as a measure of their intelligence or test-taking skills, even though the instructions specified that there was no right or wrong answer. That is, because education encourages taxonomic categorizations (Luria, 1976; E. M. Markman, 1981; Scribner, 1974; Sharp et al., 1979), these subjects might have felt the need to respond as they normally would have expected to respond in a formal classroom setting. Another possibility is that subjects had different domain knowledge, which could result in different ways of responding in such tasks. As Medin et al. (1997) have shown, tree experts with different domain knowledge (e.g., landscapers vs. maintenance workers) use different principles to construct tree categories. The current subjects might also have had different domain knowledge that led them to sort the items differently. For example, when confronting the triad {cow, buffalo, farm}, subjects with relatively extensive knowledge of biology or agriculture might have been more sensitive to the physiology or the genetics of the two animals than to the thematic relations between the cow and farm. (Recall that the majority of our subjects were college students, some of whom were likely studying these fields.) However, such subjects must have generalized this form of response across most of the tested domains because we found that responding was extremely consistent across items for subjects who showed a preference. In contrast, for students who were unfamiliar with biology or with any discipline that emphasizes hierarchical classification, their most salient knowledge of cows or other entities is likely to come from the readily accessible information about them. Much of this readily accessible information is often thematically related concepts, such as milk, saying moo, farms, or dairy products. In short, different kinds of domain knowledge, especially that developed through training in particular disciplines, may lead to different beliefs about what kinds of relations are most important.

It is also possible that more general differences in cognitive styles account for the individual differences. Dunham and Dunham (1995) found that 3-year-olds whose vocabulary consisted of more nouns than adjectives and whose pointing gestures consisted of more object reference than functional relations constructed more taxonomic than thematic categories in sorting and vice versa for children who exhibited the opposite language and gesture pattern. Even though these results are from young children, they support the notion that adults with different perceptual–cognitive processing styles should also exhibit different sorting patterns.

In sum, differences in expectations, experiences, culture (Luria, 1976; Sharp et al., 1979), background knowledge, and thinking style could all lead to individual differences in sorting strategies. One should be wary, however, of assuming that even strong differences in category construction in such a task reflect major differences in thought. Smiley and Brown (1979) showed that subjects who made one kind of choice in the triad task were extremely accurate at identifying and justifying the opposite choice. For example, college students could justify *all* of the choices that they did not select (Smiley & Brown, Table 3). Thus, a strong preference does not indicate a lack of knowledge about the other relation. As reviewed in the introduction, researchers have argued that it is very unlikely that preschool children who reliably make thematic responses do not have taxonomic categories (Fodor, 1972; E. M. Markman, 1989), and this claim is even more farfetched for adults. The importance of the present experiments is in showing that thematic categories are still salient ways of organizing objects for adults, and that these categories influence a number of category-related tasks—not in showing that some adults "think thematically" and do not use taxonomies.

Response flexibility. Sorting strategies can differ within as well as between individuals. We did not set out to examine determinants of sorting preference (except in Experiments 7 and 8), but some differences did arise as we changed the procedures in an attempt to find the expected taxonomic categorization. Overall, the results suggest that adults' preferences for thematic categories can be diminished by (a) instructions, (b) stimulus modality, and (c) prior-context manipulations (e.g., prior similarity and difference judgments). From one perspective, such flexibility in sorting is not particularly surprising, since various context effects have been well documented in the literature on similarity and categorization, starting with Tversky (1977). For example, Roth and Shoben (1983) found that people's typicality judgments of category members change depending on the situation in which the category is mentioned. Barsalou (1991) summarized evidence that categorization decisions can vary fairly drastically based on a person's current goals and conceptual orientation. In light of findings like this within taxonomic categorization, it is not surprising that preference for taxonomic or thematic categorization can also shift based on instructions or stimulus modality. However, such effects again argue against a view that adults only divide up the world in terms of taxonomic categories as part of a broad shift in conceptual organization.

Generality to other category construction tasks. Many of the category construction tasks in the adult literature require subjects to sort an array of 10 or more items rather than the present forced-choice task using triads (e.g., Ahn, 1991; Ahn & Medin, 1989; Medin, Wattenmaker, & Hampson, 1987; Regehr & Brooks, 1995; Spalding & Murphy, 1996). It is an empirical question whether the tendency to construct thematic categories would be as strong with a larger set of stimuli. The answer might depend on a combination of factors, including (a) the number and size of the potential thematic and taxonomic categories, (b) the instructions (e.g., whether the instructions explicitly specify sorting according to "alikeness" or "belongingness"), (c) stimulus modality (verbal vs. visual), and (d) the relative salience of taxonomic and thematic relations in the stimuli. No experiment seems to have been performed in which roughly equally salient thematic and taxonomic organizations of a stimulus set are contrasted. For example, one could construct a set of items that could be divided into either taxonomic categories (like animals and artifacts) or thematic categories (like dog-and-leash). Subjects would have to choose between two incompatible ways of organizing the stimuli. We suspect that very similar results would be found in such a task, in which some subjects would focus on thematic relations and others on taxonomic relations.

Unfortunately, large sets of items used in past experiments (see Appendix A) have not had equally available, multiple ways in which they could be categorized, and so such a comparison does not seem to have been made. For example, in Annett's (1959) stimuli shown in Appendix A, a subject might group the butterfly with the flower or the bird with the tree as thematic units (butterflies are attracted to flowers; birds nest in trees), but there are no other obvious thematic groupings of the remaining stimuli. As a result, adults might not even group the butterfly and flower together but instead might use the taxonomic grouping for all the items to divide them up into a small number of equal-sized categories. If such a division had been available for thematic categories, it is possible that adults would have formed considerably more thematic categories in Annett's study. A recent study found that adults often categorize thematically, even when taxonomic categorization is possible, if the thematic grouping can be made consistently (Murphy, in press).

Limitations of Thematic Categories

The present study does not imply that people perform thematic categorizations more often than taxonomic categorizations in their daily lives or that thematic categories are superior to taxonomic categories. When information that needs to be organized is vast, taxonomies can easily reduce the complexity into a manageable fashion, but thematic organizations, due to their lack of structure, are seldom sufficient. On the other hand, thematic organizations may be particularly good at representing detailed information about what happens in specific events or situations. Both kinds of knowledge are necessary. What is surprising is that the thematic knowledge had such a strong influence in our experiments in spite of instructions emphasizing taxonomic relations. Another shortcoming of thematic categories, as discussed in Experiment 9, is that they are probably limited in their inductive potential. When the constituents of a thematic category share few or no core properties and perceptual features, people cannot draw inferences concerning a vast amount of information like the internal compositions or functional and surface features of the constituents. Inferences are instead limited to the properties that are related to the particular thematic relation, such as spatial proximity.

In spite of these limitations, the results suggest some interesting possibilities for future study of induction in thematic categories. It might be interesting to discover whether some property types are more affected by thematic than by taxonomic relations in an induction task. Experiment 9 used the property of having a bacterium, but it is not known whether other properties are also extendable and on what basis thematic inductions are made. For example, people are sometimes willing to make category-based inductions on completely "blank" predicates such as "has property *X*" (e.g., "If robins have property *X*, do cardinals have property *X*?"), presumably because of the overall similarity of category members, which means that they tend to share properties. It would be interesting to see whether thematic groupings also allow purely arbitrary inductions of this sort or whether induction is completely controlled by the identity of the specific property. Another question is whether some bases for thematic groupings (e.g., functional relations) promote inductions more than other bases do (e.g., temporal relations).

Implications for Conceptual Organization

The effects of thematic relations on category construction, induction, and taxonomic category verification suggest that concepts may be intertwined with background knowledge of events and scenes. However, this does not mean that there is no structure in conceptual representation. A more likely possibility is that, in addition to taxonomies, concepts are organized via background knowledge, such as schemata (Mandler, 1979). Indeed, Lucariello and her colleagues (Lucariello et al., 1992; Lucariello & Rifkin, 1986) suggest that instances that serve the same role or function might be organized together because they fill in the same "slot" in the schema. For example, Lucariello et al. found that when 4-year-olds, 7-year-olds, and adults were instructed to produce all the instances of a superordinate category that they could think of, the productions from all age groups predominantly clustered around slot-filler categories (e.g., they produced clusters like cereal, milk, egg, and pancake—items that can fill in the breakfast slot for the food category). Ross and Murphy (1999) have shown that such relations influence people's sorting of and reasoning about food. Thus, knowledge of scenes and events does appear to constrain the organization of concepts. These slot-filler categories are not like most thematic categories, however, because they share the features of the things that fill a given slot. Cereal and eggs are both things that one might eat for breakfast, and they share a number of properties. In contrast, cereal and a bowl are thematically related by having complementary functions (cereal goes in the bowl so that one can eat it) but share few important properties.

Barsalou (1991) presented a major analysis of how concepts can form through goal-related activities. These *goal-derived* categories are often represented in memory because they derive from multiple experiences of a similar sort. People form categories such as inexpensive vacations or ways to get out of going to a party through encountering such situations repeatedly. Although inexpensive vacations may vary in a number of respects, the commonalities in the planning processes and experiences on the vacations result in a concept that is coherent and useful for planning events or in comprehending narratives of such events. That is, although one may have gone to West Virginia on one vacation and to northern Michigan on another, the commonality of both places being rural, inexpensive locations could be incorporated into the concept.

The constituents of such categories are not exactly the same as the constituents of object categories. Although there are common features that reappear (e.g., cooking for oneself on cheap vacations), some of the parts of such concepts arise through the planning process rather than through direct experience. The concepts may include negative information (e.g., air travel is not consistent with inexpensiveness). Furthermore, the concepts may include properties that are more diverse than those found in the traditional object concept. So, properties of inexpensive vacations include locations, methods of transportation, activities, things to be eaten, time parameters, and so on. This is in contrast with a typical object superordinate category, like furniture, which includes only manufactured physical objects, but it is quite similar to thematic categories, which often include very different kinds of things. Barsalou (1991) argued that the entities in a goal-derived category are bound together in a schemalike organization, in which mutual constraints are represented. For example, the method of transportation influences the possible locations, the locations influence the activities, and the price constraint sadly influences everything.

These goal-derived concepts can then lead to further concepts that are subtypes of the more general kind. For example, if one decides to go camping on a number of vacations to save money, one may form the concept of a cheap camping vacation, in which the general idea of an inexpensive vacation is specialized in terms of locations, activities, and equipment involved (Barsalou, 1991).

Barsalou's proposal suggests a very different kind of conceptual organization from the typical assumptions of the Roschian tradition. The usual view of concepts would point to structure internal to a concept like lakes, for example, that help one to identify lakes and understand what kind of things they are. But cross-cutting this concept are various goal-derived concepts that might involve lakes (such as concepts related to vacations, fishing, water sources, environmental issues, etc.). As a result, our understanding of lakes is partly determined by how they fit into such goal-derived concepts. If one has a long vacation by a lake, learning more about how to enjoy oneself there and take advantage of the lake's opportunities, one's concept of lakes has altered, although perhaps not in a way that would be obvious in a categorization task or typicality ratings.

The thematic categories discovered in the present research may be related to Barsalou's proposal. Indeed, a number of them may be goal-derived categories of the sort he describes, such as swimming and swimsuit or diamond ring and engagement. There is no inconsistency in having both a taxonomic category of dogs and a thematic category of dog-leash-collar. Like goal-derived categories, such thematic categories also presumably arise through experience with common situations and activities, and so they are helpful in planning future activities and in understanding current events. (Ross & Murphy, 1999, propose that their script categories of foods serve a planning function as well.) As Experiment 10 showed, such thematic categories may be associated with taxonomic categories in a way that helps people to perform taxonomic identification. Thus, taxonomic and thematic categories may be related, even though they are qualitatively different. Indeed, Barsalou (1991) points out that taxonomic concepts are the constituents of goal-derived concepts. The inexpensive vacation concept includes references to vehicles, locations, equipment, activities, and other taxonomic concepts.

E. M. Markman (1989) emphasized that thematic information is important for children to know, arguing against the (perhaps implicit) assumption of earlier literature that suggested that taxonomic categories are "real knowledge" and thematic relations are only a primitive, less useful form of knowledge. Our results show that even adults are strongly influenced by thematic knowledge, which suggests that theories of concepts need to understand better how these different forms of organization interact and influence one another.

A final connection of thematic categories to the mainstream categorization literature can be found in Ross's (1996, 1997, 1999) research on category use. Ross points out that when one learns to categorize an object in the real world, one is often not simply viewing it and hearing the category name. Instead, one often interacts with the object (e.g., plays with the puppy, uses the cross-cutting handsaw to cut something), which provides further information about the category as a whole. In some cases, the use provides information about the internal relations of the item's features, for example, the spine on the cross-cutting handsaw helps to keep it stiff, thereby ensuring straighter cuts. In other cases, the use provides information about how the item is related to other objects or activities, for example, a cabinetmaker might use a cross-cutting handsaw but a construction worker probably would not. Ross's work shows that using a category outside of the learning context (without feedback) can then affect how new items are categorized.

The literature just reviewed emphasizes that knowledge of concepts is not acquired simply for its own sake but has important connections to other parts of knowledge, including common activities, functions, relations, and goals. This aspect of concepts addresses the *external relations* of the concept to other concepts or knowledge structures, whereas the traditional view of concepts focuses on the *internal structure* of the concept—the category's definition or family resemblance or exemplars. As suggested in the introduction, thematic categories are also related to knowledge-based theories of concepts. In the knowledge (or theory) approach, concepts are said to be closely tied to more general knowledge of a domain, such that the concept is restricted by the prior knowledge that people have of that domain. By the same token, learning a new concept can influence one's general knowledge when it contains new information (see Murphy, 1993, 2000; or Heit, 1997, for reviews of this approach.) General knowledge is analogous to thematic relations in that it is an external source of information that impinges on the category representation. For example, general knowledge about biology constrains the kinds of concepts that people form about a new animal, not because of information directly observed about the new concept, but because the external information relates to the properties of the new concept and is used to constrain what is learned. As a result, consistent knowledge makes a new concept easier to learn, and inconsistent knowledge makes it more difficult to learn (Murphy & Allopenna, 1994). Analogously, our Experiment 10 showed that thematic information can influence people's categorization into taxonomic categories. Although a leash is not a kind of animal, the fact that leashes are used on animals may be used to help identify something as an animal. Just as general knowledge about taxonomic categories influences some induction (Heit & Rubinstein, 1994; Kalish & Gelman, 1992; Ross & Murphy, 1999) so do thematic relations.

Although thematic relations may not rise to the level of general knowledge (much less theories), the fact that they influence categorization and other tasks may be related to the phenomena discussed by knowledge-based approaches to concepts. Furthermore, they certainly reveal further the limitations of similarity-based approaches to concepts (see Kaplan & Murphy, 2000; Murphy, 1993; Murphy & Medin, 1985), because thematically related items are usually not similar.

More generally, although thematic categories have traditionally been viewed as antithetical to the development of a mature taxonomic conceptual system, our findings of consistent use of thematic categories are in fact consistent with a number of different topics in the psychology of concepts. That is, there has been a gradual increase of interest in the external relations of concepts and how these influence initial acquisition and later use of the concept. In this respect, although the present findings are at odds with the usual assumption of adults thinking taxonomically, they comport well with other recent discoveries in the psychology of concepts. Future research will have to answer the questions this work has raised about how internal conceptual structure is integrated with external structures such as thematic relations in conceptual representation.

References

Ahn, W. (1991). Effects of background knowledge on family resemblance sorting and missing features. In *Proceedings of the 13th Annual Conference of the Cognitive Science Society* (pp. 203–208). Hillsdale, NJ: Erlbaum.

Ahn, W., & Medin, D. L. (1989). A two-stage categorization model of family resemblance sorting. In *Proceedings of the 11th Annual Conference of the Cognitive Science Society* (pp. 315–322). Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1976). *Language, memory, and thought.* Hillsdale, NJ: Erlbaum.

Annett, M. (1959). The classification of instances of four common class concepts by children and adults. *British Journal of Educational Psychology, 29,* 223–236.

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition, 10,* 82–93.

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition, 11,* 211–227.

Barsalou, L. W. (1991). Deriving categories to achieve goals. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 27, pp. 1–64). New York: Academic Press.

Bassok, M., & Medin, D. L. (1997). Birds of a feather flock together: Similarity judgments with semantically rich stimuli. *Journal of Memory and Language, 36,* 311–336.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detection and judging objects undergoing relational violations. *Cognitive Psychology, 14,* 143–177.

Boster, J. S., & Johnson, J. C. (1989). Form or function: A comparison of expert and novice judgments of similarity among fish. *American Anthropologist, 91,* 866–889.

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA: MIT Press.

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.

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

Davidon, R. S. (1952). The effects of symbols, shift, and manipulation upon the number of concepts attained. *Journal of Experimental Psychology, 44,* 70–80.

Denney, D. R. (1975). Developmental changes in concept utilization among normal and retarded children. *Developmental Psychology, 11,* 359–368.

Denney, D. R., & Moulton, P. A. (1976). Conceptual preferences among preschool children. *Developmental Psychology, 12,* 509–513.

Denney, N. W. (1972). A developmental study of free classification in children. *Child Development, 43,* 221–232.

Denney, N. W. (1974). Evidence for developmental changes in categorization criteria. *Human Development, 17,* 41–53.

Denney, N. W., & Ziobrowski, M. (1972). Developmental changes in clustering criteria. *Journal of Experimental Child Psychology, 13,* 275–282.

Dunham, P., & Dunham, F. (1995). Developmental antecedents of taxonomic and thematic strategies at 3 years of age. *Developmental Psychology, 31,* 483–493.

Fodor, J. A. (1972). Some reflections on L. S. Vygotsky's "Thought and Language." *Cognition, 1,* 83–95.

Gelman, S. A., Collman, P., & Maccoby, E. E. (1986). Inferring properties from categories versus inferring categories from properties: The case of gender. *Child Development, 57,* 396–404.

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.

Gentner, D., & Markman, A. B. (1994). Structural alignment in comparison: No difference without similarity. *Psychological Science, 5,* 152–158.

Goldman, A. E., & Levine, M. (1963). A developmental study of object sorting. *Child Development, 34,* 649–666.

Goldstone, R. L. (1994). Similarity, interactive activation, and mapping. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 3–28.

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

Heit, E. (1997). Knowledge and concept learning. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts and categories* (pp. 7–41). Hove, England: Psychology Press.

Heit, E., & Rubinstein, J. (1994). Similarity and property effects in inductive reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 411–422.

Huttenlocher, J., & Smiley, P. (1987). Early word meanings: The case of object names. *Cognitive Psychology, 19,* 63–89.

Inhelder, B., & Piaget, J. (1964). *The early growth of logic in the child.* London: Routledge & Kegan Paul.

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

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

Kaplan, A. S. (1999). *Alignability and prior knowledge in category learning.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Kaplan, A. S., & Murphy, G. L. (2000). Category learning with minimal knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 829–846.

Kaplan, A. S., & Murphy, G. L. (1999). The acquisition of category structure in unsupervised learning. *Memory & Cognition, 27,* 699–712.

Keil, F. C. (1987). Conceptual development and category structure. In U. Neisser (Ed.), *Concepts and conceptual development* (pp. 175–200). New York: Cambridge University Press.

Lamberts, K., & Shanks, D. (Eds.). (1997). *Knowledge, concepts, and categories.* Cambridge: MIT Press.

Lassaline, M. E., & Murphy, G. L. (1998). Alignment and category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 144–160.

Lopez, A., Atran, S., Coley, J. D., Medin, D. L., & Smith, E. E. (1997). The tree of life: Universal and cultural features of folkbiological taxonomies and inductions. *Cognitive Psychology, 32,* 251–295.

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

Lucariello, J., & Rifkin, A. (1986). Event representations as the basis for categorical knowledge. In K. Nelson (Ed.), *Event knowledge: Structure and function in development* (pp. 189–204). Hillsdale, NJ: Erlbaum.

Luria, A. R. (1976). *Cognitive development: Its cultural and social foundations.* Cambridge, MA: Harvard University Press.

Malt, B. C. (1995). Category coherence in cross-cultural perspective. *Cognitive Psychology, 29,* 85–148.

Mandler, J. M. (1979). Categorical and schematic organization in memory. In C. R. Puff (Ed.), *Memory organization and structure* (pp. 259–299). New York: Academic Press.

Mandler, J. M. (1983). Representation. In J. H. Flavell & E. M. Markman (Eds.), *Handbook of child psychology, vol. 3: Cognitive development* (pp. 420–494). New York: Wiley.

Markman, A. B., & Gentner, D. (1993). Splitting the differences: A structural alignment view of similarity. *Journal of Memory and Language, 32,* 517–535.

Markman, A. B., & Wisniewski, E. J. (1997). Similar and different: The differentiation of basic-level categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 54–70.

Markman, E. M. (1981). Two different principles of conceptual organization. In M. E. Lamb & A. L. Brown (Eds.), *Advances in developmental psychology* (pp. 199–236). Hillsdale, NJ: Erlbaum.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction.* Cambridge, MA: MIT Press.

Markman, E. M., & Callanan, M. A. (1983). An analysis of hierarchical classification. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2, pp. 325–365). Hillsdale, NJ: Erlbaum.

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

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

Medin, D. L., Lynch, E. B., Coley, J. D., & Atran, S. (1997). Categorization and reasoning among tree experts: Do all roads lead to Rome? *Cognitive Psychology, 32,* 49–96.

Medin, D., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179–195). New York: Cambridge University Press.

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

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

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

Meyer, D. E., & Schvaneveldt, R. W. (1976). Meaning, memory structure, and mental processes. *Science, 192,* 27–33.

Murphy, G. L. (1993). Theories and concept formation. In I. Van Mechelen, J. Hampton, R. Michalski, & P. Theuns (Eds.), *Categories and concepts: Theoretical views and inductive data analysis* (pp. 173–200). New York: Academic Press.

Murphy, G. L. (2000). Explanatory concepts. In F. C. Keil & R. A. Wilson (Eds.), *Explanation and cognition* (pp. 361–392). Cambridge, MA: MIT Press.

Murphy, G. L. (in press). Causes of taxonomic sorting by adults: A test of the thematic-to-taxonomic shift. *Psychonomic Bulletin & Review.*

Murphy, G. L., & Allopenna, P. (1994). The locus of knowledge effects in concept learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 904–919.

Murphy, G. L., & Lassaline, M. E. (1997). Hierarchical structure in concepts and the basic level of categorization. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts, and categories* (pp. 93–131). Cambridge: MIT Press.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92,* 289–316.

Murphy, G. L., & Ross, B. H. (1994). Predictions from uncertain categorizations. *Cognitive Psychology, 27,* 148–193.

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,* 572–586.

Nelson, K. (1977). The syntagmatic-paradigmatic shift revisited: A review of research and theory. *Psychological Bulletin, 84,* 93–116.

Olver, R. R., & Hornsby, J. R. (1967). On equivalence. In J. S. Bruner, R. R. Olver, & P. M. Greenfield (Eds.), *Studies in cognitive growth* (pp. 68–85). New York: Wiley.

Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review, 97,* 185–200.

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition, 3,* 519–526.

Pazzani, M. J. (1991). The influence of prior knowledge on concept acquisition: Experimental and computational results. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 416–432.

Proffitt, J. B., Coley, J. D. & Medin, D. L. (2000). Expertise and category-based induction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 811–828.

Regehr, G., & Brooks, L. R. (1995). Category organization in free classification: The organizing effect of an array of stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 347–363.

Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior, 14,* 665–681.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum.

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

Ross, B. H. (1996). Category representations and the effects of interacting with instances. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1249–1265.

Ross, B. H. (1997). The use of categories affects classification. *Journal of Memory and Language, 37,* 240–267.

Ross, B. H. (1999). Postclassification category use: The effects of learning to use categories after learning to classify. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 743–757.

Ross, B. H., & Murphy, G. L. (1999). Food for thought: Cross-classification and category organization in a complex real-world domain. *Cognitive Psychology, 38,* 495–553.

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

Scribner, S. (1974). Developmental aspects of categorized recall in a West African society. *Cognitive Psychology, 6,* 475–494.

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. 148).

Skwarchuk, S., & Clark, J. M. (1996). Choosing category or complementary relations: Prior tendencies modulate instructional effects. *Canadian Journal of Experimental Psychology, 50,* 356–370.

Sloman, S. A. (1993). Feature-based induction. *Cognitive Psychology, 25,* 231–280.

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

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

Smith, E. E., Shafir, E., & Osherson, D. (1993). Similarity, plausibility, and judgments of probability. *Cognition, 49,* 67–96.

Smith, E. E., & Sloman, S. A. (1994). Similarity- versus rule-based categorization. *Memory & Cognition, 22,* 377–386.

Spalding, T. L., & Murphy, G. L. (1996). Effects of background knowledge on category construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 525–538.

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

Vygotsky, L. S. (1962). *Thought and language.* Cambridge, MA: MIT Press.

Wattenmaker, W. D., Dewey, G. I., Murphy, T. D., & Medin, D. L. (1986). Linear separability and concept learning: Context, relational properties, and conceptual naturalness. *Cognitive Psychology, 18,* 158–194.

Waxman, S. R., & Namy, L. L. (1997). Challenging the notion of a thematic preference in young children. *Developmental Psychology, 33,* 555–567.

Wisniewski, E. J., & Bassok, M. (1999). What makes a man similar to a tie? Stimulus compatibility with comparison and integration. *Cognitive Psychology, 39,* 208–238.

Appendix A

Stimuli in Previous Developmental Classification Studies That Showed Adults' Preferences for Taxonomic Categories or Similarity Justifications for Object Groupings

Annett (1959): Pictorial stimuli in object grouping and grouping justification

Animal: cow, bird, fish, butterfly
Vehicle: car, train, ship, aeroplane
Plants: tree, flower, apple, toadstool
Furniture: chair, desk, clock, television

Goldman & Levine (1963): Groups of physical objects for grouping justification

Red: red paper circle, eraser, plate, ball, red candle, red poker chip, red pencil, red paper square
Round: red paper circle, plate, ball, poker chip
Metal: all silverware, lock, keys, nails, bell, pliers
Silverware: all silverware (including toy set)
Wood: screwdriver, toy hammer, noisemaker, pipe, match box, pencil, matches, wood block
Smoking: pipe, match box, matches
Pairs: circles, forks, spoons, knives, cigarettes, poker chips, candles, sugar cubes, crackers, cigars, matches, nails, keys
Edible: sugar cubes, crackers
Squares: filing card, red square, sugar cube, crackers, match box, wood block, green rectangle
Paper: filing card, red square, green rectangle
Round: paper circles, bell, ball, poker chips, plate
Tools: pliers, small hammer, screwdriver
White: cigarettes, filing card, dog, sugar, crackers, large candle
Smoking: pipe, cigarettes, both cigars, matches, match box
Toys: small hammer, small silverware, ball, noisemaker, dog, imitation cigar, bell

Olver & Hornsby (1967)

Experiment 1: Verbal stimuli in similarity and difference judgments
Array 1: banana, peach, potato, meat, milk, water, air, germs
Array 2: bell, horn, telephone, radio, newspaper, book, painting, education
Experiment 2: Pictorial stimuli in object grouping and grouping justification
a bee, carrots, dogs, an umbrella, an airplane, balloons, a garage, a lamp, a bird, a saw, a tree, a comb, a crow, a clock, a faucet, a shoe, a doll, gloves, a screw, a fish, a thermometer, a bicycle, a coin, a barn, a coat, a house, boots, scissors, a sailboat, a candle, a pumpkin, a sword, flowers, sun, a ruler, a pie, a hammer, a taxi, a rabbit, a telephone, an apple, nails

Smiley & Brown (1979): Sample of pictorial stimuli labeled with their names in the matching-to-target task

Target	*Taxonomic*	*Thematic*
Bird	Robin	Nest
Needle	Pin	Thread
River	Lake	Boat
Net	Rope	Fish
Sheep	Goat	Wool
Bee	Butterfly	Honey
Cow	Pig	Milk
Crown	Hat	King
Spider	Grasshopper	Web
Dog	Cat	Bone

(Appendixes continue)

Appendix B

Critical Pairs in Experiment 10's Category Verification Task

ANIMAL	*OCCUPATION*
LEASH–DOG	DRUG STORE–PHARMACIST
CHEESE–MOUSE	POST OFFICE–MAILMAN
DESERT–CAMEL	MEDICINE–DOCTOR
NEST–BIRD	CIRCUS–CLOWN
JUNGLE–LION	DRAMA–ACTOR
SADDLE–HORSE	COURT–JUDGE
BARN–PIG	RESTAURANT–WAITRESS
LITTER BOX–CAT	GOVERNMENT–POLITICIAN
HUMAN DWELLING	*BEVERAGE*
ESKIMO–IGLOO	CALCIUM–MILK
STUDENT–DORMITORY	PARTY–BEER
PRINCE–CASTLE	BREAKFAST–JUICE
MORTGAGE–HOUSE	CAFFEINE–COFFEE
BUILDING	*VEHICLE*
TEACHER–SCHOOL	PILOT–PLANE
PRIEST–CHURCH	RAILROAD–TRAIN
TRAVEL–AIRPORT	STRETCHER–AMBULANCE
NUN–CONVENT	HYDRANT–FIRE TRUCK
INSECT	*WRITING INSTRUMENT*
HONEY–BEE	BLACKBOARD–CHALK
WEB–SPIDER	PAPER–PENCIL
COUNTRY	*CLOTHING*
PARIS–FRANCE	HANDS–GLOVES
TOKYO–JAPAN	FEET–SHOES
FURNITURE	*CITY*
SLEEP–BED	DISNEY–ORLANDO
LIGHT–LAMP	CASINO–LAS VEGAS
SPORT	*MEAT*
STADIUM–FOOTBALL	MUSTARD–HOT DOG
BAT–BASEBALL	A1 SAUCE–STEAK

Received August 5, 1997
Revision received August 12, 1999
Accepted February 1, 2000 ■