

## **MORE EVIDENCE FOR A RELATIONAL SHIFT IN THE DEVELOPMENT OF ANALOGY: CHILDREN'S PERFORMANCE ON A CAUSAL-MAPPING TASK**

**Mary Jo Rattermann**  
Swarthmore College

**Dedre Gentner**  
Northwestern University

Gentner (1988) has proposed a relational shift whereby children interpret analogy and metaphor first in terms of object similarity and then in terms of relational similarity. Goswami (1996) argues against the relational shift hypothesis, citing as evidence a study performed by Goswami and Brown (1989) in which 3-, 4-, and 6-year-old children were able to correctly complete pictorial A:B::C:? analogies based on familiar causal relations, and, contrary to the predictions of the relational shift hypothesis, made very few object-similarity errors despite the presence of an object-similarity choice. In the present experiment we obtained similarity ratings of Goswami and Brown's stimuli and found that the materials did not contain a true object similarity choice and therefore that study was not an adequate test of the rela-

---

This research was supported by an award from the National Science Foundation (SBR-95-11757) to Gentner, a grant from the National Institute of Education under Contract 400-31-0031 to the Center for the Study of Reading at the University of Illinois at Urbana-Champaign and by Training Grant HD 07205 from the National Institute of Mental Health to the Psychology Department of the University of Illinois. The studies were carried out at the Beckman Institute at the University of Illinois. Preparation of the manuscript was supported in part by a fellowship to Gentner from the Max Planck Institute for Psycholinguistic Research in Nijmegen, the Netherlands.

We thank Usha Goswami and Ann Brown for helpful discussions of these issues and for generously allowing us access to the stimulus materials and the individual data for Experiment 2 of the Goswami & Brown (1989) paper. Special thanks to Rebecca Campbell for her valuable contributions to the design and implementation of these experiments. We also thank Graeme Halford and an anonymous reviewer for their helpful comments on this manuscript.

---

Direct all correspondence to: Mary Jo Rattermann, Whitely Psychology Laboratories, Franklin & Marshall College, P.O. Box 3003, Lancaster, PA 17604-3003.

---

tional shift hypothesis. After appropriate modifications to their methodology, we found that 4- and 5-year-old children's performance was consistent with the relational shift hypothesis: First, object-similarity errors were highly frequent initially and decreased with age; second, the rate of relational (correct) responding increased with age; and third, performance on the analogues was positively related to children's knowledge about the participating causal relations. We conclude by proposing an explanation for the relational shift based on an alignment view of similarity comparison and, further, suggest a new role for object similarity in children's analogical development.

There is little doubt that children can, and do, reason by analogy. Countless examples, from both empirical studies and anecdotal stories, demonstrate the ability of very young children to reason about the unfamiliar based on the familiar. The question addressed in this paper is not whether children can reason analogically, but how analogical reasoning develops. Different theoretical viewpoints have emphasized different aspects of this development, focusing, for example, on very young children's use of analogy as a primary tool for acquiring knowledge (Brown, 1990; Goswami & Brown, 1989); the role of cognitive capacity in the use of relational similarity (Halford, 1993); or the importance of knowledge accretion and representational change in analogical reasoning (Gentner & Rattermann, 1991; Gentner, Rattermann, Markman, & Kotovsky, 1995; Kotovsky & Gentner, 1996; Markman & Gentner, 1993ab).

Studies using a variety of methods have shown that young children can respond to relational similarity (e.g., Brown, Kane, & Echols, 1986; Gentner, 1977a, 1977b; Holyoak, Junn, & Billman, 1984), suggesting that analogical reasoning is a skill possessed at a very early age. For example, Gentner (1977a, 1977b) has demonstrated that young children can perform a spatial analogy between a human body—which is a highly familiar domain, even for preschoolers—and pictured objects, such as trees and mountains. She showed children simple pictures, such as a picture of a tree, and asked, "If a tree had a knee, where would it be?" Even 4-year-olds (as well as 6-year-olds and 8-year-olds) were able to perform the mapping of the human body to the tree. They were as accurate as adults, even when the orientation of the tree was changed or when confusing surface attributes were added to the pictures.

Yet there is also evidence that young children are limited in their analogical reasoning abilities. For example, Sternberg and his colleagues presented grade-school children and adults with  $A:B::C:D$  analogies (Sternberg & Downing, 1982; Sternberg & Nigro, 1980). They found that young children (grades 3 and 6) responded based on thematic, rather than relational, similarity while ninth-grade and college students produced relational responses. A similar shift from thematic to relational was also found by Piaget, Montangero, and Billeter (1977). Further evidence of limitations on children's abilities is presented by Winner, Rosenstiel,

and Gardner (1976), who found that young children interpreted comparisons like "The prison guard's heart was hard as stone" to mean that the person was literally turned to stone. The relational interpretation—that like a stone the guard's heart was impossible to alter or break, etc.—was made only by much older children and adults. While few researchers would agree with Piaget's assessment that before the advent of formal operations (at about 11 years) children do not possess the basic cognitive competence required to perform a relational mapping (Inhelder & Piaget, 1958; Piaget, Montangero, & Billeter, 1977), there remains the challenge of finding a unified account of both children's abilities and their limitations.

**The Career of Similarity.** Gentner (1988) proposed that the development of similarity and analogy shows a relational shift: a shift from early attention to common object properties to later attention to common relational structure. Gentner and Rattermann (1991) amplified this framework to encompass the career of similarity. They proposed that the development of similarity proceeds from overall similarity to object similarity—for example, a *round, red ball* and a *round, red apple*—to similarity based on common relations between objects—for example, "*apple ON table*" and "*book ON bureau*"—to similarity based on common higher-order relations (relations between relations)—for example, "*sun MELTS snow CAUSES snow CHANGE TO puddle*" and "*flame MELTS candle CAUSES candle CHANGE TO mound*."<sup>1</sup> They further argued, based on a survey of developmental literature, that the mechanism underlying this evolution is epistemological, not maturational. Thus, there is no absolute level of maturity or experience that will enable children to process relational similarity; rather, the kind of similarity children can perceive is determined by the nature of their domain representation, and, in particular, by the amount and kind of relational knowledge they possess in the domain (Gentner, Rattermann, Markman, & Kotovsky, 1995; Kotovsky & Gentner, 1996; Rattermann, Gentner, & DeLoache, 1989; Rattermann & Gentner, in preparation). The claim that changes in knowledge shape children's analogical development offers a potential resolution to the question of why analogical ability emerges at very different times in different domains (Gentner & Rattermann, 1991).

There is considerable evidence for such a pattern of performance in children's use of similarity and analogy (see Gentner & Rattermann, 1991, for a full discussion). Support for the relational shift portion of this hypothesis is provided by Gentner (1988), who asked children and adults to interpret metaphors such as "A cloud is like a sponge." The younger children (5- to 7-year-olds) produced *mere-appearance* interpretations: interpretations based on shared object descriptions,

<sup>1</sup>The term 'higher-order relation' has been used in two conflicting ways in the developmental literature. For Piaget, higher-order relations referred to the relation of identity between two lower-order relations, such as that which applies in classic A:B::C:D analogies. To avoid confusion, it is worth noting that our use of the term "higher-order relation" encompasses not only this identity relation but also other relations between relations.

such as "Both are round and fluffy." Older children and adults gave interpretations based on common relational structure: for example, "Both can hold water and then later give it back." Chen and Daehler (1989) found that children easily became confused when given problem sets in which surface similarity did not match with structural similarity, suggesting that they relied on surface similarity in their mapping. Rattermann, Gentner, and DeLoache (1989; in preparation) presented 3- and 4-year-old children with an analogical mapping task in which the correct answer was based on relational similarity, but there was a competing (incorrect) object-similarity response. Both age groups, but particularly 3-year-olds, had difficulty learning the relational mapping and made many object-based errors, particularly when the objects were perceptually rich (so that object similarity was salient). These results bear out three predictions of the relational shift hypothesis, which proposes that when relational similarity competes with object similarity: (1) young children (and adults) will initially make many object-similarity responses; and with age/experience, (2) object-similarity responding should decrease, and (3) relational responding should increase. A further prediction, because the relational shift is not age-determined but knowledge-related, is that it can occur at different ages in different domains, depending on domain knowledge (Gentner & Rattermann, 1991).

In her relational primacy view, Goswami (1992, 1996) proposes an alternative to the relational shift hypothesis. In common with the view proposed by Gentner and Rattermann (1991), Goswami (1996) states that "analogical reasoning (relational or structural mapping) can be used to solve problems and to acquire new knowledge as soon as the infant or young child has represented the relevant relational knowledge" (p. 98).<sup>2</sup> A fundamental difference between the two positions, however, is in the proposed role of object similarity. According to the relational shift hypothesis, the use of object similarity is a crucial and necessary step in the progression from comparisons based on overall similarity to comparisons based on relational similarity. In contrast, the relational primacy view proposes no special role for object similarity in early comparison processing; object similarity is merely one of several performance factors that can impede children's ability to perform relationally. Because the only role of object similarity is that of a "performance factor" no relational shift from object similarity to relational similarity is predicted. In fact, based on this proposal, Goswami has stated that the relational shift is "more apparent than real" (Goswami, 1996, p. 110). In sum, while both the relational shift and the relational primacy views propose that the ability to make

---

<sup>2</sup>Note that while both Gentner and Ratterman (1991) and Goswami (1996) consider analogical reasoning to be the mapping of relational similarity, when discussing the analogical abilities of children Goswami often describes the "relation of similarity;" that is, similarity of identity between lower-order predicates, such as "same color," "same shape," or "identical clown." We prefer to use the more specific term *literal similarity* when describing situations in which both objects and relations are identical.

relational comparisons appears early and increases with the accretion of relational knowledge within a domain, they differ in the proposed role of object similarity.

Thus, although both the relational primacy view and the relational shift hypothesis assert that knowledge of domain relations is a crucial determinant of analogical performance, the two views differ in the developmental course they predict. The relational shift hypothesis predicts a shift from interpretations based on object commonalities to interpretations based on relational commonalities as domain experience increases. The relational primacy view predicts that children should either process relational comparisons correctly from the beginning or, if errors are made before the child possesses adequate domain knowledge, there should be a variety of errors, with no one type predominating. Further, once children understand the relevant relations, no further object similarity matches should occur. Thus Goswami argues against the relational shift in analogical development.

### Object Similarity and Relational Primacy

The evidence cited previously suggests a shift in children's analogical processing from object commonalities to relational commonalities (for a complete review see Gentner & Rattermann, 1991). To account for these findings, Goswami (1992, 1996) proposes that object similarity is merely a "performance factor" that impedes children's ability to perform a relational mapping. To support this view of object similarity, Goswami cites an experiment performed in collaboration with Brown (Goswami & Brown, 1989). In this research, they tested the relational shift hypothesis against the relational primacy claim that from the beginning (or at least as soon as any domain knowledge is present) children can overlook object similarity and respond based on relational similarity. To do this they presented young children (2-, 3-, and 4-year-olds) with 4-term (A:B::C:??) causal analogies in which relational similarity and object similarity were in conflict. The relations that formed the basis for the analogies were familiar causal relations, such as burning and melting. For example, an analogy based on the cut relation was

loaf of bread:single slice of bread::lemon:?? (*slice of lemon*)

The child was to select the correct completion (*slice of lemon*) from a set of five alternatives:

- D - correct (the correct object with the correct causal transformation)
- E - the wrong object with the correct transformation
- F - the correct object with the wrong transformation
- G - a high-similar object match to the C term ("mere-appearance")
- H - an identity match to the C term

Goswami and Brown included the *wrong object-correct transformation* and *correct object-wrong transformation* response alternatives in order to test whether children had partial knowledge of the analogical rule. Finally, to test the relational shift hypothesis, the "*mere-appearance*" choice was included. If, as this hypothesis predicts, children rely on object similarity before they understand the relational structure, then the high-similarity object match to C (a "*mere-appearance match*" in the terms used by Gentner (1983, 1988)) (a) should be a frequent error among young children, and (b) should decrease with age and experience as the rate of correct relational responses increases. In contrast, if the relational primacy view is correct, then even very young children will overlook salient object similarity choices and respond relationally to causal analogies. When young children do make errors, they should be evenly distributed among the error types; there should be no special tendency for young children to make object-similarity errors.

In addition to the analogy task, Goswami and Brown (1989) conducted a causal reasoning control task to test children's understanding of the domain relations. For instance, for the "cut" relation the children were shown a lemon and a knife cutting the lemon, and were asked to select from among several alternatives the final picture to complete this sequence: specifically, a slice of lemon.

Goswami and Brown (1989) performed two experiments using this basic paradigm, and the results of these experiments supported their major prediction: both age groups performed significantly above chance in both the analogy and causal reasoning tasks. Thus Goswami and Brown's first conclusion from this research was that children as young as 3 years old were able to resist the tempting object similarity and respond based on a common causal relation. Of course, this conclusion is not problematic for the relational shift hypothesis, which makes no predictions regarding the age of relational responding in any given domain; rather, its predictions concern the developmental sequence of similarity types.

Goswami and Brown's (1989) second conclusion, however, is more troublesome. Based on results from their Experiment 2, they noted that there was no special tendency to make mere-appearance errors in either age group (mere-appearance matches accounted for only 10% of the 3-year-olds' responses and 2% of the 4-year-olds') and further that the 3-year-olds made several types of errors. From this they concluded that "when the 3-year-olds cannot solve the analogies they select randomly between the different answer options." This pattern of results produced what appeared to be a clear disconfirmation of the relational shift hypothesis; when young children did make errors they were distributed across a variety of alternatives, with no special tendency to make object similarity errors.

The conclusion that young children did not show a tendency to make object similarity errors does bear directly on the relational shift hypothesis, which predicts that object errors should be predominant. Note, however, that Goswami and Brown (1989) base their conclusion on the relatively low number of responses to the mere-appearance choice; a choice that they assumed provided a tempting object match to the C term. Only if this assumption holds true does the failure of



young children to choose the mere-appearance alternative count against the relational shift hypothesis. An examination of the stimuli used in their second experiment (which is the focus of this work) suggests that this was not the case.<sup>3</sup> The mere-appearance matches did not seem particularly similar to the C-term and, further, in some cases other alternatives appeared more similar to the C-term than the mere-appearance matches. If this were true, the presence of these competing high-similarity alternatives would of course diminish the rate of responding to the designated "mere-appearance" matches. Consequently, the rate of mere-appearance responding would not be a fair assessment of children's tendency to make object similarity responses.

To settle this matter, we took two further steps. In Experiment 1a we obtained similarity ratings for the stimuli used by Goswami and Brown (1989) to determine whether the alternatives presented to the children contained other high similarity matches to the C term—matches that could present viable alternatives to the mere-appearance choice. We then compared the results of this rating task to the performance of the children in Goswami and Brown's Experiment 2. In Experiment 1b we repeated Goswami and Brown's study, altering the stimuli to better achieve the desired tension between relational and object similarity.

### EXPERIMENT 1A

In order to verify whether the similarity structure in Goswami and Brown's (1989) study truly tested the relational shift hypothesis, we presented adults with the stimuli from Goswami and Brown's second experiment and asked them to judge the similarity between the C term and each of the possible choices.<sup>4</sup> We then reanalyzed the data from Goswami and Brown's Experiment 2 to determine whether the children's responses varied based on object similarity.

### METHOD

**Participants.** The participants were 30 undergraduates from the University of Illinois who were paid for their participation.

**Materials.** The materials were the pictures used in Goswami and Brown's (1989) Experiment 2.

**Procedure.** Participants were shown the C term of the analogy and asked to rate its similarity to each of our response alternatives: *correct*, *wrong object-cor-*

---

<sup>3</sup>We are very grateful to Usha Goswami for making the stimuli and the individual data from Experiment 2 available to us.

<sup>4</sup>We attempted to obtain similarity ratings from young children, however, we found that it was difficult for them to make explicit judgments of similarity; they were either nonresponsive or highly variable in their responses. Consequently, we used adult ratings for this experiment.

rect transformation (WO-CT), correct object-wrong transformation (CO-WT), and mere-appearance (MA). (The fifth response alternative, an identity match to the C picture, was assumed to be of maximum similarity and was not rated.) Participants were told to "rate the similarity of the picture on a scale of 1 (least similar) to 5 (most similar). You should rate them according to how much they look alike—that is, how similar they are."

## Results

The participant ratings confirmed that the materials used by Goswami and Brown (1989) did not achieve the necessary stimulus structure. (See Appendix for item-by-item ratings of these stimuli.) The MA match should have been significantly more similar to the C term than the correct choice. Instead, the mean rating for the MA match ( $M = 3.00$ ,  $SD = .36$ ) was not significantly higher than that of the correct choice ( $M = 2.56$ ,  $SD = .73$ ),  $t(10) = 1.29$ , NS. The possibility of a dilution effect was also confirmed. The CO-WT item ( $M = 3.85$ ,  $SD = .88$ ) had high similarity to the C term (significantly more so than the MA item,  $t(10) = 2.20$ ,  $p < .05$ ), as did the identity match (which was not rated but is assumed to have a similarity to the C term of 5.0). Thus, object similarity responses could be distributed across the mere-appearance, CO-WT, and the identity match (as well as the correct choice). In fact, the high-similarity CO-WT alternative was the dominant error for both 3- and 4-year-olds.<sup>5</sup> Goswami and Brown interpreted children's propensity to choose the CO-WT match as evidence that they understood the basic notion of a causal transformation but had failed to maintain the proper transformation. However, given its strong similarity to the C alternative, it is clear that object similarity could also have accounted for its dominance.

Based on the previous analysis, it is clear that the children's tendency to make object similarity responses cannot be judged by the frequency of the mere-appearance choice; rather, we need to examine their performance in light of the actual similarity structure of the alternative choices as revealed by the similarity ratings. Using these ratings, we can compare the frequency of a particular alternative choice and its rated similarity to determine whether the children's choices were affected by the presence of object similarity. Goswami and Brown (1989) would predict no specific pattern of responses based on object similarity, while the relational shift hypothesis would predict a pattern of responses in which the object similarity of an alternative affects the likelihood of it being chosen. We found that, consistent with the relational shift hypothesis, these similarity ratings were highly predictive of children's responses. A Pearson-product moment correlation between the mean similarity rating of each alternative and the mean number of

<sup>5</sup>For 3-year-olds, but not for 4-year-olds, the identity choice was also a frequent error. Though not rated, this is of course the alternative most similar to the C term. Given feedback on each trial, 4-year-olds were evidently able to learn not to choose this highly distinctive match.

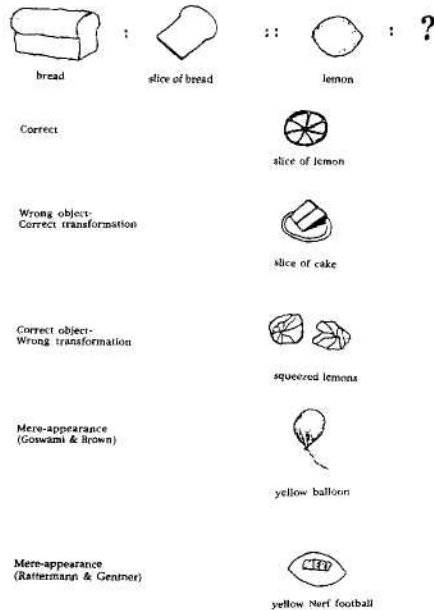


responses (regardless of correctness) for that alternative was significant for both 3-year-olds,  $r(24) = .52$ , and 4-year-olds,  $r(24) = .53$ .

## Discussion

These analyses confirmed the hypothesized difficulties with Goswami and Brown's study. The similarity ratings obtained demonstrate that similarity was not properly controlled. At least two factors mitigated against obtaining an object similarity effect: (1) The relatively low similarity of the mere-appearance choice and the relatively high similarity of the correct choice to the C term (consequently the desired contrast between object similarity and relational similarity did not obtain); and (2) the presence of other incorrect alternatives—the CO-WT choice and the identity choice—that were much higher in object similarity than the MA match, produced a dilution effect. These factors suggest that the children's rate of selecting MA matches substantially underestimated their propensity to choose on the basis of object similarity.

Thus, the conclusion that young children were resistant to object similarity matches is not supported by the data. The children's low level of mere-appearance responding can be accounted for by the low similarity of the intended object similarity choice and by the dilution effect of other competing object matches. More-



**Figure 1.** Sample stimuli from Goswami and Brown's (1989) Experiment 2, together with the amended MA choice used in our experiment. (The stimulus sets were otherwise identical.)

over, when ratings of the stimuli's actual similarity structure were used to test the predictions of the relational shift hypothesis we found evidence suggesting that children's responses were predominantly similarity based.

This evidence, however, is based on post-hoc analyses of Goswami and Brown's (1989) data. In order to properly test the predictions of the relational shift hypothesis we used Goswami and Brown's paradigm with modifications aimed at creating the necessary similarity structure. Specifically, we presented children with a true mere-appearance choice, one that, relative to the correct answer, was highly similar to the C choice. We also reduced the number of alternative object similarity choices, thus mitigating the dilution effect. The analogies were adapted from those used by Goswami and Brown, with the change that the mere-appearance choices were altered to be highly similar to the "C" terms. For example, the mere-appearance match for a lemon was changed to a yellow, lemon-shaped, Nerf football, and the mere-appearance match for an orange box of soap powder was changed to a red box of mashed potatoes. Figure 1 shows a sample stimulus set from Goswami and Brown's second experiment as well as our altered version of the set. (See Appendix for the complete set of stimuli.)

Similarity ratings for the new MA alternatives were obtained from a new group of 15 paid adult participants using the procedure described previously. We found that our alterations of the MA stimuli had the desired effect; the average rating of the MA matches was 3.88 ( $SD = .39$ ) (compared to 3.0 ( $SD = .36$ ) for the MA match and 2.56 ( $SD = .73$ ) for the correct alternative used by Goswami and Brown, 1989).

An additional change was that we eliminated the identity match to the C term, reducing the number of possible response alternatives from five to four choices: *correct*, *wrong object-correct transformation*, *correct object-wrong transformation*, and *mere-appearance*. This was done chiefly to reduce the number of highly similar alternatives, thus reducing the dilution effect referred to earlier, wherein children's high-similarity responses are spread over several alternatives. In addition, four alternatives seemed an easier array for preschoolers to manage than five.

### Causal Reasoning Control Task

For the causal reasoning control (CRC) task we tested the children's knowledge of the causal relations by constructing a causal sequence based upon the first two terms of each analogy. In this way we ensured that the correct answer in the CRC trial was not the correct answer in the corresponding analogy trial.

For example, in the analogy for the relation "cut"

bread (A) : slice of bread (B) :: lemon (C) : ??

our causal reasoning control was

bread : knife cutting the bread : ??

with the correct answer being "a slice of bread." We also created for the CRC trials a new *wrong object-correct transformation* (slice of pie), *correct object-wrong transformation* (bread wrapped in plastic package), and *semantically-related end-state* (garlic toast)<sup>6</sup> to fit with the new causal event.

Had we constructed the CRC task using the last two terms of the analogy (e.g., lemon: slice of lemon), the answer to the causal task would have also been the answer in the corresponding analogy trial. Note that in their experiment Goswami and Brown (1989) did use the last two terms of the analogy in the CRC, leaving open the possibility of a fairly sizable cross-task effect on the children's performance in the analogy task, an issue we will discuss more fully below.

Although the general procedure was based closely on that used in Goswami and Brown's (1989) Experiment 2, the practice task was slightly different, as described in the Method section. Finally, instead of using 3-, 4- and 6-year-olds, as in Goswami and Brown's study, we used 4- and 5-year-olds. We had originally intended to use 3-year-olds as well, but pilot studies proved discouraging; we had a high rate of nonperformance. (Our participants in general performed at a slightly lower rate of correct responding than Goswami and Brown's, an issue we will return to later). As in Goswami and Brown's Experiment 2 the children received feedback after every trial. When they made a mistake, the experimenter showed them the correct choice and stated the causal chain underlying the analogy.

To review, the relational shift hypothesis predicts that the younger children should show higher rates of object-similarity errors and lower rates of relational responding than the older children. Thus, mere-appearance responses in particular, and high-similarity errors in general, should diminish and correct relational responding increase as the children gain in age and experience. A second related prediction is that to the extent that children fail to select the correct relational response, high object-similarity responses should constitute a significant portion of their errors. The relational primacy view predicts no special propensity to make object-similarity errors, even in the youngest group, and no overall shift from object-based to relational responding. Rather, children should give a high level of relational responses throughout and their errors should be distributed nonsystematically across error types. Both views predict that children's performance on the analogy items should increase with knowledge of domain relations, as assessed by their performance on the causal-control task.

---

<sup>6</sup>The new stimuli were tested on five 3-year-olds and five 4-year-olds to be sure that these pictures were readily identifiable. All the children could readily identify all the objects pictured.

## EXPERIMENT 1B

### Method

**Participants.** The participants were twelve 4-year-olds (mean age 4 years, 6 months) and twelve 5-year-olds (mean age 5 years, 5 months) from middle-class families in the Champaign-Urbana area.

**Materials.** We used the six analogies and the six corresponding causal-reasoning items that were used in Goswami and Brown's (1989) Experiment 2, with the modifications described above. For the analogy task, the four response alternatives were (D) correct, (E) wrong object-correct transformation (WO-CT), (F) correct object-wrong transformation (CO-WT), and (G) mere-appearance (MA). For example, for the relation "cut," the children were shown:

bread : slice of bread :: lemon : ??

with the response alternatives of

- D: slice of lemon (Correct)
- E: slice of yellow cake (WO-CT)
- F: squeezed lemons (CO-WT)
- G: yellow Nerf football (MA).

For the CRC task, the four alternatives were (C) correct, (D) wrong object-correct transformation (WO-CT), (E) correct object-wrong transformation (CO-WT), and (F) semantically-related end-state (SEM). For example, for the relation "cut," the children were shown:

bread : knife cutting the bread : ??

with the response alternatives of:

- C: slice of bread (Correct)
- D: slice of cherry pie (WO-CT)
- E: bread in a white bag (CO-WT)
- F: garlic bread (SEM).

As in the Goswami and Brown (1989) studies, the stimuli were color pictures 10.16 cm by 15.2 cm in size, mounted on cardboard and covered with plastic laminate. The complete set of materials is given in the Appendix.

**Design.** The design was a 2 x 2 between-participants design: Age (4- and 5-years) x Order (Analogy to CRC and CRC to Analogy). Thus, half the children

received the analogy trials first and half received the CRC trials first. There were six analogy trials and six CRC trials.

### Procedure

**Practice Task.** The practice task consisted of a verbally presented causal chain. The children were told that they would play a game about choosing pictures and were asked questions like "If I show you a picture of a watermelon and a picture of someone throwing the watermelon off of a building, what would the next picture look like?" After responding correctly, they were then shown the first trial of either the analogy task or the causal reasoning task, depending upon condition. (In contrast, in Goswami and Brown's (1989) practice task, children were shown pictures of a sample analogy and asked to complete the series.)

**Analogy Trials.** The children were shown the pictures of the first three terms, placed in a straight line, and were asked "to choose the card that finishes the pattern. See if you can work out how the pattern goes." The experimenter then laid out the response alternatives in a random order. When children picked an incorrect alternative, they were shown the correct response and the causal chain of the analogy was explained: for example "The bread is cut to make a slice of bread and the lemon is cut to make a slice of lemon."

**Causal Reasoning Control Trials.** In the CRC trials participants were shown the first two terms of the causal sequence placed in a straight line and again asked to "finish the pattern." The experimenter then presented the four possible responses. As in the analogy task, children received corrective explanatory feedback on their responses.

### Results

**Testing the Relational Shift.** As predicted by the relational shift hypothesis, the proportion of relational (correct) responses increased with age (from .35 for

**Table 1. Mean Similarity Ratings for Stimuli Used in Experiment 1b, Together With the Children's Choice Proportions**

	Similarity Ratings	Choice Proportion	
		4-yr-olds	5-yr-olds
Correct	2.56 (.73)	.35	.67
Correct object-			
Wrong transformation	3.85 (.88)	.35	.28
Mere Appearance	3.88 (.39)	.22	.03
Wrong object-			
Correct transformation	1.13 (.23)	.07	.03

*Notes.* The scale was 1 (least similar) to 5 (most similar). The identity choice was not rated.

Standard deviations appear in parentheses.

the 4-year-olds to .67 for the 5-year-olds,  $t(22) = 3.37, p < .01$ ) and the number of mere-appearance choices decreased with age (from .22 for the 4-year-olds to .03 for the 5-year-olds,  $t(22) = 3.76, p < .01$ ). Only the 5-year-olds scored significantly above chance on the analogy trials ( $t(11) = 7.369, p < .0001$ ). The performance of the 4-year-olds was not significantly above chance ( $t(11) = .968, p > .18$ ). (Chance here is .25.) Table 1 shows the children's choice proportions along with the mean similarity ratings; item-by-item ratings are given in the Appendix.

A 2 (age)  $\times$  2 (order) mixed-measures analysis of variance (ANOVA) on the number of correct responses on the analogy trials across participants revealed a significant main effect of age ( $F(1, 20) = 9.48, p < .01, MSE = 2.33$ ). There were no significant effects or interactions with order; performance on the analogies task was the same whether it preceded (.53 proportion correct) or followed (.49 proportion correct) the causal knowledge task.

Because we had retained as much of Goswami and Brown's (1989) design as possible, our choice sets included other high-similarity alternatives besides the MA match. To assess the effects of object similarity, we separated the alternatives into three similarity categories based on the adult similarity ratings: *correct*, *high-similarity errors* (errors with a mean similarity-to-C rating of 3.5 or higher), and *low-similarity errors* (those rated 2.0 or lower). (Note that several alternatives fell between these two categories and were not included in the similarity analyses.) As shown in Table 2, the proportion of high-similarity errors was quite substantial for the youngest group and decreased with age, from .39 for 4-year-olds to .22 for 5-year-olds ( $t(22) = 2.37, p < .01$ ).<sup>7</sup> Thus, the results show a strong influence of object similarity on the performance on the younger group, as predicted by the relational shift hypothesis.

### Effects of Causal Knowledge

To test whether the degree of causal knowledge predicted performance on the analogies task, we computed the correlations between performance on the CRC task and performance in the analogy task. These were not significant for the separate age groups, but reached significance when we combined both age groups, ( $r(24) = .60$ ).

As a further check on the effects of causal knowledge, we divided the items into two halves based on the children's ability to correctly complete the causal sequence in the CRC task. We then submitted the two groups of items to a 2  $\times$  2 mixed-measures ANOVA (Age  $\times$  Familiarity) on the number of correct responses. This revealed a significant main effect for age, ( $F(1, 20) = 9.83, p < .005, MSE = 2.05$ ), and a marginal main effect for familiarity, ( $F(1, 20) = 3.98, p < .06, MSE = 2.05$ ). Although the difference between the familiar ( $M = 3.58, SD = .82$ ) and the

<sup>7</sup>Chance levels for the three categories were as follows: correct, .25; high-similarity errors, .30; low-similarity errors, .20. Because medium-similarity errors were omitted from the analysis the levels do not add up to 1.0.



**Table 2. Response Types in Experiment 1b Categorized by Similarity to the "C" Choice**

	Correct	High Sim Errors	Low Sim Errors
4-year-olds	.35	.39	.07
5-year-olds	.67*	.22	.02*

Note. \* = above chance.

unfamiliar ( $M = 2.42$ ,  $SD = .96$ ) causal relations was only marginally significant, the direction is consistent with the claim that children did better on analogies when they understood the causal relation.

### Exactly What Can a Young Child Do?

In the previous experiment we found that the 4-year-olds we tested were unable to perform the relational mapping at an above chance level (.35 correct relational responses, where chance was .25), in contrast to the 3-year-olds tested by Goswami and Brown (1989; .36 correct relational responses, where chance was .20). This leaves us with a puzzle; given the high level of consistency between our experiment and Goswami and Brown's, why is there such a large discrepancy between the performance of the two groups of children? While our manipulation of the object similarity choice can account for a portion of this difference, there is still a considerable amount of variation in the performance of the two populations that is left unexplained. It is important to know at what age and experience level children can carry out a causal analogy, for the precise age at which children acquire this ability has ramifications for many theories of children's intellectual development.

The reason for the large gap between the performance of the children in our experiment and those tested by Goswami and Brown (1989) appears to be a cross-task effect from the CRC task to the analogy task. As we briefly noted in our description of the causal reasoning task we designed, within each set of stimuli used by Goswami and Brown the correct choice in the analogy task was also the correct choice in the CRC task. For example, the causal reasoning control for the analogy concerning cutting was:

lemon:knife cutting the lemon : ??

The correct answer, the alternative "slice of lemon," was also the correct answer to the corresponding analogy:

bread(A):slice of bread (B)::lemon (C):??

Because the same picture was correct for both tasks, and because children were always shown the correct response on the CRC trials, children's greatly superior performance when the CRC task preceded the analogy task could have resulted

from a tendency to select the prior correct answer from the CRC task. Because the CRC task preceded the analogy task on two-thirds of the trials, such cross-task effects would have had a large influence on the overall results. Given these concerns about the stimulus structure and about the possible cross-task effects, Goswami and Brown's (1989) conclusions must be reexamined.

**Examination of Order Effects.** In their research, Goswami and Brown (1989) reported a main effect of task order, but noted that in the overall Age  $\times$  Order  $\times$  Condition analysis there were no significant interactions between order and the other factors, and so did not pursue the order effect further. Our further analysis of the data, however, revealed that the children's proportion correct for analogy trials that followed CRC trials was about twice as high as for analogy trials that preceded CRC trials (.91 vs. .48 for the 4-year-olds and .54 vs. .19 for the 3-year-olds). Specifically, in Experiment 2, Goswami and Brown (1989) used a within-participants design that included appearance-same (AS) analogies and appearance-differs (AD) analogies in addition to the causal-reasoning control task. The appearance-same analogies (which were used in both Experiments 1 and 2) were modified to form the appearance-differs analogies (which were the focus of our experiment). *T*-tests performed on the data from the analogy task confirmed the superiority of performance when the causal control preceded the analogy task (CRC-AD-AS versus AD-AS-CRC:  $t(12) = 2.00$ ,  $p < .05$  for the 3-year-olds,  $t(11) = 3.07$ ,  $p < .01$  for the 4-year-olds; and AS-CRC-AD versus AD-AS-CRC:  $t(12) = 2.33$ ,  $p < .05$  for the 3-year-olds and  $t(12) = 3.23$ ,  $p < .01$  for the 4-year-olds). This striking change in performance suggests that carryover from the CRC task inflated the results of the AD task.

**Re-Analysis of Performance.** To accurately evaluate the effect of object similarity on children's performance in Goswami and Brown's (1989) experiment, we examined the one-third of the AD trials that were clearly free of cross-task effects: namely, those on which the AD task preceded the CRC task. To assess the effects of object similarity, we again separated the alternatives into three similarity categories based on the adult similarity ratings. The proportion of responses for each response choice revealed a marked effect of object similarity on the children's AD performance when AD preceded CRC (i.e., when no cross-task effects were possible). For 3-year-olds, the dominant response (.47) was high-similarity errors, as compared to .19 correct responses and .14 low-similarity errors. Indeed, the high-similarity error category was the only category significantly above the chance level of .33.<sup>8</sup> The 4-year-olds performed better (.48 correct, significantly above chance), but still produced .31 high-similarity errors, as compared with .09 low-similarity errors. Thus, both age groups showed an effect of object similarity. Fur-

<sup>8</sup>Chance levels for the three categories were as follows: correct, .20; high-similarity errors, .33; low-similarity errors, .20. Because medium-similarity errors were omitted from the analysis, the levels do not add up to 1.0.

ther, consistent with the predictions of the relational shift hypothesis, the object-similarity effect was strongest for 3-year-olds. Also as predicted, there was a shift with age and experience from predominantly object-based responding among the younger children to predominantly relational responding among the older children.

In summary, we first note that the cross-task effects in Goswami and Brown's Experiment 2 apparently led to an inflation in the correct response rate and an underestimation of the rate of similarity-based responding, and second that both age groups showed an effect of object similarity. Thus the findings do not constitute evidence against the relational shift hypothesis. In fact the results of this re-analysis bear out the relational shift hypothesis.

### GENERAL DISCUSSION

The goal of these experiments was to test the relational shift hypothesis and, specifically, to respond to Goswami and Brown's (1989) empirical studies that seemed to falsify the hypothesis. In Experiment 1a we demonstrated the effects of object similarity on children's analogical reasoning through a re-analysis of Goswami and Brown's own data. In Experiment 1b we partially replicated and extended their work in order to provide a true test of the relational shift hypothesis. The results of both experiments support the existence of a relational shift in children's analogical reasoning.

The relational shift hypothesis predicts (1) an initially high level of object-similarity responses, dropping with age; (2) an increase in relational responding with age. In Experiment 1b we found that, when the response alternatives from Goswami and Brown's (1989) experiment were grouped according to their rated similarity to the C term and the analysis is restricted to data that are free from possible cross-task effects, the results fit the relational shift hypothesis. Object-similarity responses are the dominant response category (47% among 3-year-olds, diminishing to 31% among 4-year-olds). Relational responding shows the predicted increase, from 19% among 3-year-olds to 48% (the dominant category) for 4-year-olds. Also in Experiment 1b, we altered Goswami and Brown's method to create a stimulus structure more in accord with the logic of the hypothesis. The results again supported the relational shift: relational responding rose from 35% for the 4-year-olds to 67% for the 5-year-olds and object-similarity responses declined from 39% for the 4-year-olds to 22% for the 5-year-olds. Thus, contrary to Goswami and Brown's conclusion, there was both a drop in object-similarity responding and a rise in relational responding across the age range studied.

These results are consistent with our *career of similarity* account (Gentner & Rattermann, 1991) in which we propose that there is a systematic evolution in the kinds of relational comparisons that can be made as knowledge within a domain deepens; from overall similarity to object similarity, to relational similarity, and finally, to higher-order relational comparisons. This evolution of similarity com-

parisons is a reflection of a multipurpose similarity mechanism, present at birth, that aims for all common matches. This mechanism is dependent upon domain knowledge; children must possess the appropriate relational knowledge in order to make a relational comparison. On Goswami's (1992, 1996) account, children are from the beginning actively seeking relational interpretations of analogies. There is no change in the kind of comparisons that can be made, only an increase in the domains in which relational comparisons can be performed. This account suggests that humans possess a mechanism dedicated to forming relational comparisons, and that object similarity merely hinders this mechanism's performance. What is not clear from Goswami's relational primacy account is how ordinary mundane similarity is computed. What do we do when matching based on common object properties as well as common relational similarity is appropriate? It would seem that separate mechanisms for analogy and for other more mundane similarity types, such as literal similarity or object-based matches, would be required. While such a situation is logically possible, it is counterintuitive in that ordinary similarity and analogy lie along a continuum of similarity types—thus one multipurpose similarity mechanism could compute both.

Goswami's (1992, 1996) relational primacy view agrees in some respects with our *career of similarity* account (Gentner & Rattermann, 1991). Both theories are knowledge-driven rather than stage-based and, in contrast to more purely stage-based views of analogy, both propose that the chief limit on children's ability to arrive at relational interpretations is their early lack of knowledge. The fundamental difference between the two views lies in our differing accounts of the developmental course of similarity. On our account, objects and relations both participate in the mapping throughout. This follows from structure-mapping theory, in which carrying out a comparison process involves matching both the objects and the representational structure within which they are embedded (e.g., Forbus, Gentner, & Law, 1995; Gentner, 1989). This view is supported by empirical evidence and computational simulations showing that objects and relations both participate integrally in the process of structural alignment (Gentner & Markman, 1993; Goldstone, 1994; Goldstone, Medin, & Gentner, 1991; Markman & Gentner, 1993ab, 1996; Medin, Goldstone, & Gentner, 1993; See also Halford, 1993; Halford, Wilson, Guo, Gayler, Wiles, & Stewart, 1994; Holyoak & Thagard, 1989, 1995; Keane, 1988; Spellman & Holyoak, 1992, 1993). On this view, a comparison becomes more similar if either object commonalities or relational commonalities are added to the shared system.

How can we decide between our account of the development of similarity in which object similarity and relational similarity work together to promote a match, and Goswami's (1992, 1996) account in which relations are integral and objects are not? Zelazo and his colleagues (Zelazo, Carter, Reznick, & Frye, 1997) have suggested that to untangle factors relevant to a process it is necessary to compare the operations of these factors in oppositional and in facilitating roles. For instance, the sentences below place object similarity and relational similarity

in both an oppositional (cross-mapped) relation and a facilitating (literal similarity) relation. On the structure-mapping theory, children and adults will both find a literal similarity match, such as that between sentences 1 and 2, more similar than a purely relational match, such as that between sentences 1 and 4, and further, will have difficulty ignoring extraneous similarity, such as that in the cross-mapped sentences 1 and 3.

- (1) The squirrel helped the robin escape from the evil trout.
- (2) The chipmunk helped the bluebird escape from the malevolent salmon.
- (3) The trout helped the squirrel escape from the evil robin.
- (4) The secretary rescued her boss from the taxman.

Goswami's relational primacy account proposes that children (and presumably adults) actively seek and focus on relational commonalities regardless of the match or mismatch among the objects. Object commonalities can affect the match process, but only as a *performance variable*—as a task feature that can either help or hinder the ability to perform a relational mapping—or as a “fallback” strategy to be used when relational knowledge fails, and not as an integral part of the process. Thus sentence (1) should match equally well with sentences (2), (3), and (4), because the relational match is the same for all these pairs (and are familiar to normal adult participants). Adults might take longer to comprehend the pairs with inconsistent object matches, but the end result should be equivalent.

However, there is evidence from adults that suggests that object matches are integral to the computation and evaluation of comparisons (Gentner, Rattermann, & Forbus, 1993; Markman & Gentner, 1993ab; Medin, Goldstone, & Gentner, 1993). Processing a comparison appears to involve interactive levels of matching, with object matches generally computed before relational matches (Goldstone & Medin, 1994; Ratcliff & McKoon, 1989). Further evidence indicates that object similarity is crucial in similarity-based access from memory (Gentner, Rattermann, & Forbus, 1993; Holyoak & Koh, 1987; Keane, 1985; Novick, 1988; Ross, 1989), as well as transferring a prior solution (Gentner & Schumacher, 1986; Reed, 1987; Ross, 1989). Conceivably, some of these results could be construed as reflective of mere performance factors. However, what is especially troublesome for the view that object commonalities are merely a performance factor is the pattern provided by adults' judgments of similarity. In accord with the intuition that relational matches are privileged, Gentner, Rattermann and Forbus (1993) found that, in general, adults considered analogies (pure relational matches) to be more similar than purely object-level matches. However, these same adults rated overall literal similarity (both objects and relations in common) as still more similar than analogies. It is not clear how this pattern of performance could result from a mechanism dedicated to relational similarity; rather, this suggests a general similarity mechanism that considers both object and relational similarity.

Other developmental findings support our view that object similarity and relational similarity work together in the mapping process. We have performed experiments that manipulate whether object similarity and relational similarity play a facilitating or oppositional role (cf., Zelazo et al., 1997). For example, Gentner and Toupin (1986) asked children to act out a simple story with a set of toy characters and then to re-enact the same plot using a new set of toys. When given literal similarity matches 6-year-old children could accurately retell the story (approximately 90% correct), however, their accuracy dropped (approximately 60% correct) when given no supporting object similarity. The worst performance (approximately 50% correct) occurred when object similarity was cross-mapped (i.e., in opposition to the relational match) between the sets. Although the 9-year-olds generally performed better, they showed the same pattern; better performance when a facilitory relation held (approximately 90%), and poorer performance when an oppositional relation held (approximately 60%) between object similarity and relational similarity. Rattermann, Gentner, and DeLoache (1991, in preparation) also found that when given a simple mapping task between ordered sets of objects, 4-year-old children were most accurate in a literal similarity condition in which object and relational similarity were in harmony and least accurate when object similarity and relational similarity were in opposition. These results run counter to predictions of Goswami's relational primacy view; if objects are extraneous to the comparison of relational similarity there should be no facilitory effect of object similarity in a relational similarity task.

The assumption that children focus solely on relational commonalities seems implausible on two counts. First, the evidence just discussed demonstrates that adults use both object commonalities and relational commonalities in processing comparisons. It seems unlikely that children would forego the use of object commonalities and adopt a more abstract relational form of processing than adults. On the contrary, empirically children appear to be *more* influenced by object-level matches than are adults. The second difficulty is that it is not clear how ordinary mundane similarity, such as that in sentences (1) and (2) above, is processed on the relational primacy account, whether by separate mechanisms or by the relational primacy mechanism with object similarity added. The relational primacy account would seem to lead to the prediction that children disregard object similarities when processing overall similarity comparisons. It seems more intuitive to assert that because there is no way for a child (or an adult) to know in advance whether a comparison will be an analogy or a literal similarity statement, the initial processes must be uncommitted to a particular type of similarity.

Our results are relevant to another important theory of analogical development. Halford (1989, 1992) has proposed a sequence of development that also includes a relational shift; children shift from being able to process *unary relations* (which correspond to object attributes) to *binary relations* (relations between two elements), to *ternary relations* (relations between three elements) and finally to *quaternary relations* (relations between four elements). In Halford's account an



important driver of the relational shift is changes in cognitive capacity; in order to carry out similarity matches children must have the capacity to represent and map the number of arguments necessary for a match. Children show a developmental increase in cognitive capacity that permits them to make more complex matches. However, Halford also postulates that changes in domain knowledge are important in the development of analogical skill. Our data agree with Halford's predicted pattern of performance in that object-based responding preceded relationally based responding. While the current experiment does not allow us address the issue of changes in cognitive capacity, the fact that increases in causal knowledge contribute to relational responding is consistent with the claim made by both theories that knowledge plays an integral role in children's ability to complete causal analogies.

### **Knowledge and the Relational Shift**

To recapitulate, the relational shift refers to a change from early attention to common object properties to later attention to common relational structure (Gentner, 1988; Gentner & Rattermann, 1991). This shift is assumed to follow an earlier focus on overall similarity, and is assumed to be based on changes in knowledge representation rather than on changes in cognitive competence. It is assumed that children in the initial stages of learning have better representations of objects than of the relations between them. This assumption stems from the more general novice-expert shift assumption that knowledge of objects tends to be acquired earlier than knowledge of the relational structure for both child and adult learners (e.g., Chi, Feltovich, & Glaser, 1981). How these changes in knowledge representation come about is beyond the scope of this paper. In our prior work, we have suggested two mechanisms that may promote the learning of relational structure: (a) the acquisition of relational language, and (b) progressive abstraction brought about by structural alignment (Gentner & Rattermann, 1991; Gentner, Rattermann, Markman, & Kotovsky, 1995; Rattermann, Gentner, & DeLoache, in preparation).

### **The Role of Causal Knowledge**

Despite our disagreement with the role assigned to object similarity in the relational primacy view, there are some points of fundamental agreement between our position and that of Goswami. Because we believe that the relational shift results primarily from knowledge accretion, rather than from maturation (Gentner & Rattermann, 1991), we agree with Goswami (1992, 1996) and with Goswami and Brown (1989) in predicting that children's performance on the analogy items should improve with their level of knowledge of the causal relations. Consistent with this prediction, we found that children's performance on the analogy task was correlated with their performance on the causal-knowledge assessment task, supporting the hypothesis that children's ability to process causal analogies depends largely on knowledge of the underlying causal mechanisms. This result

accords with other research on the role of domain knowledge in analogical development (Brown, 1989, 1990; Brown & Campione, 1984; Gentner, 1977a, 1977b; Gentner & Rattermann, 1991; Gentner, Rattermann, Markman, & Kotovsky, 1995; Goswami, 1991, 1992; Vosniadou, 1987).

### Reconsidering the Role of Similarity

Similarity is often cast as a kind of malign force in cognitive development: a misleading factor that interferes with children's ability to perceive deeper relationships. This feeling is engendered partly by the many experiments in which children's ability to make an insightful response is pitted against their attraction to incorrect object-similarity lures. Of course, this is a useful tension to exploit, and we have done our share of such experiments. But we should not lose sight of the fact that similarity, including object similarity, is far more often benevolent than malevolent. Perceptual similarity is often correlated with conceptual similarity (Medin & Ortony, 1989), and object similarity with relational similarity (Gentner, 1989; Gentner, Rattermann, & Forbus, 1993), thus attention to concrete similarity often results in deeper insights (Gentner & Medina, 1998). We close with an observation that illustrates this point:

A two-year-old boy is walking through the zoo with his mother and grandmother when he sees the tigers. His grandmother says, "See, the tiger has a big tongue. He's licking his paws." The child sticks his tongue out, points to it and asks his grandmother to stick hers out too. As the child continues to look at the tiger his mother says, "See, the tiger has a big tail that he's waving." The child watches for a moment as the tiger's tail waves back and forth. Then he glances over his shoulder to where his own tail would be if he had one. (Marilyn Shatz, personal communication, April, 1989).

The child's noticing of the object similarity between his tongue and the tiger's led him to compare and align the tiger's body with his own and to attempt a further inference from the tiger to himself. By testing and disconfirming that possible inference he learned something about the morphology of the human body relative to that of other mammals. Far from being harmful, similarity can lead children to consider new hypotheses that refine and enrich their knowledge.

### REFERENCES

- Brown, A. L. (1989). Analogical learning and transfer: What develops? In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 369-412). New York: Cambridge University Press.
- Brown, A. L. (1990). Domain-specific principles affect learning and transfer in children. *Cognitive Science*, 14, 107-133.
- Brown, A. L., & Campione, J. C. (1984). Three faces of transfer: Implications for early competence, individual differences, and instruction. In M. E. Lamb, A. L. Brown, & B. Rogoff (Eds.), *Advances in developmental psychology* (Vol. 3, pp. 143-192). Hillsdale, NJ: Erlbaum.

- Brown, A. L., Kane, M. J., & Echols, C. H. (1986). Young children's mental models determine analogical transfer across problems with a common goal structure. *Cognitive Development, 1*, 103-121.
- Chen, Z., & Daehler, M. W. (1989). Positive and negative transfer in analogical problem solving by 6-year-old children. *Cognitive Development, 4*, 327-344.
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.
- Forbus, K. D., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science, 19*, 144-206.
- Gentner, D. (1977a). Children's performance on a spatial analogies task. *Child Development, 48*, 1034-1039.
- Gentner, D. (1977b). If a tree had a knee, where would it be? Children's performance on simple spatial metaphors. *Papers and Reports on Child Language Development, 13*, 157-164.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155-170.
- Gentner, D. (1988). Metaphor as structure mapping: The relational shift. *Child Development, 59*, 47-59.
- Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199-241). New York: Cambridge University Press.
- Gentner, D., & Medina, J. (1998). Similarity and the development of rules. *Cognition, 65*, 263-297.
- Gentner, D., & Markman, A. B. (1993). Analogy - Watershed or Waterloo? Structural alignment and the development of connectionist models of cognition. In S. J. Hanson, J. D. Cowan, & C. L. Giles (Eds.), *Advances in neural information processing systems 5* (pp. 855-862). San Mateo, CA: Kauffman.
- Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on thought and language: Interrelations in development* (pp. 225-277). London: Cambridge University Press.
- Gentner, D., Rattermann, M. J., & Forbus, K. D. (1993). The roles of similarity in transfer: Separating retrievability and inferential soundness. *Cognitive Psychology, 25*, 524-575.
- Gentner, D., Rattermann, M. J., Markman, A. B., Kotovsky, L. (1995). Two forces in the development of relational similarity. In T. Simon and G. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 263-313). Hillsdale, NJ: Erlbaum.
- Gentner, D., & Schumacher, R. M. (1986). Use of structure mapping theory for complex systems. *Proceedings of the 1986 IEEE International Conference on Systems, Man, and Cybernetics* (pp. 252-258). Atlanta, GA. New York: IEEE.
- Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10*, 277-300.
- Goldstone, R. L. (1994). Similarity, interactive activation, and mapping. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20*, 3-28.
- Goldstone, R. L., & Medin, D. L. (1994). Time course of comparison. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20*, 29-50.

- Goldstone, R. L., Medin, D. L., & Gentner, D. (1991). Relational similarity and the non-independence of features in similarity judgments. *Cognitive Psychology*, 23, 222-264.
- Goswami, U. (1991). Analogical reasoning: What develops? A review of research and theory. *Child Development*, 62, 1-22.
- Goswami, U. (1992). *Analogical reasoning in children*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Goswami, U. (1996). Analogical reasoning and cognitive development. In H. Reese (Ed.), *Advances in child development and behavior* (pp. 92-135). San Diego, CA: Academic Press.
- Goswami, U., & Brown, A. L. (1989). Melting chocolate and melting snowmen: Analogical reasoning and causal relations. *Cognition*, 35, 69-95.
- Halford, G. S. (1987). A structure-mapping approach to cognitive development. *International Journal of Psychology*, 22, 609-642.
- Halford, G. S. (1992). Analogical reasoning and conceptual complexity in cognitive development. *Human Development*, 35, 193-218.
- Halford, G. S. (1993). *Children's understanding: The development of mental models*. Hillsdale, NJ: Erlbaum.
- Halford, G. S., Wilson, W. H., Guo, J., Gayler, R. W., Wiles, J., & Stewart, J. E. M. (1995). Connectionist implications for processing capacity limitations in analogies. In K. J. Holyoak & J. Barnden (Eds.), *Advances in connectionist and neural computation theory, Vol. 2: Analogical connections* (pp. 363-415). Norwood, NJ: Ablex.
- Holyoak, K. J., Junn, E. N., & Billman, D. O. (1984). Development of analogical problem-solving skill. *Child Development*, 55, 2042-2055.
- Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition*, 15, 332-340.
- Holyoak, K. J., & Thagard, P. R. (1989). A computational model of analogical problem solving. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 242-266). New York: Cambridge University Press.
- Holyoak, K. J., & Thagard, P. R. (1995). *Mental leaps: Analogy in creative thought*. MIT Press; Cambridge, MA.
- Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.
- Keane, M. (1985). On drawing analogies when solving problems: A theory and test of solution generation in an analogical problem-solving task. *British Journal of Psychology*, 76, 449-458.
- Keane, M. T. (1988). *Analogical problem solving*. New York: Wiley.
- Kotovskiy, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, 67, 2797-2822.
- Markman, A. B., & Gentner, D. (1993a). Splitting the differences: A structural alignment view of similarity. *Journal of Memory and Language*, 32, 517-535.
- Markman, A. B., & Gentner, D. (1993b). Structural alignment during similarity comparisons. *Cognitive Psychology*, 25, 431-467.
- Markman, A. B., & Gentner, D. (1996). Commonalities and differences in structural comparisons. *Memory and Cognition*, 24, 235-249.
- Medin, D., Goldstone, R., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, 100, 254-278.

- 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.
- Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 510-520.
- Piaget, J., Montangero, J., & Billeter, J. (1977). La formation des correlats. In J. Piaget (Ed.), *L'Abstraction reflechissante* (pp. 115-129). Paris: Presses Universitaires de France.
- Ratcliff, R., & McKoon, G. (1989). Similarity information versus relational information: Differences in the time course of retrieval. *Cognitive Psychology*, 21, 139-155.
- Rattermann, M. J., Gentner, D., & DeLoache, J. (1989, April). Effects of competing surface similarity on children's performance in an analogical task. Poster presented at the biennial Meeting of the Society for Research in Child Development, Kansas City, MO.
- Rattermann, M. J., & Gentner, D. (in preparation). The effect of relational language on children's performance in an analogical mapping task.
- Rattermann, M.J., & Gentner, D., & DeLoache, J. (in preparation). Effects of object and relational similarity on children's performance in a spatial analogies task.
- Reed, S. K. (1987). A structure-mapping model for word problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 124-139.
- Ross, B. (1989). Distinguishing types of surface similarity: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 456-468.
- Spellman, B. A., & Holyoak, K. J. (1992). If Saddam is Hitler then who is George Bush? Analogical mappings between systems of social roles. *Journal of Personality and Social Psychology*, 62, 913-933.
- Spellman, B. A., & Holyoak, K. J. (1993). An inhibitory mechanism for goal-directed analogical mapping. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 947-952). Hillsdale, N.J.: Lawrence Erlbaum.
- Sternberg, R. J., & Downing, C. J. (1982). The development of higher-order reasoning in adolescence. *Child Development*, 53, 209-221.
- Sternberg, R. J., & Nigro, G. (1980). Developmental patterns in the solution of verbal analogies. *Child Development*, 51, 27-38.
- Vosniadou, S. (1987). Children and metaphors. *Child Development*, 58, 870-885.
- Winner, E., Rosenstiel, A., & Gardner, H. (1976). The development of metaphoric understanding. *Developmental Psychology*, 12, 289-297.
- Zelazo, P. D., Carter, A. S., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem solving approach. *Review of General Psychology*, 1, 198-226.

## APPENDIX

Mean rated similarity to the C term for Goswami and Brown's appearance-differs alternatives and for the altered mere-appearance alternatives used in our Experiment 1b.

CUT: A (bread) : B (cut bread) :: C (whole lemon) : ?		
COR	slice of lemon	3.8
WO-CT	slice of cake	1.0
CO-WT	squeezed lemons	2.7
MA (GB)	yellow balloon	3.4
MA (RG)	yellow Nerf football	4.1
BREAK: A (gray vase) : B (broken gray vase) :: C (gray egg) : ?		
COR	broken gray egg	2.4
WO-CT	broken blue plate	1.1
CO-WT	multi-colored egg	4.1
MA (GB)	yellow lemon	2.8
MA (RG)	tan potato	4.1
WET: A (coffee crystals) : B (mug of coffee) :: C (orange box of Tide) : ?		
COR	soapy water	1.6
WO-CT	glass of Koolaid	1.1
CO-WT	spilled Tide	4.7
MA (GB)	blue box of cat chow	3.0
MA (RG)	red box of mashed potatoes	4.1
BURN: A (tree) : B (burnt tree stump) :: C (pile of leaves) : ?		
COR	smoking pile of leaves	2.9
WO-CT	burnt paper	1.0
CO-WT	swirling leaves	2.8
MA (GB)	pile of sticks	2.4
MA (RG)	pile of brown socks	3.9
OPEN: A (tub of ice cream) : B (ice cream cone) :: C (yellow box of Cheerios) : ?		
COR	blue bowl of Cheerios	2.4
WO-CT	blue bowl of Chex	1.6
CO-WT	torn Cheerios box	4.3
MA (GB)	red and blue box of crackers	3.3
MA (RG)	yellow Sunlight detergent	4.1
MELT: A (frozen river) : B (rushing river) :: C (snowman) : ?		
COR	melted snowman	2.3
WO-CT	melted chocolate	1.0
CO-WT	dirty snowman	4.5
MA (GB)	scarecrow	3.1
MA (RG)	marshmallow man	3.1
MEAN:		
COR	2.56	
WO-CT	1.13	
CO-WT	3.85	
MA (GB)	3.00	
MA (RG)	3.88	

---

Note. The scale was 1 (least similar) to 5 (most similar).