# How Inferences About Novel Domain-Related Concepts Can be Constrained by Structured Knowledge

Michelene T. H. Chi, Jean E. Hutchinson, and Anne F. Robin
*University of Pittsburgh*

Three studies are focused on (a) the definition of *structure* in a specific domain of knowledge (in this case, dinosaurs), and (b) the relationship between how knowledge is structured and how it is used. The evidence suggests that the knowledge of children who are experts on dinosaurs is structured hierarchically into well-defined families and family groups, and, within each level of the hierarchy, such knowledge appears more locally cohesive. Greater hierarchical structure allows expert children to use domain features to generate causal explanations, use categorical reasoning, induce attributes about novel dinosaurs, and sort dinosaurs based on well-defined family types. Novices, in contrast, often use irrelevant features for causal explanations and infer attributes about novel dinosaurs based on explicitly depicted features. The consequences of hierarchically structured knowledge is that expert children can use it to constrain their inferences about novel dinosaurs, whereas novices must rely on their general world knowledge, thereby making less accurate and often inappropriate inferences.

A central concern for developmental psychologists is the issue of structural change. Do knowledge structures undergo major changes as children mature? According to Piaget, structural change is "global" in the sense that children at each developmental stage are supposed to have acquired a particular representational structure within which they can reason logically across all domains. The particular structure within which children are capable of reasoning is represented by the logical inferences in which children can engage, such as transitive inferences or reversible operations, and so forth. In a sense, the structures in Piaget's representations are operators that can act on any do-

main of knowledge. Thus, when children have reached the stage in which they are capable of making transitive inferences, they should be able to infer in a transitive way for all domains. Because these structures constrain children's ability to make logical inferences across all domains, restructuring in cognitive development constitutes a global change.

It is increasingly clear that Piaget's notion of global restructuring is too general. In many instances, having knowledge in a specific domain can overcome any limitations that could have been imposed by the lack of global operators. Yet lacking knowledge in a specific domain also can prevent adults from reasoning logically, even though they are presumed to have the logical operators. (For summaries of this kind of evidence, see review papers by Gelman & Baillargeon [1983] on Piagetian tasks, and Chi & Ceci [1987] on memory-related tasks.) The dilemma created by the role of domain knowledge necessitates the examination of the following questions:

1. What is the structure of a domain of knowledge?
2. What does it mean to say that knowledge of a domain becomes more structured?
3. How does the structure of domain knowledge relate to the child's ability to remember and reason logically?
4. What are the mechanisms that produce the greater structure? This paper focuses on all of these issues to some degree, but principally on the first three issues.

## HOW DOMAIN KNOWLEDGE IS REPRESENTED

To answer the questions relating to how domain knowledge is structured, a representation of the domain knowledge must first be created. There are many ways to represent knowledge of static concepts such as dinosaurs. A popular means is in terms of a network of nodes and links. A *node* usually corresponds to any concept that is designated, and a *link* corresponds to the relation between nodes. However, little empirical work has been done to define what it means to say that one network is more or less structured than another.

In our first study (Chi & Koeske, 1983) the focus was primarily on how to describe the degree of structure that exists within a domain of knowledge. In our work when the degree of structure within a representation of static concepts is discussed, reference is not primarily to the quantity of knowledge. Although the amount of knowledge necessarily determines to some degree the structure of the knowledge, it is important to unconfound the two factors if possible. Of most concern are those characteristics of a structure that can affect the use of

the knowledge. Hence, the kinds of representational properties predicted as important in our work are the coherence of a structure and the hierarchical nature of a particular structure. *Coherence* refers to both the integratedness of the whole structure (henceforth called *global coherence*) and the integratedness of the substructures that form the hierarchy (*local coherence*). In this paper, because the concepts discussed have a clear classification scheme, *global coherence* refers to the presence of *hierarchical organization.* What constitutes evidence for *coherence* and *hierarchy* will be described later.

It is postulated in this study that two sets of identical concepts and attributes can have different degrees of local and global coherence simply by the way the concepts are linked. In Figure 1, for example, both the top and the bottom networks have identical concepts (the circles) and attributes (the triangles), and the same number of links. However, because the concepts and attributes are linked in different ways, it is hypothesized that the behavioral manifestations corresponding to the two networks will be different. Thus, it could be said that Structure I is "better" or "more structured" than Structure II. In a sense, in these hypothetical networks, the content and amount of knowledge have been controlled.

Within these hypothetical networks, *local coherence* of the substructures can be defined in terms of the patterns of interlinkages and attribute-sharing among concepts, and *hierarchy* (or *global coherence*) in terms of the patterns of relationships among substructures. Comparing Structures I and II of Figure 1, it might be said that Structure I is more globally as well as more locally coherent than is Structure II. Structure I is more hierarchical than Structure II because it has well-differentiated substructures which together constitute the larger structure (Concepts A and C form one substructure; Concepts B and D form another). Structure II does not have any hierarchical embedding. The substructures in Structure I are also more locally coherent in the sense that the pairs of concepts share attributes to a greater degree than any subsets of concepts in Structure II. Thus, it might be said that Concepts A and C in Structure I are more related than Concepts A and C in Structure II. (Such relatedness can be manifested in clustered recall, for example.)[1]

One way to describe the relationship between local and global coherence is in terms of differentiation. Structure I is more differen-

<hr />

[1]Structure II, of course, could be considered locally coherent in the context of a larger structure. However, if both Structures I and II were embedded within the context of a larger structure, Structure I would probably still maintain greater global coherence than Structure II in the sense of having a more well-defined hierarchical structure.

STRUCTURE I
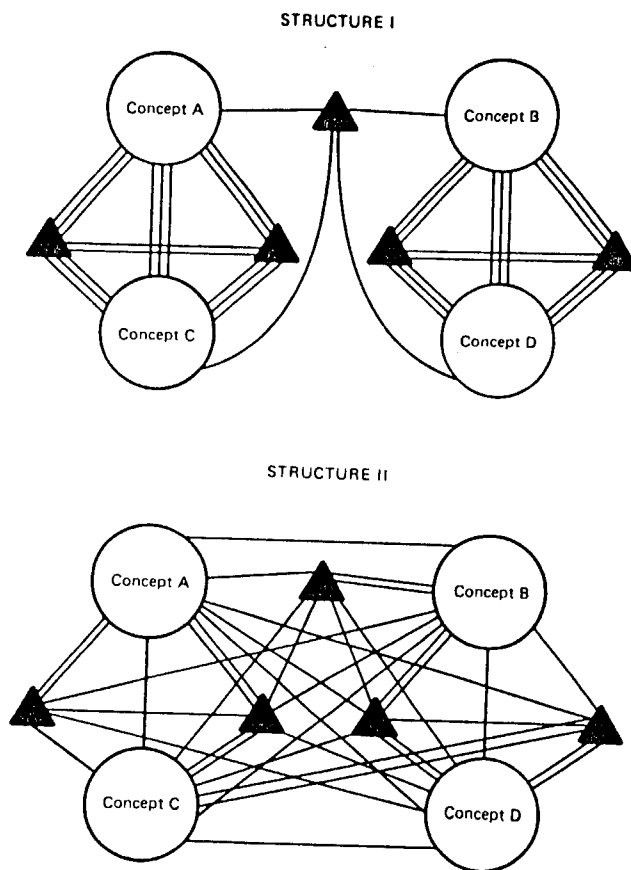
STRUCTURE II

**Figure 1.** Two possible structures containing the same concepts and attributes and the same number of links.

STRUCTURE I A

STRUCTURE II A

**Figure 2.** How structures depicted in Figure 1 can be made more or less coherent.

tiated than Structure II in that it has substructures that share more attributes within themselves than they do between themselves. The differentiation of Structure II can be increased, thereby causing it to become more hierarchical if direct and indirect links are added between Concepts A and C and between Concepts B and D to form coherent substructures (see Structure IIA, Figure 2). The differentiation of Structure I can also be decreased, causing it to become less hierarchical if links from the concepts and attributes of one substructure are added to the concepts and attributes of the other substructure (see Structure IA, Figure 2).

Data have been collected to support this notion that structure may be defined in terms of coherence and hierarchy. In our first study (Chi & Koeske, 1983), a 4½-year-old child's knowledge of dinosaur concepts was elicited by asking him to: (a) generate a sequence of dinosaur names, (b) identify a dinosaur from dinosaur attributes that we provided, and (c) generate dinosaur attributes when given dinosaur names. The sequential generation of dinosaur names provided dinosaur-dinosaur linkages. The identification of a dinosaur from its attributes, as well as the generation of attributes from dinosaur names, provided the data for the dinosaur-attribute links. In mapping the net-

work, the only assumption made was that no redundant nodes were depicted. Thus, if the child generated the attribute of "sharp teeth" for two different dinosaurs, then these two different dinosaurs were shown to share the same "sharp teeth" node.

For simplification and comparison purposes, the total set of dinosaurs that the child generated was divided into two subsets of 20 each; this was done on the basis of how familiar the child was with each of the dinosaurs. The classification into each of the two categories, called *better-known* and *lesser-known,* was done on the basis of two external sources: the mother's judgment of how much her child knew about the dinosaurs, and the frequency with which a given dinosaur was mentioned in the books that were read to the child. Each group of 20 dinosaurs could also be divided into seven subgroups on the basis of their family membership, as discussed in the dinosaur books. Members of a dinosaur family share a common set of explicit attributes (i.e., physical features) and implicit attributes (i.e., diet, habitat).

For further simplification, the networks drawn for the better- and lesser-known dinosaurs consisted of the attributes and links of a subset of seven dinosaurs from each set of 20. These seven dinosaurs (the targets) were chosen so that they were matched on the number of attributes that the child produced and recognized, as assessed in the tasks *b* and *c* mentioned earlier. The goal was to compare two subsets of dinosaurs for which the child knew approximately the same number of attributes, because we were not interested in the effect produced by having a greater quantity of knowledge. In this case, the child knew about five attributes for each of the seven target dinosaurs from both the better-known and lesser-known sets. However, although the two subsets of seven targets were matched on the number of attributes known about each, each subset of seven belonged to a different number of families. The seven better-known target dinosaurs belonged to only two families—armored dinosaurs and giant plant-eaters—whereas the seven lesser-known target dinosaurs belonged to five families: armored dinosaurs, small bird eaters, water dwellers, duckbills, and early meat-eaters.

Once the target dinosaurs were represented in terms of their links and attributes in network form, coherence and hierarchy could be quantified in an operationally definable way. Basically, the networks corresponded to the hypothetical ones depicted in Figure 1. The network of the seven better-known dinosaurs (see Figure 3) corresponded to the network depicted in Structure I, with two families as substructures. (Target dinosaurs are enclosed by rectangles and families are enclosed by dotted circles in Figure 3.) Which subset of target

dinosaurs was included within a family was determined independently from information provided in books. Families in the network were thus represented by clusters of interrelated concepts and attributes. A family was not represented by a single node, as is often done in simple hierarchical representation of concepts. The network of the lesser-known dinosaurs (shown in Figure 4) corresponded to the network depicted in Structure II, but with potentially five substructures corresponding to the five families.

For the better-known dinosaur network in Figure 3, then, evidence of global coherence can be gleaned from the existence of multiple links among target dinosaurs within the same family, in conjunction with only single links between target dinosaurs from different families. Evidence of local coherence can be seen in the greater extent to which dinosaurs of the same family share attributes. This pattern of greater within-family and lesser between-family linkages was evident for the entire set of 20 better-known dinosaurs as well, not just the seven target ones. But the pattern of linkages for the lesser-known dinosaurs was more diffuse and less differentiated. No sharp contrast existed between the linkages within and the linkages between families. Like Structure II of Figure 1, the connections for the lesser-known dinosaurs were much more uniformly distributed among the families. Thus, it was concluded that there was no evidence of local or global coherence in the network of the lesser-known dinosaurs.

In this first study (Chi & Koeske, 1983), there were also three additional measures of the child's dinosaur knowledge. The child was asked to (a) categorize the 40 dinosaurs in a free sorting task; (b) freely recall each list of 20 dinosaurs after it was read out loud to him; and (c) a year later, name each of the 40 dinosaurs when a picture of it was shown. In the free recall task, the child could recall about 45% of the dinosaurs taken from the better-known set, and only 22% of the dinosaurs taken from the lesser-known set. Furthermore, the recall from the better-known list had a higher clustering score using the Bousfield (1953) ratio of repetition, RR = .67, than the clustering score of the lesser-known set, RR = .17. The better recall and clustering for the better-known set was interpreted as a validation of the analyses of the differential coherence and hierarchy of representations as a function of knowledge. The 20 better-known dinosaurs must have had greater local coherence (more shared links within each family) and more differentiated family structures in order to produce a higher clustering score and better recall. The 20 dinosaurs of the lesser-known set may not have had as cohesive a pattern of linkages. That after a year, the child could name many more dinosaurs from the
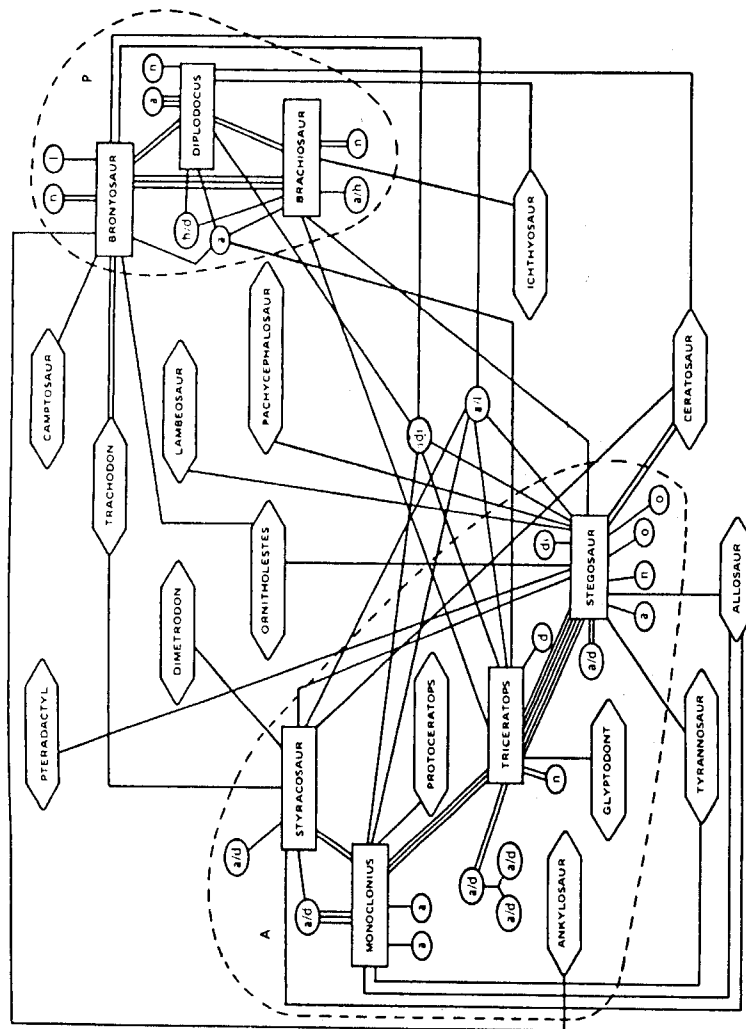
**Figure 3.** Network for the better-known set of dinosaurs. From "Network Representation of a child's dinosaur knowledge" by M. T. H. Chi & R. Koeske, 1983, *Developmental Psychology, 19,* p. 33. Copyright 1983 by the American Psychological Association. Reprinted by permission of the American Psychological Association.
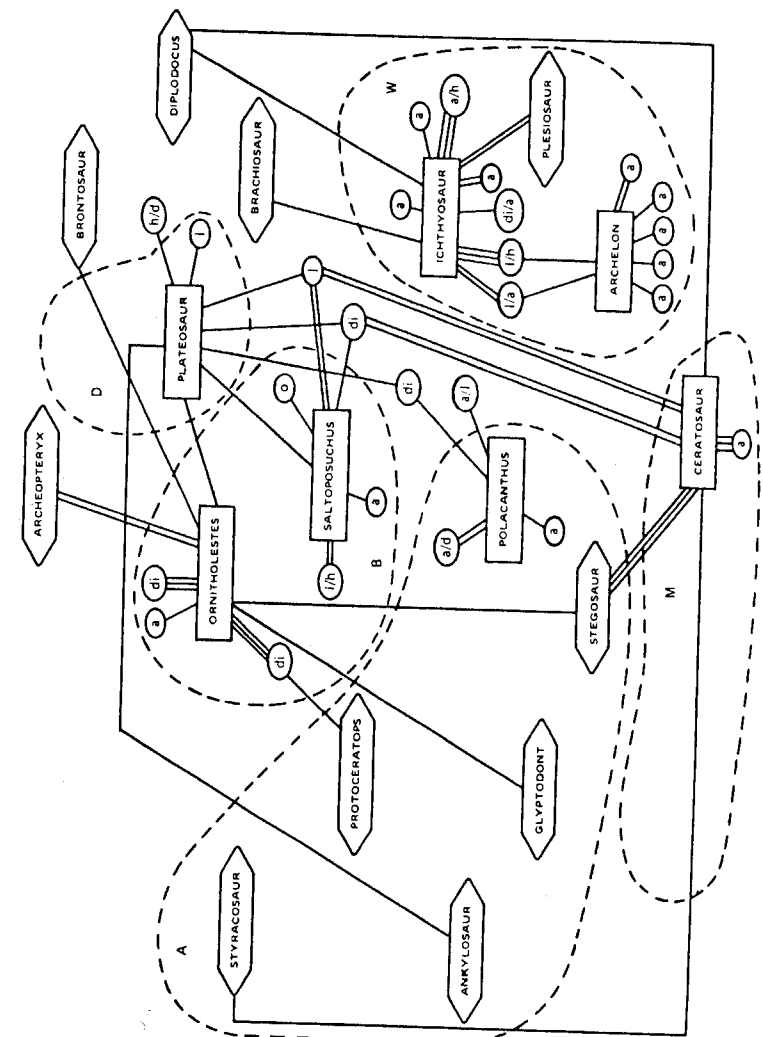
**Figure 4.** Network for the lesser-known set of dinosaurs. From the same source as cited for Figure 3, p. 34. Copyright 1983 by the American Psychological Association. Reprinted by permission of the American Psychological Association.

better-known than the lesser-known set (11/20 vs. 2/20) also suggested that the attributes required for identification were not as strongly associated for the lesser-known dinosaurs. (This distinction was not evident a year earlier, when the child could easily identify all 40 dinosaurs by name.)

The results from the sorting task were the most intriguing because they were not predictable from the network representation. (The sorting data were not included in the Chi & Koeske paper [1983] because at the time it was not a robust result. But data from two other studies, reported later in this paper, now replicate this finding.) The child sorted the 20 better-known dinosaurs very quickly, without any hesitation or pauses, into two groups: meat-eaters and plant-eaters. This pattern of sorting was consistent across two separate trials. In contrast, he sorted the 20 lesser-known dinosaurs with a great deal of hesitancy and uncertainty. Furthermore, in three trials, he could not reach a consistent sort. The first sort was divided between plant- and meat-eaters, the second sort combined meat-eaters with land-dwellers, and included a new group of water-dwellers, and the third sort included land-plant-eaters, water-plant-eaters, and water-meat-eaters. Thus, the child's sorting behavior for the better-known dinosaurs resembled that of older children's sorting behavior in general, in that it was very fast, exhaustive, and stable across trials. His sorting for the lesser-known dinosaurs typifies the behavior of younger children: It was hesitant, inconsistent, and slow.

The dilemma about the child's sorting behavior was that the groups he created did not map directly onto the family structures that were represented in the networks nor in the family structures that were manifested in the clustering pattern of the free recall. A higher level of the hierarchy may have existed that was not captured. That is, the families may have coalesced into two superordinate groups: meat- and plant-eaters. The more consistent sorting of the better-known dinosaurs could be interpreted to mean that the families of these 20 dinosaurs have coalesced clearly into these two groups. Had all 20 of the better-known dinosaurs been represented in a network, how the families related to each other in a hierarchical way might have been demonstrated.

How children coalesce families into higher-level groups that contrast with each other can be inferred. Children may come to realize that there are one or more features that the majority of the members of one set of families possess, and the majority of the members of a contrasting set of families do not. If the network for the better-known dinosaurs (see Figure 3) is re-examined, it is seen that most of the target dinosaurs have links to the same diet attribute, in this case plant

eating (as represented by *di,* and highlighted there by shading). These linkages can be re-interpreted to mean that this is the critical attribute which one family (the armored dinosaurs) shares with the other family (the giant plant-eaters). That is, the "plant-eating diet" node can be thought of as the main attribute that is shared by the two families of dinosaurs depicted in Figure 3; perhaps this is the way that hierarchical groups are formed.

In contrast to the network for better-known dinosaurs, the network for the lesser-known dinosaurs shown in Figure 4 depicts many *di* nodes. These nodes correspond to the mention of different diets associated with different dinosaurs. For example, some dinosaurs were said to eat eggs, others to eat meat, and still others to eat mushy plants. The child clearly had not generalized the similarities of the diet across dinosaurs, and did not see the families as sharing the same diet.[2] Furthermore, many of the lesser-known dinosaurs' diets were incorrect and inconsistent. For example, Ornitholestes is indicated to eat plants as well as eggs, and Plateosaur is indicated to eat meat as well as plants. These inconsistencies also must make it difficult to clearly segregate the families in the child's representation of the lesser-known dinosaurs. The presence of both lack of generalization and inconsistency suggests that well-defined hierarchical clustering of families has not emerged for the lesser-known dinosaurs.

## RELATIONSHIP BETWEEN STRUCTURE AND USE OF KNOWLEDGE

The major focus of the first study was to define a way of representing declarative knowledge of concepts (such as dinosaurs). To give some validity to the representations depicted, several measures (recall, retention, and sorting) were taken on a single child. The results of these measures all corresponded to some degree to the distinction made between more- or less-structured knowledge, a distinction based primarily on the existence of links among concepts, rather than the existence of concept nodes.

The second study (Gobbo & Chi, 1986) added greater support for the idea that a more locally coherent and integrated knowledge structure is one in which there are more connections or stronger links among the concepts and attributes within a family. In this second study, the effects of knowledge structure on the use of that knowl-

---

[2]Generalizing eating eggs and eating meats into one meat-eating node is consistent with the macrorule of *generalization* proposed by Kintsch and van Dijk (1978) for the substitution of a sequence of propositions by the general proposition which denotes their immediate superset.

edge was also explored. Five 7-year-old children who were experts in the domain of dinosaurs were compared with five 7-year-olds who were novices. Expertise was defined on the basis of a pretest of dinosaur knowledge.

In one task, children were asked to sort 20 pictures of better-known dinosaurs, similar (but not identical) to those used in the first study. These results will be reported in conjunction with the results from the third study, because the same task was also presented to additional subjects in that study. Two other measures were used to evaluate the cohesiveness of a representation.

First, it was found that when children were simply asked to generate comments about dinosaurs, the expert children's protocols contained more connecting words (such as *because* and *if* ) than the protocols of the novice children. This finding suggests that when experts activate a dinosaur concept node, several other related attributes and concepts also get activated with high strength and, thus, they feel compelled to continue and state them. The connecting words serve the function of accessing additional knowledge that is associated with the information that is presented in the dinosaur picture. Consider the quote: "And he had webbed feet, *so* that he could swim, and his nose was shaped like a duck's bill, *which* gave him his name." The propositions generated prior to the connecting words (*webbed feet, duckbill-like nose*) were derived from the external features explicitly depicted in the picture, whereas the propositions generated after the connecting words (*he could swim, gave him his name*) were derived from implicit knowledge associated with those features.

In contrast, novice children's productions tended to be a listing of explicit features that were depicted in the picture of the dinosaur, such as, "He has sharp teeth. He has three fingers. He has sharp fingers, sharp toes, a big tail." Because novice children lack a strongly interconnected set of attributes and concepts, their protocols do not contain a great number of connecting words, nor do they contain inferences derived from the explicit external features. Consistent with this interpretation is the second finding: Expert children are much more likely than novices to change topics in their discussion of each dinosaur. This finding suggests that each topic, such as the manner of defense of a dinosaur, is strongly related to other features of the dinosaur, such as its diet, its habitat, its social behavior, and its physical appearance, in the child's representation. For example, the way a dinosaur defends itself depends on where it lives. If the dinosaur lives in the swamp, it might hide in the water for defense. How the dinosaur defends itself also relates to the kind of protective armor that it has (its appearance). It is easier to understand the attributes of a dino-

saur if it is known how the dinosaur relates in a causal or correlated structure. Expert children, it is suggested, have represented these causal and correlated associations, so that it is difficult for them to discuss one topic without relating it to another, which results in the greater frequency of transitions among the different topics in experts' discussions of each dinosaur.

What is the effect of the structure of children's knowledge on their use of that knowledge? In the second study (Gobbo & Chi, 1986), usage of knowledge in a less traditional way, as opposed to the more traditional measures of recall and retention used in the first study, was examined. One effect of a more integrated and highly associated knowledge structure is hinted at in the analysis of the use of connecting words. That is, expert children, when confronted with a picture of a familiar dinosaur, often mentioned attributes of that dinosaur that were not pictorially depicted (*implicit attributes*) because a visual depiction of the dinosaur triggered other associated attributes and knowledge of that dinosaur. This pattern is seen even more clearly in a propositional analysis of the children's protocols. Both the expert and novice children stated the same total number of propositions concerning the explicit features of a dinosaur (such as the sharpness of its teeth), but the expert children stated four times as many propositions concerning implicit features of a dinosaur (such as where it might live, or how it might attack an enemy).

Another major difference in the way expert and novice children used their knowledge lies in the number of comparisons they made among dinosaurs. Experts not only used comparisons more frequently than novices, but their comparisons included mention of differences as well as similarities. For example, when dinosaurs shared common features, expert children might generate comparative comments such as "They are pretty much like Diplodocus" or "They eat plants, kind of like Stegosaurus." Yet, they were also capable of contrasting dinosaurs: "And it goes in the water, but it doesn't go in like Brontosaurus." These two uses of comparisons by experts suggests that their dinosaur concepts are both more interrelated and more differentiated than those of novices.

How knowledge is used and represented is further revealed by the way children decided upon the diet of a dinosaur. Expert children were found to use predominantly four features to derive diet information. Furthermore, they discriminated contrastive diet classes by using a single feature in both an inclusive and exclusive way. That is, if a feature was present, then the diet was meat-eating, but if the feature was absent, then the diet was plant-eating (or vice versa). The presence and absence of a single feature correlated with diet may have

been the factor that led experts to subcategorize dinosaurs into complementary diet classes. It is speculated that the formation of hierarchical clusters is based on the development of these contrastive discriminations, a view that is consistent with the speculation from the first study that families coalesce into higher-order categories if members of all the families share similar attributes. (For other suggestive evidence, see Chi, 1988.)

Novices did not use features in such a contrastive way. They had different sets of features to determine whether a dinosaur was meat-eating or plant-eating. Consequently, the novice children used twice as many features as the expert children (9 vs. 4). This finding suggests that, for novices, the two diet classes were not complementary, and thus were not well-differentiated, but this is also evident in the expert child's representation of the lesser-known dinosaurs in the Chi and Koeske study (Figure 4). There, the child would sometimes mention contradictory diets for a given dinosaur. Thus, the representation of novices is similar to the expert child's representation for the lesser-known dinosaurs.

In summary, the results of the second study (Gobbo & Chi, 1986) provide additional evidence to suggest that the experts' representation of families was locally coherent; this is shown, for example, by the frequent use of comparisons among dinosaurs of the same family. The results of this second study also indicate that the knowledge about individual dinosaurs was more integrated for the experts; this is shown by the way connecting words were used to draw out knowledge of a dinosaur's implicit attributes, as well as by the way experts frequently changed topics in their discussion of a particular dinosaur. At the same time, experts' representations were also more differentiated than those of the novices, as evidenced by their use of a single set of features to contrast two diet classes, suggesting that they could combine families to form the higher-level categories of plant- and meat-eaters. Thus, the experts' representation of dinosaur knowledge seems to be more hierarchical and more locally coherent than the novices' representation.

## HOW KNOWLEDGE IS USED TO LEARN NEW CONCEPTS

In the third and just completed study of this series, the focus was on the understanding of the same two issues, namely, to define *structure* as it applies to a conceptual domain, and to relate the way the knowledge is structured to how it is used. There were, however, three additional specific goals: (a) to examine how knowledge is used to constrain learning about new domain-related concepts; (b) to formu-

**Table 1.** Profiles of Means (and Standard Deviations) of Expert and Novice Children's General Abilities: Third Study

| Tests | Expert | Novice | U |
|---|---|---|---|
| Forward digit span | 5 (0.7) | 4 (1.0) | 5.5 |
| Backward digit span | 3 (0.7) | 2 (1.8) | 9.5 |
| Raven's colored progressive matrices | 27.2 (3.8) | 26.0 (6.3) | 12.5 |
| Peabody picture vocabulary | 117.6 (25.5) | 96.4 (19.7) | 6.0 |

*Note.* Means are based on raw scores. No significant differences were found between the expert and novice children for any of the tests, using Mann-Whitney $U$ $(p > .1)$.

late a clearer conception of the nature of the knowledge that is embodied in a family (or more broadly speaking, a schema); and (c) to control more carefully the selection of the subject sample, so that they were matched on all dimensions except for their knowledge about dinosaurs. Because this third study has not been previously published, it is reported in greater detail than the first two studies.

Two groups of five children were selected to participate on the basis of their dinosaur knowledge, which was assessed by a pretest (the same as that used in the second study). The expert group scored a mean of 59%, and the novice group scored a mean of 13% on the pretest. Each group of children included 3 boys and 2 girls, and the mean ages were 5 years 11 months for the expert group, and 6 years 3 months for the novice group. The two groups of children were further matched on a number of other measures, such as their forward and backward digit spans, and their scores on the Raven's Colored Progressive Matrices and the Peabody Picture Vocabulary test (see Table 1).

The stimulus materials consisted of line drawings of a set of 20 familiar dinosaur illustrations (the same as the better-known dinosaurs used in the second study), and a set of 20 novel dinosaur illustrations. The familiar set was composed of prototypical members of five taxonomic dinosaur families: duckbills, giant plant-eaters, horned dinosaurs, armored dinosaurs, and early meat-eaters. These familiar dinosaurs were ones that both experts and novices had commonly seen in children's dinosaur books.

The novel set included only dinosaurs which were unfamiliar to both the experts and the novices (i.e., they were not commonly found in children's dinosaur books). Five of the novel dinosaurs were similar to the *prototypical* members of the five dinosaur families just

described. Ten others were constructed, by physically transposing parts of dinosaur illustrations, to represent five *peripheral* family members (one from each family) and five *conflicting* family members. A peripheral dinosaur was composed mostly of features specific to one family with the exception of one conflicting feature that was salient to another family. Members of the conflicting set were made up of physical features characteristic of several families. And the remaining five dinosaurs, named *newly-discovered,* were recently discovered dinosaurs that did not belong to any of the five common families. These dinosaurs were found in sophisticated currently published dinosaur books that children were not likely to have seen.

The experimenter met with each child for approximately four half-hour sessions, over a 2- to 3-week period, to complete the following five tasks (in the following fixed order). In the Sort with Familiar Dinosaurs Task, the 20 familiar dinosaurs were randomly displayed on a table in front of the child. Instructions were to "Put the dinosaurs that you think belong together in the same group." If a child constructed very large groups, he or she was asked to break them down further. Upon completion of the sorting, the child was asked to explain each grouping. No limit was placed on either the number or size of the groups to be formed. This task was designed to indicate the extent to which experts and novices categorize dinosaurs on the basis of family distinctions, as well as their differential use of more perceptually based classification.

In the Free Generation Task, eight of the novel stimuli, randomly selected for each child so that two were chosen from each category (i.e., prototypical, peripheral, conflicting, and newly-discovered), were presented in random order with the instruction to "Tell me the name of this dinosaur and everything you know about it." The purpose of this task was to assess how children infer information about novel domain related concepts.

In the Probe Question Task, children were asked specific questions, in case they were reticent and did not spontaneously generate many remarks in the free generation task. For each of the remaining 12 novel stimuli not used in the free generation task, children were asked five probe questions:

1. What would this dinosaur eat? Why?
2. Describe the place this dinosaur might live? Why?
3. What would this dinosaur's cousin be like? Why?
4. How did this dinosaur defend himself? Why?
5. If this dinosaur only ate _____ (contingent upon child's response to the 1st question) would it be a peaceful dinosaur or would it attack other dinosaurs? Why?

The purpose of this task was to determine the degree to which children integrated and related knowledge about a family of dinosaurs. For example, if they knew about the diet of a dinosaur, they might have classified the dinosaur into either a meat-eating or a plant-eating class. If so, then any categorical knowledge about the meat- or plant-eating classes would constrain additional inferences that they might make about other attributes of the dinosaur. This task was designed to assess this kind of integrated knowledge.

In the Forced Choice Sort, the original groups into which the familiar dinosaurs were initially sorted by each child were re-displayed on a table. Then the child was asked to sort the set of novel dinosaurs either into these existing groups or into new groups. The purpose of this task was to examine the extent to which children perceived the novel dinosaurs as fitting into the family groupings of the familiar dinosaurs.

Finally, in an Oddity Task, the experimenter displayed 10 sets of four familiar dinosaurs one set at a time, saying "Three of these dinosaurs are alike in some way. The other one is different. Which one is different from the others? Why?" Within each set, three of the four dinosaurs were alike on a conceptual basis (i.e., shared habitat, defense mechanisms, diet, and/or family membership). The fourth dinosaur did not fall in the same conceptual grouping, but often shared physical attributes with one or more of the other three dinosaurs. This task should indicate which dimensions (i.e., physical attributes, diet, habitat) the child uses to categorize dinosaurs. The hypothesis was that experts are more likely to categorize only on the basis of conceptual features, whereas novices are more likely to categorize only on the basis of explicit physical features.

### Results From the Free Generation Task

The free generation task was in a sense the purest measure of how children used knowledge to infer new attributes about novel dinosaur concepts because it was the first task utilizing novel stimuli; and because the children were not probed, they were free to say anything they liked about the novel dinosaurs. The free generation task was analyzed in several different ways. These analyses serve to illustrate some general themes that will then be discussed in relation to the other tasks.

*Kinds of knowledge used to generate inferences.* First, the protocols produced by the children were analyzed at the level of propositions, or the smallest possible meaningful statements (see examples in text later). Figure 5 is a classification scheme for the analyses con-
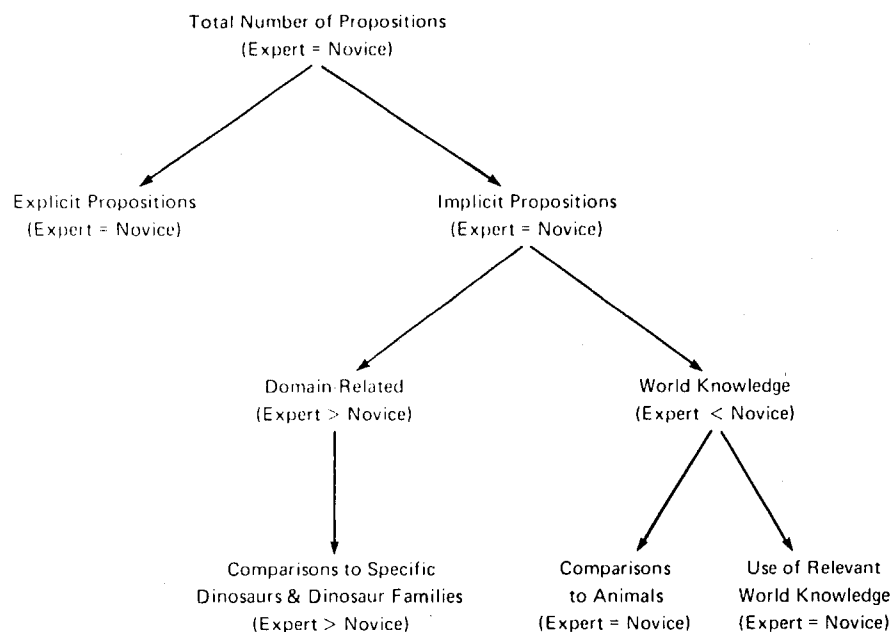
Figure 5. Analyses of propositions produced by experts and novices in the free generation task.

ducted. To assess the reliability of the coding, two independent raters coded one page (approximately 27% of the data) randomly selected from each subject's protocol. The agreement was 83%, and differences were resolved through discussion.

Most of the propositions could be classified as either implicit or explicit, defined in the same way as in the second study. Experts and novices produced about the same total number of propositions across the eight novel dinosaurs (about 66 propositions), which provides assurance that both the novices and the experts tried equally hard to generate inferences about novel dinosaurs. In general, there were no differences between experts and novices in the number of implicit propositions or explicit propositions produced. The results for the explicit propositions were consistent with the results of the second study, but the results for the implicit propositions were inconsistent. The discrepancy most likely resulted from the fact that in the previous study the production of knowledge about familiar dinosaurs was studied, whereas in this study, the dinosaurs were novel. Explicit propositions do not rely on knowledge an individual already has about the dinosaurs, whereas implicit ones do. Thus, the introduction

of novel dinosaurs in the third study had no effect on experts' and novices' production of explicit statements, but did erase the distinction between the number of implicit statements made by experts and novices.

Although there was no difference in the number of implicit propositions produced, there were notable differences in the content of those implicit statements. Implicit propositions were categorized into those that were related to the domain of dinosaurs (e.g., "It eats meat"), and those that were related to world knowledge (e.g., "like a rhinoceros"). The majority of the experts' implicit propositions were domain-related ($M = 19$, of 27), whereas the majority of the novices' implicit propositions were related to world knowledge ($M = 26$, of 33), and the interaction between type of implicit proposition and knowledge was significant, $F(1, 8) = 13.40$, $p < .01$, as shown in Figure 6. Thus, the advantage of having domain knowledge is that experts can use it to make inferences and attributions about new dinosaurs, whereas novices must rely more on world knowledge.

Within the domain-related implicit propositions, the number of comparisons to specific familiar dinosaurs and to dinosaur families were counted. Any comparison that children drew between a novel dinosaur and an existing dinosaur that they knew about was coded as a specific comparison (e.g., "It's like Tyrannosaurus"). Family comparisons were statements such as "like a duckbill" in which the child compared a novel dinosaur to a dinosaur family. Overall, the experts made significantly more Dinosaur and Family comparisons than novices ($M = 9$, of 19 implicit statements for experts vs. $M = 1$, of 7 implicit statements for novices). In other words, about half of the experts' inferences were generated by comparing the novel dinosaur to dinosaurs and families that they knew about, whereas the novices rarely made comparisons for the purpose of generating implicit inferences. (A similar analysis conducted across all the tasks produced comparable results.) This result corroborates findings on comparisons from the second study. Thus knowledge about familiar dinosaurs and dinosaur families constrained experts' inferences about the novel dinosaurs.

It might be argued that the novice children were generally deficient in making any kind of comparisons, in contrast to the experts. When the implicit world knowledge propositions were examined, however, it was found that experts and novices were equally proficient in making comparisons based on world knowledge. Most of these comparisons were comparisons to animals such as "like a rhinoceros." (A similar analysis drawing upon data from all the tasks produced the same results.) Thus, there was no fundamental difference
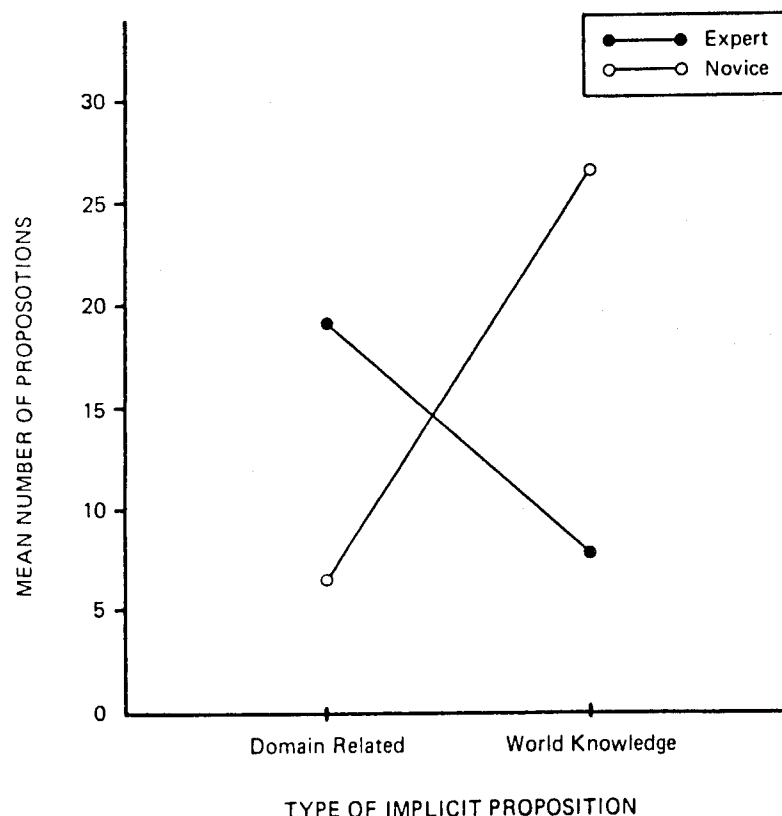
Figure 6. Types of implicit propositions produced by expert and novice children.

in the way expert and novice children generated inferences by using comparisons. The only difference seems to be in the content of their inferences, that is, to what they compared dinosaurs. The expert children had a specific advantage in their ability to compare the novel dinosaurs with known dinosaurs in order to generate potentially useful information about the novel dinosaurs. In contrast, the novice children must have used animals other than dinosaurs as their source of comparison.

It is important to know whether novices can learn anything by drawing inferences based on world knowledge, because they do so much more frequently than experts. After all, an individual must draw on some knowledge in order to learn about a new domain. It has already been seen that novices often compared dinosaurs to animals,

which could yield some useful information. Therefore, the number of implicit world knowledge statements that may be relevant to learning about dinosaurs or animals in general (e.g., "and he's definitely a land animal"), excluding those which were direct comparisons to specific animals, were examined.

On the whole, there was no difference between the experts and the novices in the proportion of implicit attribute statements that made use of "relevant" world knowledge (expert, $M = 55.4\%$; novice, $M = 75.6\%$), $U = 4.5$, $p > .05$, by a Mann-Whitney test. This finding suggests that although novices lacked domain knowledge, they were just as capable of making use of relevant world knowledge to constrain their inductions as expert children. In any context in which totally new information must be learned, one productive way is to look for similarities between what is already known and what is being learned. This result, then, is not surprising. Of course, novices' inductions could not possibly be as accurate as those that expert children generated based upon relevant domain knowledge.

In summary, two conclusions can be drawn from the analyses of the propositions generated in the free generation task. First, no differences appeared between experts and novices in the number of implicit and explicit propositions produced. The equivalence of these measures validates the assumption that we succeeded in obtaining a homogeneous sample of children. For instance, neither group of children was more articulate than the other. Furthermore, there was no difference in the way expert and novice children used general world knowledge to constrain their inferences. That is, when expert and novice children did use general world knowledge to infer new information, they were equally proficient at making comparisons, and they both relied to the same extent on relevant animal-related knowledge to make attributions about the novel dinosaurs. However, the expert children excelled in one regard: They used predominantly domain knowledge (specific dinosaurs and families of dinosaurs) to constrain their inferences. These findings suggest that young children rely on concepts they know the most about when generating inferences about a similar novel concept. For the expert children, other dinosaurs were more useful than animals in general as a source for inferences about novel dinosaurs. This result is consistent with the Carey (1985) finding that young children use people (a category they know a great deal about) as a prototype on which to base attributions about unfamiliar animals.

*Reasoning structures.* To examine more directly how expert and novice children used their knowledge to infer information about the novel dinosaurs, the free generation protocols were coded at a higher

level of analysis: that of reasoning structures. Reasoning structures consisted of one or two sentences that together completed a reasoning chain. Reasoning chains were classified into four types: family/superordinate, dinosaur, animal, and attribute.

*Family* or *superordinate reasoning structures* included all inferences that invoked a dinosaur family or superordinate class (such as meat-eating) from an attribute or a dinosaur, and vice versa. Hence, they included all of the following:

1. Inferences from a superordinate class to an attribute (e.g., "He's pretty dangerous." [*Why?*] "'Cause he's a meat-eater").

2. Inferences from an attribute to a superordinate class (e.g., "I figure it's a plant-eater 'cause it needs weapons to protect from a meat-eater").

3. Inferences from a superordinate class to a dinosaur (e.g., "It's not Triceratops." [*Why?*] "'Cause they're new dinosaurs").

4. Inferences from a family to an attribute (e.g., "I guess he's probably a good swimmer." [*Why?*] "'Cause duckbills are good swimmers").

5. Inferences from an attribute to a family (e.g., "Duckbill." [*Why?*] "'Cause it has this [*beak*]").

*Dinosaur* or *animal reasoning structures* were inferences from an attribute to a dinosaur or to an animal, respectively (e.g., Dinosaur: "He's like Tyrannosaurus Rex." [*Why?*] "'Cause I saw his teeth." Animal: "A lizard." [*Why?*] "'Cause he has this long tail").

Inferences from one attribute to another attribute were called *attribute reasoning structures.* An example: "He could walk real fast." (*Why?*) "'Cause he has giant legs." The "giant legs" led directly to the "walking fast" inference, rather than proceeding via a family/superordinate, dinosaur, or animal inference.

The important difference between the first three types of inferences (family/superordinate, dinosaur, and animal) and the attribute inferences is that, for the first three types of inferences, the child engages in a type of *hierarchical reasoning,* whereas in the attribute inferences only *linear reasoning* is involved. In hierarchical reasoning, a specific feature in the picture of the dinosaur triggers the identification of a familiar animal, dinosaur, dinosaur family or superordinate class, and this identification then generates a new inference. For example, in the first quote just listed, the child said, "He's pretty dangerous. 'Cause he's a meat-eater." Some feature(s) in the dinosaur picture must have led the child to identify the dinosaur as a meat-eater, which in turn generated the inferred knowledge that this particular dinosaur must be dangerous, because meat-eaters in general are dangerous. Similarly, for the dinosaur inferences, to say that "He's

like Tyrannosaurus Rex. 'Cause I saw his teeth" implies that the "teeth" in the picture allowed the child to identify the novel dinosaur as similar to Tyrannosaurus Rex. Presumably, upon further investigation, the child might then have made other attributions based on the inference that the dinosaur was like Tyrannosaurus.

Attribute inferences, in contrast, do not involve an initial categorical or instance identification. They are *linear* in nature, rather than hierarchical. To say that a dinosaur "could walk real fast 'cause he has giant legs" implies that the explicit feature "giant legs" leads directly to the inference that the dinosaur "walks fast." That is, "walking fast" is a property or consequence of having "giant legs." This inference is not hierarchical because the child did not identify "giant legs" as an attribute of a certain class of dinosaurs, and then make the inference that because this class of dinosaurs walks fast, this particular dinosaur walks fast.

Analysis of the reasoning chain at this sentence level revealed that experts and novices made approximately the same number of inferences (expert, $M = 9.2$; novice, $M = 11.4$), $U = 12.0$, $p > .10$. Consistent with the propositional analysis, this level of analysis again showed that both experts and novices were equally capable of generating information about the new dinosaurs by using knowledge that they had. In particular, experts and novices did not differ in the number of times they reasoned analogically to animals (expert, $M = 0.8$; novice, $M = 0.6$). Thus, this finding validates the assumption that the two groups had equivalent amounts of knowledge about animals in general, and that they were equally proficient at reasoning from animal instance knowledge.

However, because the experts had more knowledge about dinosaurs, they were able to make use of this knowledge by generating a significantly greater number of dinosaur reasoning chains than the novices (expert, $M = 3.6$; novice, $M = 0.2$), Mann-Whitney $U = 3.5$, $p < .05$, one tailed. In fact, this was their largest category, accounting for 39.2% of the experts' inferences. Experts also made a larger number of superordinate/family inferences than the novices, although the difference was not significant (expert, $M = 2.2$, novice, $M = 1.0$). It is suspected that a significant difference would have been reached had more novel dinosaurs that were prototypical been used. At any rate, if both the dinosaur inferences and the family/superordinate inferences are combined, the expert children made a significantly greater number of these domain-related hierarchical type inferences than the novices (expert, $M = 5.8$; novice, $M = 1.2$), $U = 2.5$, $p < .05$, one tailed. In contrast, novices made many more attribute inferences than experts (expert, $M = 2.6$; novice, $M = 9.6$), $U = 4.0$, $p < .05$, one

tailed. The attribute inference category was clearly the largest for the novices, accounting for 78.6% of their inferences.

Thus, the main finding from this analysis is that experts' reasoning structures are derived from categorizing the novel dinosaur into a domain-related type or a class, whereas novices' inferences are not constrained by such instance or family knowledge. What this finding means is that expert children, in examining a picture of a novel dinosaur, seek critical features that can enable them to identify it as similar to a particular dinosaur instance or member of a family. Additional inferences are then generated on the basis of this initial classification. Because novice children lack the domain related instance or family knowledge, they cannot seek specific features that can lead to classification of the dinosaur, which forces the novices to generate inferences on the basis of individual attributes. Their inferences are necessarily more limited in scope, because, for example, there are only a few things one can say about a dinosaur having "giant legs." No class of dinosaurs has giant legs, which could lead to a large number of inferences. Thus, the use of attribute reasoning necessarily limits the potential inferences novices can generate about a novel dinosaur.

It seems that young children often reason in a naive way because they lack the relevant domain knowledge. Carey (1985) found that 4- to 7-year-olds did not use hierarchical reasoning and deductive inferences to attribute properties to animals such as flies and worms; this is in part because 4- to 7-year-olds have not acquired an independent domain of intuitive biology. But the direct evidence of our present study shows that 4- to 7-year-olds can reason deductively for domains (such as dinosaurs) in which they have acquired an independent and coherent theory. These young experts reason much like the way 10-year-olds and adults reasoned in the Carey study. Thus, our evidence directly confirms Carey's assertion that: "It is not that young children are incapable of deductive inferences from category membership; it is just that their knowledge of biological properties does not make it clear to them that such inferences are warranted" (1985, p. 193).

*Causal statements.* Another way to analyze children's reasoning in the free generation task is to look at their causal explanations. Thus causal statements in our study were classified into all statements that were spontaneous (i.e., not prompted by an experimenter) and that were connected by the words *because, 'cause, so,* and *since.* A typical causal statement identified a dinosaur feature or placed the dinosaur into a category in the premise, then followed with an explanation for the identification, such as "It's a meat-eater because it has sharp teeth."

In general, there was no difference between the number of causal statements generated by the expert and novice children (expert, $M = 2.8$; novice, $M = 6.6$), not significant by a Mann-Whitney test. (The higher mean for the novices is inflated by one subject's data.) Causal statements were classified into three categories: (a) domain-related premise and domain related explanation (DR-DR: e.g., "It's a meat-eater because it has sharp teeth"); (b) domain-related premise and explanation drawn from world knowledge (DR-WK: e.g., "It's a meat-eater because it has 14 toes"); and (c) World knowledge premise and world knowledge explanation (WK-WK: e.g., "It lives in the desert because it looks like a lizard").

No difference appeared between experts and novices in the number of WK-WK causal statements (expert, $M = 0.2$; novice, $M = 3.0$), not significant by a Mann-Whitney test. For those causal statements in which the children identified a domain-related feature or classification in the premise (categories a and b), there was a difference in the kinds of explanations novices and experts gave. Experts gave explanations based predominantly on domain-related knowledge (category a), whereas novices gave explanations based predominantly on world knowledge (category b). (See Figure 7.) The interaction between the kind of explanation and knowledge was significant (expert, *Ms*: DR-DR = 2.2, DR-WK = 1.0; novice, *Ms*: DR-DR = 0.4, DR-WK = 2.6), $F(1, 8) = 7.71$, $p < .05$. Hence, even though both novices and experts are equally facile at producing causal inferences, the explanations given by novices are less appropriate for justifying the identified feature.

*Accuracy within dinosaur taxonomy.* Most of the analyses have focused on the extent to which experts' inductions are hierarchical in nature and are based on domain-related knowledge. Although experts, by definition, have greater domain knowledge, it is not necessarily clear that they use it in a way that leads to inferences that are more accurate than novices' inferences with respect to an absolute scientific standard. Consequently, the accuracy of children's generated knowledge in the free generation protocols was examined. For example, a statement such as "Usually dinosaurs live on land if they have small heads like that" is clearly incorrect because some dinosaurs with small heads live in the water. A statement such as "[It eats] plants because it's low down and plants are low down" is not a scientifically acceptable reason for classifying a dinosaur as plant-eating. Similarly, the statement, "He'd be able to eat animals, might eat plants too" is not scientifically acceptable because no dinosaur eats both animals and plants.
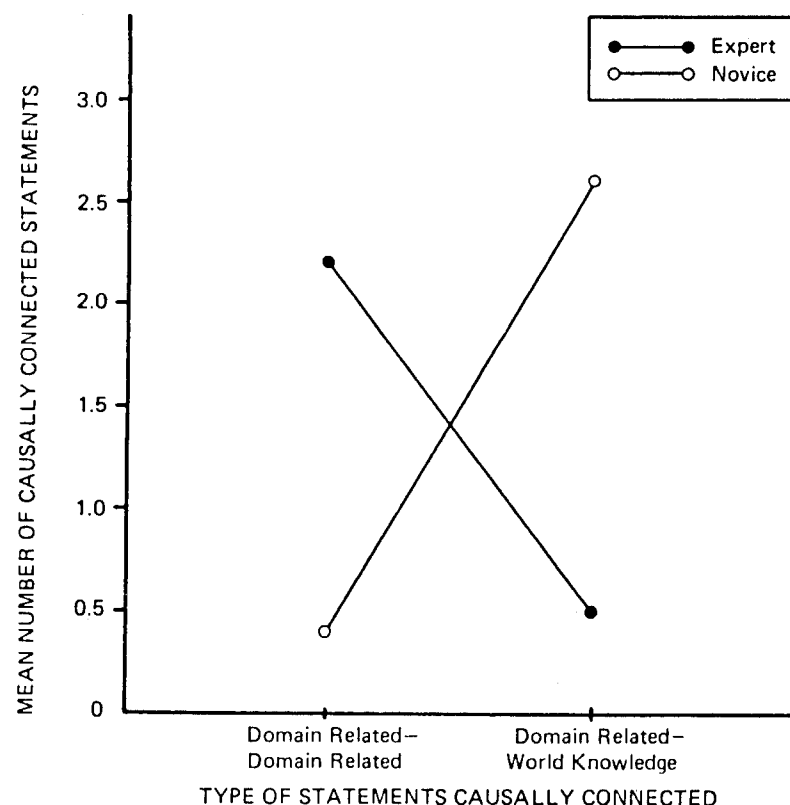
**Figure 7.** Types of causally connected statements produced by expert and novice children.

The experts did not produce any incorrect statements in the free generation task. Novices produced an average of 1.8 incorrect statements, Mann-Whitney $U = 5.00$, $p < .05$. (A similar analysis of accuracy for the probe question task revealed the same pattern of results.) The accuracy analysis indicates that experts have a stable body of domain knowledge that allows them to generate correct inferences about novel stimuli that are consistent with what they know. Novices, even though they actively attempt to generate inferences and generalizations, produce incorrect inferences because they lack the relevant domain knowledge to constrain them.

## Results from the Sorting Tasks

*Sorting of familiar dinosaurs.* From the results of the first study, it was speculated that children, as they acquire expertise, will structure their dinosaur knowledge hierarchically. Individual dinosaurs are grouped into families, and the families are then coalesced into the higher categories of meat- and plant-eaters, a differentiation zoologists also use to classify mammals (Storm, 1980). This tentative conclusion was based on the evidence provided by a single child. Across the second and third studies, data were accumulated on 20 additional children's sorting of the 20 better-known dinosaurs. Ten of the children were novices and 10 were experts, and they ranged in age from 4 to 7 years.

To gain a rough idea of what hierarchical level (if any) children's sorting patterns fit into, each child's sorting into superordinate, emerging superordinate, family, emerging family, or no classification, categories was first classified. To be classified as *superordinate,* a child's groups had to satisfy the following three criteria: (a) half or more of the members of each group must be from the same superordinate category; (b) half or more of the members of each superordinate category must be grouped together; (c) members of each superordinate category must form a majority (more than half) of at least one group of two or more members. The same criteria were used for classification as family, substituting "family" for "superordinate." If a child's sorting satisfied 2 out of the 3 criteria for superordinate and consisted of 2 or 3 groups, it was classified as *emerging superordinate.* Similarly, if a child's sorting satisfied 2 out of the 3 criteria for family and consisted of 4 or more groups, it was classified as *emerging family.* Sortings that did not conform to these criteria were designated as *no classification,* and were not considered to fit a hierarchical scheme.

Based on these criteria, 3 of the 10 experts classified the 20 dinosaurs into the superordinate meat-eating and plant-eating categories; this replicates the sorting result found for the 4½-year-old expert in the first study. None of the novices spontaneously grouped the dinosaurs on the abstract dimension of diet, and only one novice had an emerging superordinate grouping. When asked to subdivide their superordinate groupings, the three experts created family or emerging family groupings. It is known, then, that at least some experts can group dinosaurs into both superordinate diet categories and families, forming a hierarchical structure. Novices' groupings do not show evidence of a hierarchical representation.

Four experts initially grouped the dinosaurs by families and an additional one had an emerging family grouping. Only one novice's grouping could be classified as family, but four novices showed emerging family distinctions. It is not surprising that novices have an emerging ability to sort dinosaurs on the basis of family membership because family membership is determined to a large extent by visual

similarity. Dinosaurs belonging to the same family tend to have similar visual features, such as duck-like bills for the duckbill family, and notably distinct horns for the horned dinosaurs. Wattenmaker, Nakamura, and Medin (1988) also believe that one can use apparent family resemblance to sort objects into categories. To sort strictly by family or superordinate category as the experts did, however, may require the use of deeper principles.

If the results are combined, it is found that 7 experts sorted strictly by family or superordinate category, whereas only 1 novice did so; in contrast, 5 novices showed emerging family or emerging superordinate distinctions, whereas only 1 expert did so, a significant difference by a Fisher's Exact test, $p < .05$. This result suggests that novices formed groups on-line, as the task of sorting demanded of them. In contrast, more of the expert children appeared to sort on the basis of some predefined representation of family or superordinate structure.

A more detailed quantitative analysis, performed only on the data from the 5 novices and 5 experts from the third study, compared the number of "perfect" groups formed. The analysis was performed on the groupings children constructed after they were asked to break down any initial large groups. A *perfect group* was defined as one that includes all of the dinosaurs from one family and no dinosaurs from other families. Experts produced significantly more perfect groups ($M = 2.2$) than novices ($M = 0.6$), Mann-Whitney $U = 3.00$, $p < .05$, one tailed. Conversely, the novices had three times as many groups with intrusions (groups with a mixture of dinosaurs from different families) as experts ($Ms = 0.8$ mixed groups for experts and 2.4 mixed groups for novices), $U = 4.00$, $p < .05$, one tailed. This finding implies that the expert children are less likely to make "errors" in their sorting, which would presumably occur if one were relying strictly on visual similarity as a basis for the groups. This result was interpreted as further evidence that family structures that already existed in the experts' knowledge base guided their sorting, whereas the novices had to rely on properties they observed on-line to come up with family-like groupings.

To capture the differences in the principles which guide the sortings of expert and novice children, the explanations they gave for their groupings were also examined. Even though both experts and novices appeared to sort on the basis of family resemblance, the novices relied more on the visible properties in the pictures of dinosaurs. They explained by saying: "They all look alike. Their heads look alike," or "They have small hands . . . And their skin is rough," or "They have the same heads . . . And they have horns." By contrast, expert children explained their groupings on the basis of more ab-

stract dimensions that were implicit rather than explicitly visible in the picture: "This group . . . is protected and harmful dinosaurs that are plant-eaters," or "These dinosaurs are plant-eaters. They aren't harmful," or "These are the only meat-eaters I could find." This particular expert child's protocol not only shows that the child attended to more implicit, abstract features, but it also shows some hierarchical embedding. That is, plant-eaters can be harmful or not harmful.

The difference between experts' and novices' explanations for their sortings were consistent across both the second and third studies. In the second study, novices gave explicit physical features 83.2% of the time, whereas experts did so only 42.6% of the time, $F(1, 8) = 7.176$, $p < .05$. Experts mentioned implicit features such as diet, defense, and habitat more often than novices (expert, $M = 77.6\%$; novice, $M = 16.8\%$), $F(1, 8) = 21.605$, $p < .01$. In the third study, 40% of the expert children's reasons concerned implicit features as compared to 13% of the novices' reasons, $F(1, 8) = 3.495$, $.05 < p < .10$.

Thus, the analyses of the explanations further support the interpretation that the novices sort on-line, basing their groupings on the visible explicit features, in contrast to the experts' more principled sorting based on implicit features. In conclusion, the analyses of the sorting of the familiar dinosaur data suggest the following: It is possible for novices to sort objects on the basis of family resemblance, without knowing or understanding the underlying principle for the categorization; this reflects an ability to perceive similarity in the visual stimuli, and is a critical attribute for humans to possess in order to learn categories for objects. Expert children, however, have additional knowledge beyond the explicit visual features of the dinosaurs. They sort on the basis of their internal representation of the family structures. In addition, a few expert children sort according to diet, an even higher level of classification that coalesces family structures, in the same way that zoologists classify animals.

*Forced choice sort.* The effect of having a coherent representation of family structures can also be detected by examining the way children sort unfamiliar dinosaurs: The children had the choice of either inserting the unfamiliar ones into the existing categories created by their initial familiar dinosaur sortings (redisplayed for them), creating new categories, or partially dissolving the existing categories and forming revised ones. The end result was that both groups of children formed a total of 9 to 10 groups. Although there were no differences between the expert and novice children in the number of groups formed, there were some differences in how accurately they sorted the unfamiliar dinosaurs, especially the prototypical ones. One

measure of accuracy is the degree of overlap among the features of each group's members. A rough measure of the overlap of features of the members of a group is to measure the degree to which members match on the following four attributes: body, head/teeth, legs, and miscellaneous features (armor, horn, neck, tail, crest, etc.). Each of the four attributes of each dinosaur was given a family classification. For example, one of the conflict dinosaurs had the body of a giant plant-eater, had the head/teeth of a horned dinosaur, had meat-eater legs, and had miscellaneous features like a giant plant-eater. To determine the degree of overlap for a child's grouping, each attribute for each dinosaur in a group was given a score representing the percentage of dinosaurs in the group having that particular family classification for that attribute. Then all the scores for a group (four for each dinosaur) were averaged to yield a group mean. Finally, the group means were averaged for each child to represent an overall percentage of overlap. The five newly-discovered dinosaurs were excluded from the analysis because their features do not fit into any of the common family types. Groups with only one member were also excluded.

For the expert children, the percentage of overlap was 78.5%, whereas for the novice children, the overlap was 61.4% (Mann-Whitney $U = 2.0$, $p < .05$). By using this more sensitive analysis, it was shown once again that even though category membership can be determined to a large extent by the visual similarity among the dinosaurs, the expert children were still superior at detecting such similarity. Experts were probably better at such detection because they imposed their schemata of family structures to help them encode the relevant features. That is, they based their groupings of the dinosaurs on features that were predetermined by their family schemata. The novice children had to rely more on "bottom-up" processing of the physical features to make decisions about category membership, and as a result may have missed or underemphasized some of the relevant features.

As anticipated, experts were especially good (compared to novices) at classifying the unfamiliar prototypical dinosaurs. Experts placed an average of 36% of the new prototypical dinosaurs with a group which included all the familiar dinosaurs from the same family as the prototypical dinosaur, without any intrusions of familiar dinosaurs from other families. (Sometimes children placed some other unfamiliar dinosaurs with the group as well.) Novices, in contrast, did not insert any of the new prototypical dinosaurs in such groups. The difference was significant by a Mann-Whitney test, $U = 0$, $p < .01$. Thus, consistent with the theoretical assumptions, experts were able to use their knowledge of families to infer the correct family grouping for unfamiliar dinosaurs.

Experts also clearly distinguished the newly-discovered dinosaurs from the prototypical dinosaurs. They classified 80% of the newly-discovered dinosaurs into new groupings which were not initially formed during the sorting of the familiar dinosaurs. These new groups did not include any of the original set of 20 familiar dinosaurs. In contrast, they classified only 20% of the prototypical dinosaurs into such new groupings. The difference in placement of newly-discovered and prototypical dinosaurs was significant by a Wilcoxon's test, $p < .05$. Novices, on the other hand, did not differ in their placement of the newly-discovered and prototypical dinosaurs into new groupings (newly-discovered: 80%, prototypical: 52%).

In sum, experts were able to utilize their family schemata to sort dinosaurs they had never seen before. They not only were better at detecting similarities between the dinosaurs, but they placed prototypical dinosaurs into the correct existing family groupings, and formed new groups for the newly-discovered ones. Such discrimination between the prototypical and newly-discovered dinosaurs was not evident in novices' sorting; this suggests that the experts would be more likely to use categorical reasoning to constrain their inferences for the prototypical than for the newly-discovered dinosaurs.

### Analyses of Protocols Across All Tasks

Analyses of the protocols generated by the children across all the tasks will now be reported. The remarks and explanations given in the free generation, probe, sorting, and the oddity tasks are combined in various different ways for the following analyses. The purpose of these analyses is to provide additional converging support for experts' knowledge of dinosaur families and superordinate diet classes.

*Inferences based on families.* Of the 20 novel dinosaurs, 5 prototypical and 7 nonprototypical ones had features that were all consistent with either the meat-eating or the plant-eating categories. Across the free generation and probe tasks, experts were more correct than novices in classifying the 5 prototypical dinosaurs according to diet (expert, $M = 95\%$ correct; novice, $M = 44.6\%$ correct), $U = 1.0$, $p < .05$. When determining the diet class of the 7 nonprototypical dinosaurs, experts and novices were equally correct (expert, $M = 70.0\%$, novice, $M = 70.6\%$). That experts had an advantage for prototypical dinosaurs but not for nonprototypical dinosaurs suggests that experts cannot rely strictly on individual attributes (such as "big head") to make diet classifications, because individual attributes can be used equally effectively for prototypical and nonprototypical dinosaurs. Rather, experts may be able to draw upon their knowledge of families to categorize the prototypical dinosaurs and thereby identify

their diets. This result is consistent with the previous analysis showing that experts sort prototypical dinosaurs better than novices. The finding is also consistent with our previous conclusion that expert children use reasoning structures that are hierarchical, whereas novices use linear reasoning, in which inferences are based on individual attributes.

Additional evidence that the experts in this third study used family-level knowledge at least part of the time to infer the dinosaurs' diet can be gleaned from the reasons they gave for their diet classifications in the free generation and probe tasks. The expert children used membership in a family to justify diet classifications 14.4% of the time, whereas novices never did (Mann-Whitney $U = 2.5$, $p < .05$). For example, an expert might say, "He eats plants because he looks like a duckbill." In the Carey (1985) study, only about 6.6% of the 4- to 7-year-olds appealed to inclusion in a subclass of animals to justify an attribute (e.g., "Dodos breathe because birds breathe," p. 85). The difference between our experts and the Carey subjects in the frequency of use of the intermediate level of the classification hierarchy may result from the fact that our experts were more knowledgeable about dinosaurs than the average child is about animals. When the data from our novices and experts are averaged, the resulting percentage of classification on the basis of dinosaur families (7.2%) is comparable to the Carey children's classification on the basis of a subclass of animals (6.6%).

The total number of family-related statements that children produced (e.g., "He's a horned dinosaur") was also examined. Tabulating statements generated across all the tasks of this study, the expert children produced an average of 13.6 family-related statements, whereas the novices produced an average of only 1.0 (Mann-Whitney $U = 5.0$, $p < .05$). Again, this result suggests that experts more often make inferences about a novel dinosaur on the basis of its relation to a certain family.

*Presence of superordinate diet classes.* As in the second study, it was found that experts had a sense of the contrastive nature of the two superordinate diet classes, meat-eating and plant-eating, whereas novices did not. Across all the tasks, all five experts used the presence of a feature as justification for placement of a dinosaur in one diet class, and the absence of the feature as justification for placement of the dinosaur in the complementary diet class. Only one novice used features in this constrastive way ($p < .05$ by a Fisher Exact Probability test).

The presence of superordinate meat- and plant-eating categories in the oddity task was also detected. This task included four trials on which one dinosaur could be segregated from the other three on the

basis of diet category. Experts grouped meat- and plant-eaters in separate groups on 85% of these trials, while novices grouped meat- and plant-eaters separately on only 55% of the trials (Mann-Whitney $U = 1.5$, $p < .05$).

### Summary

In the third study of this series, the focus was on how expert and novice children infer new domain-related information; children were asked to generate attributions about novel dinosaurs. The expert and novice children were selected so that they were matched on a number of dimensions indicating that they had equivalent abilities to learn. As the result of obtaining a matched sample, no differences were detected between the experts and the novices on a number of quantitative measures, such as the total number of propositions produced, the number of groups created when sorting dinosaurs, and the number of causal statements made. Nor were any differences detected in the use of general learning strategies when domain knowledge was not involved, such as the use of comparisons to animals, and the use of relevant features of world knowledge to constrain inferences.

However, experts showed a great advantage over novices in using domain knowledge to generate inferences. In particular, the expert children used domain-related knowledge to generate comparisons whereas novice children used world knowledge. Experts based their explanations of causality on relevant domain features to a greater extent than novices. The expert children knew what explicit features to attend to in the pictorial representations, and were able to use categorical reasoning to induce attributes about novel dinosaurs. Expert children based their sorting on well-defined schemata of family types, whereas the novice children tended to sort on-line, on the basis of visual similarity. The evidence for experts' use of existing family schemata to guide sorting includes (a) formation of "perfect" groupings without intrusions, (b) greater overlap in features among members of a group, and (c) placement of novel prototypical dinosaurs into existing groups from the same family, and newly-discovered dinosaurs into new groups. Although no longitudinal learning data is available, our results imply that such domain-constrained inferences would allow expert children to learn about new domain concepts more quickly and accurately. Thus, the third study in this series has shown how background knowledge per se can enable the expert children to learn new domain-related concepts more readily, despite the fact that both expert and novice children have the same fundamental learning skills.

# GENERAL DISCUSSION

In a series of three studies, a small, simple, and confined domain —dinosaur concepts—was examined to assess how it may be structured in children's representations, and how such knowledge may be used. Even for such a simple domain, in which the knowledge is basically organized in a classification scheme, significant differences were found in the way experts' and novices' representations of such knowledge are structured. Across the three studies, evidence suggests that the structure of the expert children's knowledge is more coherent, both locally and globally (hierarchically). Local coherence was assessed in a number of ways: by the sequential generation of dinosaur names and the identification of dinosaur features, by the use of connecting words in the production of discourse about a familiar dinosaur, and by the frequency with which the discussion of one attribute (such as defense) leads to the discussion of another attribute. Experts' family structures were better-defined than those of novices, as could be seen by the number of links between concepts within the same family and the sharing of features among family members, by the amount of overlap of features of dinosaurs grouped together, and by the extent to which groups conformed to perfect family structures. Global coherence or hierarchy was assessed by the presence of diet-level classification schemes as well as family-level schemes. Some experts were able to sort dinosaurs at both the superordinate and family levels, whereas no novices did this. Expert children also were able to use a single feature to contrast superordinate meat-eating and plant-eating classes.

In a matched sample of children, novice children were shown to be as competent as expert children in using general learning skills, such as making comparisons and giving causal explanations, when these skills were assessed in a domain in which both the expert and novice children had equivalent knowledge (such as the animal domain). The use of learning skills seemed to differentiate the experts from the novices only when they interact with dinosaur knowledge. For instance, using a single feature to contrast two diet classes can only be accomplished if one has a representation in which the features shared by members of the same class have been generalized. Classes share few features. Hierarchical reasoning can be manifested only if children have the categories or family types represented in memory, otherwise they would have to resort to linear reasoning. Likewise, sorting based on family types can be manifested only if these family types exist in memory, otherwise sorting has to be based on visual similarities that can be detected on-line. Finally, inferences

derived from comparisons to dinosaurs and dinosaur families can be made only if dinosaur knowledge is available and hierarchically organized in memory so that it can be accessed.

In sum, sufficient evidence has been marshalled from the present set of three studies to suggest that children can act more or less intelligently on the basis of what knowledge they have. Even for such a small and confined domain of knowledge, a modest mastery of the domain is manifested in intelligent explanations, constrained inferences, categorical reasoning, and hierarchical classification based on well-defined family structures and superordinate diet categories. Yet novices' explanations seem inconsequential, and their inferences are not constrained and therefore appear incorrect and inappropriate. Their linear reasoning is limited in scope, their sorting is errorful in that it includes a number of intrusions, and they have no awareness that newly-discovered dinosaurs should be discriminated from prototypical ones.

Because the subject samples were carefully controlled so that they differed mainly in the amount of domain knowledge, a number of quantitative measures did not exhibit differences. Hence, the differences that are present reflect the presence and structure of domain knowledge. The conclusion is that one reason that children generally display global inadequacy across a number of domains is that they lack the relevant knowledge in a number of domains. By selecting a domain that some children know something about, qualitatively superior abilities that can be attributed only to domain-specific knowledge and the way that it is organized have been demonstrated.

## REFERENCES

BOUSFIELD, W. A. (1953). The occurrences of clustering in the recall of randomly arranged associates. *Journal of General Psychology, 49,* 229–240.

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

CHI, M. T. H. (1988). Children's lack of access and knowledge reorganization: An example from the concept of animism. In F. Weinert & M. Perlmutter (Eds.), *Memory development: Universal changes and individual differences* (pp. 169–194). Hillsdale, NJ: Erlbaum.

CHI, M. T. H., & CECI, S. J. (1987). Content knowledge: Its representation and restructuring in memory development. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 20, pp. 91–142). New York: Academic Press.

CHI, M. T. H., & KOESKE, R. (1983). Network representation of a child's dinosaur knowledge. *Developmental Psychology, 19,* 29–39.

GELMAN, R., & BAILLARGEON, R. (1983). A review of some Piagetian concepts. In J. H. Flavell & E. M. Markman (Eds.), *Handbook of child psychology: Vol. 3. Cognitive development* (4th ed., pp. 167–230). New York: Wiley.

GOBBO, C., & CHI, M. T. H. (1986). How knowledge is structured and used by expert and novice children. *Cognitive Development, 1,* 221–237.

KINTSCH, W., & van DIJK, T. (1978). Toward a model of text comprehension and production. *Psychological Review, 85,* 363–394.

STORM, C. (1980). The semantic structure of animal terms: A developmental study. *International Journal of Behavioral Development, 3,* 381–407.

WATTENMAKER, W. D., NAKAMURA, G. V., & MEDIN, D. C. (1988). Relationships between similarity-based and explanation-based categorization. In D. Hilton (Ed.), *Science and natural explanation: Commonsense conceptions of causality.* London: Harvester Press.