Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Experimental Child Psychology 94 (2006) 249–273
Journal of Experimental Child Psychology
www.elsevier.com/locate/jecp

Children's development of analogical reasoning: Insights from scene analogy problems

Lindsey E. Richland [a,*], Robert G. Morrison [b], Keith J. Holyoak [c]

[a] *Department of Education and Department of Psychology and Social Behavior, University of California, Irvine, Irvine, CA 92697, USA*
[b] *Xunesis, Chicago, IL 60626, USA*
[c] *Department of Psychology, University of California, Los Angeles, Los Angeles, CA 90095, USA*

Received 2 August 2005; revised 1 December 2005
Available online 19 April 2006

Abstract

We explored how relational complexity and featural distraction, as varied in scene analogy problems, affect children's analogical reasoning performance. Results with 3- and 4-year-olds, 6- and 7-year-olds, 9- to 11-year-olds, and 13- and 14-year-olds indicate that when children can identify the critical structural relations in a scene analogy problem, development of their ability to reason analogically interacts with both relational complexity and featural distraction. Error patterns suggest that children are more likely to select a distracting object than to make a relational error for problems that present both possibilities. This tendency decreases with age, and older children make fewer errors overall. The results suggest that changes in analogical reasoning with age depend on the interplay among increases in relational knowledge, the capacity to integrate multiple relations, and inhibitory control over featural distraction.

Keywords: Analogy; Relational complexity; Similarity; Inhibitory control; Working memory

Introduction

Analogical reasoning is an important component of children's higher order cognitive development. Analogy is a conceptual strategy enabling children to make inferences about novel phenomena, to transfer learning across contexts, and to extract relevant information

* Corresponding author. Fax: +1 949 824 2965.
E-mail address: l.e.richland@uci.edu (L.E. Richland).

0022-0965/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jecp.2006.02.002

from everyday learning experiences on the basis of relational similarity (Chen, Sanchez, & Campbell, 1997; Gentner, 1977; Goswami, 2001; Halford, 1993; Holyoak, Junn, & Billman, 1984). It has been argued that this sophisticated conceptual process is central to children's everyday learning; however, the underlying mechanisms that support the development of analogical reasoning are not yet well understood.

Possible mechanisms of developmental change

Within the literature on cognitive development, three major hypotheses have been advanced to account for age-related differences in analogical reasoning: increased domain knowledge, a relational shift from object similarity to relational similarity, and increased working memory capacity for manipulating relations.

Increased domain knowledge

Goswami and colleagues have proposed a relational primacy hypothesis, arguing that analogical reasoning is fundamentally available as a capacity from early infancy but that children's analogical performance increases with age due to the accretion of knowledge about relevant relations (Goswami, 1992, 2001; Goswami & Brown, 1989). Goswami's proposal for knowledge acquisition as a mechanism for development emerged in reaction to Piagetian studies suggesting that children are unable to reason analogically prior to achieving formal operations at approximately 13 or 14 years of age (Piaget, Montangero, & Billeter, 1977). Piaget's tasks frequently involved uncommon relations, such as "steering mechanism," which would likely have been unfamiliar to younger children. In contrast, Goswami and Brown (1989) found that children as young as 3 years could be successful on analogical reasoning tasks when they demonstrated knowledge about the relevant relations.

In a series of studies, Goswami, Leevers, Pressley, and Wheelwright (1998) presented children with complex versions of analogy tasks in which two physical causal relations (e.g., cutting and wetting) were manipulated to change one object, "A," into another object, "B." Children were required to map the relation between A and B to a different object, "C," and its transformed version, "D." They were given a set of alternatives and asked to identify the D object. On a second task, the children were tested to assess their knowledge of the causal relations used in each problem. Goswami and colleagues found that children as young as 4 years were fairly competent on these problems with two relational changes when they showed knowledge of the relations on an additional task, although 3-year-olds did not perform as well. The authors interpreted these data as evidence that domain knowledge is the primary constraint on children's analogical reasoning. However, as Goswami and colleagues noted, the knowledge-based account cannot fully account for age-related effects in young children's performance on analogical reasoning tasks. In particular, these authors pointed out that children seem to fail on analogies in systematic ways even when the children possess relational knowledge relevant to the task.

Relational shift

In an alternative formulation of young children's observed age-related increase in analogical reasoning performance, Gentner and Rattermann (1991; see also Rattermann &

Gentner, 1998) hypothesized that a "relational shift" occurs. They suggested that before the relational shift, children attend primarily to featural similarity between objects and will reason on the basis of perceptual features rather than on the basis of relational similarity. Following the relational shift, children can and will reason on the basis of relational features. Several studies have supported the relational shift hypothesis, demonstrating the deep interrelations between young children's processing of object similarity and their processing of relational similarity (Gentner, 1988; Gentner & Toupin, 1986). Importantly, however, the mechanisms underlying the observed relational shift remain unclear. Rattermann and Gentner (1998) proposed that knowledge of the relevant relations is central and argued that the relational shift is therefore domain specific in nature. They suggested that the relational shift occurs at different ages for different domains, depending on the child's knowledge of that domain.

Relational complexity

A third explanation for developmental changes in analogical reasoning highlights limits on children's working memory capacity that affect their ability to process multiple relations simultaneously. Halford and colleagues (Andrews & Halford, 2002; Halford, 1993; Halford, Andrews, Dalton, Boag, & Zielinski, 2002) have defined relational complexity in terms of the number of sources of variation that are related and must be processed in parallel. For example, a binary relation is defined as a relation between two arguments, both of which are sources of variation. Thus, "boy chases girl" specifies a single relation (chase) between two arguments (boy and girl). A reasoner would need to hold both arguments and the relevant relation in mind to reason on the basis of this relation. Similarly, a ternary relation includes three arguments as sources of variation. A special case of a ternary relation is formed by two integrated binary relations with three arguments such as "mom chases boy who chases girl." Using this metric of relational complexity, Halford (1993) argued for a developmental continuum in children's working memory capacity, such that children can process binary relations (a relation between two objects) after 2 years of age and can process ternary relations after 5 years of age.

Using an alternative formulation of a relational complexity metric, Zelazo and Frye (1998; see also Frye and Zelazo, 1998; Frye et al., 1996) have identified similar age-related developmental progressions. According to their cognitive complexity and control theory, complexity is defined by the number of hierarchical rules that must be considered to accomplish a task. For example, in the Dimensional Change Card Sort task, children were asked to follow a rule to sort by color (e.g., "if red ... here," "if blue ... here") and a rule to sort by shape (e.g., "if rabbit ... here," "if boat ... here"). The 3- and 4-year-olds were successful on these sorting tasks when performing them separately but failed when required to switch between tasks, integrating them with a higher order rule. Zelazo and Müller (2002) hypothesized that change with age depends on children's development of executive function and particularly the ability to reflect on the relation between two rules so as to develop and use a higher order rule that integrates the rule pair. Zelazo, Frye, and Rapus (1996) found that knowledge of relevant rules is not sufficient to allow young children to solve problems with higher levels of complexity; rather, this ability develops with age and increased control over thoughts and actions.

Investigating multiple factors in analogical development

We believe that it is useful to consider the potential role of multiple factors in the development of analogical reasoning. Although acquisition of relational knowledge doubtless is essential, it seems likely that additional developmental changes are also important. Constructing an analogy requires a reasoner to represent source and target analogs and to construct a mapping between elements of the source and target based on correspondences between relations in each (Gentner, 1983; Gick & Holyoak, 1980). These processes have been shown to depend on working memory functions (Morrison, 2005; Morrison, Holyoak, & Truong, 2001; Waltz, Lau, Grewal, & Holyoak, 2000), which in turn depend on developmental changes in the prefrontal cortex (Diamond, 2002). Using an analogy frequently involves mapping multiple relations, a process that has been shown to depend critically on areas of the prefrontal cortex associated with working memory (Christoff et al., 2001; Kroger et al., 2002; Prabhakaran, Smith, Desmond, Glover, & Gabrieli, 1997; Waltz et al., 1999). It follows that increases in capacity to cope with relational complexity (Halford, 1993; Zelazo & Müller, 2002) would be expected to lead to increased analogical ability.

In addition, responding on the basis of relational correspondences may compete with tendencies to respond on the basis of more superficial featural or semantic similarities between individual objects (Gentner & Toupin, 1986). Therefore, inhibitory control in working memory will be required when relational and more superficial responses conflict (Morrison et al., 2004; Viskontas, Morrison, Holyoak, Hummel, & Knowlton, 2004). Although inhibitory control has not been previously discussed directly as a factor in the development of analogical reasoning, this hypothesis is consistent with results from other cognitive tasks that explore developmental changes in children's ability to use inhibitory control. Children at 3 to 4 years of age may fail at a task that requires them to inhibit a salient response despite a complete understanding of the demands of the task. For example, Diamond, Kirkham, and Amso (2002) manipulated the day–night task, a Stroop-type task in which children are instructed to say "day" when shown a picture of a moon and "night" when shown a picture of sun. This task is designed to test inhibitory control because presumably children's semantic category of "day" is activated when they are shown a scene depicting a sun, but they are required instead to generate a word with the opposing semantic meaning, namely "night." The opposite case was also tested with children who were shown the moon and required to say the word "day." Children at 4 to 5 years of age failed on this task consistently; manipulations intended to reduce working memory load and instructions to "say the opposite" did not alter children's performance. Diamond and colleagues then reduced the inhibitory requirements of the task by asking participants to say "dog" and "pig" instead of "day" and "night." Under these conditions, there no longer was competition between the pictures and the semantic meanings of the words. Children's accuracy improved with the change, indicating that they understood the task and were capable of maintaining the rules of the task in mind. It appears that their inability to inhibit the more salient response produced their failure in the day–night version of the task.

Similarly, to select a relational match that is in competition with a more salient featural match in an analogy, the reasoner must inhibit the response to the featural match (Morrison et al., 2004; Viskontas et al., 2004). Thus, developmental changes in inhibitory control may explain data supporting the transition from children's selection of featural matches to children's selection of relational matches, as documented by Rattermann and Gentner (1998; see also Gentner & Rattermann, 1991).

Accordingly, acquisition of the full capacity to reason analogically seems likely to require both the ability to integrate multiple relations and the ability to inhibit tendencies to respond on the basis of competing superficial similarities (for a review, see Morrison, 2005). Both of these working memory processes feature prominently in a neurally plausible computational account of analogical reasoning, namely learning and inference with schemas and analogies (LISA) (Hummel & Holyoak, 1997, 2003). This model has been used to simulate patterns of disruption in analogy use associated with frontal and temporal lobe damage (Morrison et al., 2004) as well as with normal aging (Viskontas et al., 2004).

Overview of the current study

Although the roles of relational complexity, featural distraction, and relational knowledge in children's development of analogical reasoning have been investigated and widely debated, each possible explanation has been generally construed as mutually exclusive and has been largely examined separately from the others. Potential relations between the various hypotheses have not been thoroughly investigated, and none of the hypotheses alone appears to explain all of the systematic difficulties observed in young children's analogical reasoning.

The current study used scene analogy problems to examine how children's ability to notice and analogically map simple relations is influenced by variations in featural distraction and relational complexity of the stimuli. The scene analogy problems consisted of pairs of illustrated scenes depicting simple relations among objects, similar to stimuli originally developed by Markman and Gentner (1993). Unlike the Markman and Gentner stimuli, we varied relational complexity by designing matched problems that required mapping either one or two instances of the same relation. Featural distraction was manipulated by varying the identity of one object in the second (target) scene. Specifically, the second scene in the pair either included an object with great featural similarity to the object to be mapped in the first (source) scene or substituted another featurally dissimilar object. Because the same basic relation was used in all conditions, relational knowledge was held constant at a level that could be assessed by performance in the easiest condition (one-relation/no-distractor condition) (Experiments 1 and 2) or by asking children to correctly identify the relation present in a scene (Experiment 2).

Experiment 1

Experiment 1 used scene analogy problems (available from the first author on request) to investigate relational complexity and featural distraction within a single analogical reasoning task while controlling for knowledge of relations. Relations corresponded to motion verbs familiar to children by 3 years of age (e.g., *kiss, chase, feed*) (Gentner, 1978; Golinkoff, Hirsh-Pasek, Mervis, Frawley, & Parillo, 1995; Golinkoff, Jacquet, Hirsh-Pasek, & Nandakumar, 1996). The objects used to represent these relations were items regularly encountered by preschool age children, including humans, animals, and dolls.

Method

Participants

The participants were 68 children (22 3- and 4-year-olds, 21 6- to 8-year-olds, and 25 13- and 14-year-olds). Participants were recruited from area child care centers in New York City and Los Angeles, an elementary academic summer school program in New York City, and junior high school programs in the Los Angeles area. Although demographic data were not collected systematically, children were from both upper and lower middle-class neighborhoods. Children recruited from the New York City area were primarily of African American or Caucasian descent, and children recruited in Los Angeles area were of diverse ethnicities with a majority of Caucasian descent.

Materials and design

Fig. 1 depicts an example of the four counterbalanced versions that were created for each of the 20 picture sets. Each set factorially varied (a) the number of instances of the relevant relation that needed to be mapped (one or two) and (b) the presence of an object in the target scene that was either featurally similar to (distractor) or dissimilar to (nondistractor) the object to be mapped in the source scene. Two-relation problems were created by having one inactive object in the one-relation problems participate in the relation for the two-relation version. For example, a woman who observes a boy chasing the girl in the one-relation version of the problem (i.e., chase (boy girl)) now joins in the chase in the two-relation version (i.e., chase (woman boy girl)). Distractor and nondistractor versions were created by having an extra object in the same picture that was either similar to (sitting cat) or dissimilar to (sandbox) the item to be mapped in the source picture (running cat). Distractors were either exact matches to the object to be mapped in the source (e.g., running cat) or slight variations of the same object (e.g., cat chasing, cat sitting).

The objects that were used as featural distractors were varied to ensure that participants would not learn, for example, that the correct answer was always a human. Distractors were animate objects (adults, children, or animals) or inanimate objects (mobile automobiles, stationary furniture, or cooking items). To ensure that the featural distractors were indeed featurally similar to the highlighted source objects, 10 undergraduates were asked to select the most featurally similar object to the target in the two-relation/distractor version of each stimulus. Participants selected the intended featural match 96% of the time, indicating that the manipulation of featural similarity was valid.

All pictures contained extra items not depicting the relevant relation, and the number of total objects was controlled across picture pairs. The spatial locations of the extra objects and distractor/nondistractor objects were held constant across the two conditions. In addition to the previously noted differences, our scene analogy problems differed from the Markman and Gentner (1993) stimuli in that distractors were never placed in key relational roles (allowing featural and relational errors to be coded separately) and the number of objects in each picture was controlled. Most image sets contained a total of five objects.

Given the dual nature of the relations depicted, the objects that played the pivotal role in the two-relation problems (e.g., the cat in the source and the boy in the target picture presented in Figs. 1C and D) were relationally located between the other two objects involved in the critical relation. Care was taken to ensure that the spatial

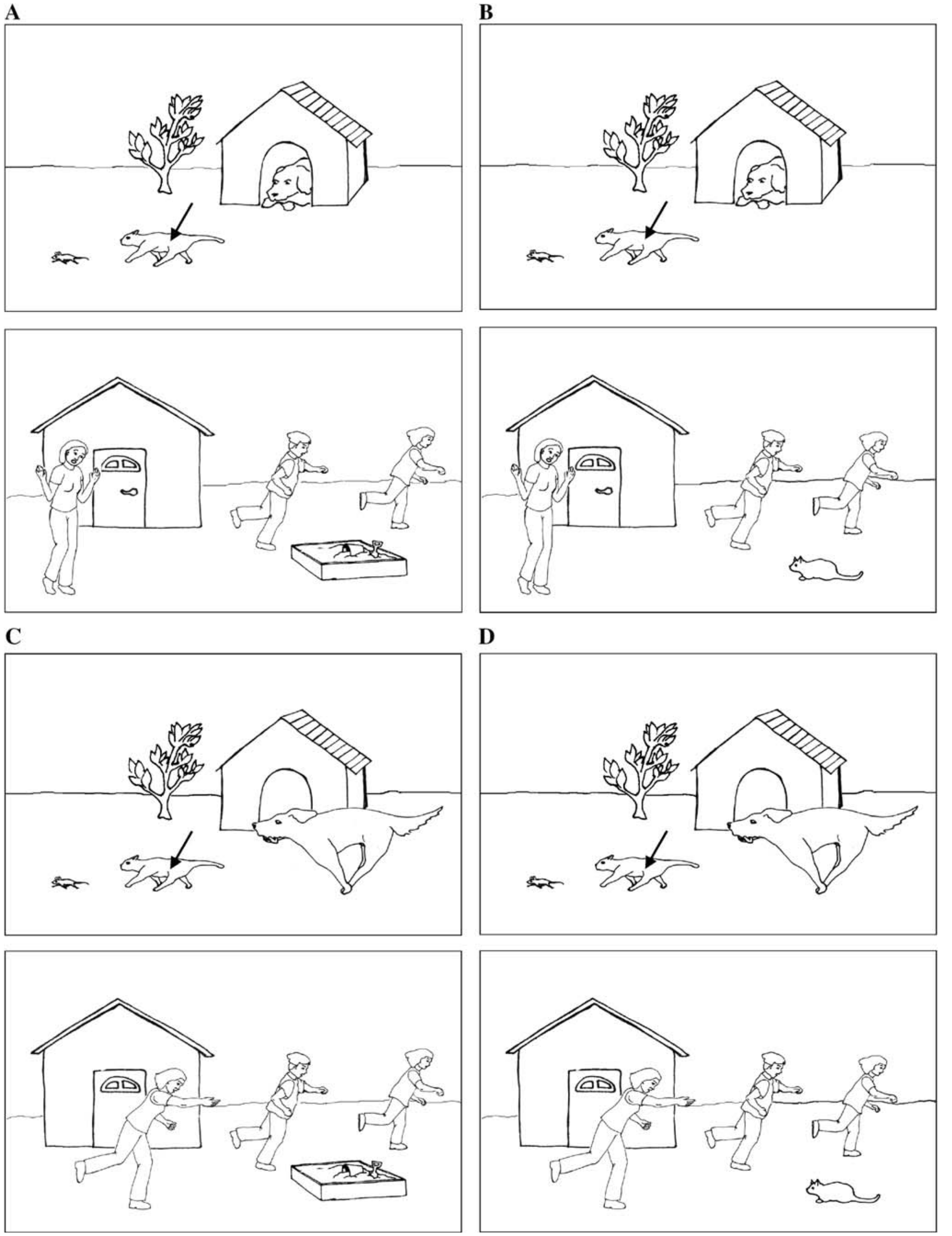

Fig. 1. Sample stimuli for "chase" problem: (A) one relation/no distractor; (B) one relation/distractor; (C) two relations/no distractor; (D) two relations/distractor.

organization of the pictures was varied such that the correct relational choice was not always located between the other two relational objects. Even so, this spatial arrangement held for 15 of the 20 scene pairs. Thus, it is possible that children sometimes could have solved two-relation problems by mapping the "middle" item in the source and target (although they first would need to identify the three relevant relational objects because there would be no way in which to reliably pick the middle object out of the five total objects in each scene). It should be noted that such a strategy would work for two-relation problems but not for one-relation problems and, hence, could lead to relatively higher performance in the two-relation problems than in the one-relation problems. This result would be in contrast to the relatively greater difficulty we predicted for that condition.

The 2×2 repeated-measures design generated four conditions: one relation with no featural distractor (one relation/no distractor), one relation with featural distractor (one relation/distractor), two relations with no featural distractor (two relations/no distractor), and two relations with featural distractor (two relations/distractor). Packets of picture pairs for each participant were organized such that five examples of each condition were included in a random order. The assignment of specific picture pairs to each of the four conditions was counterbalanced across participants in each age group. The three age groups constituted an additional between-subjects factor. The dependent variable was the participants' accuracy in choosing the correct object in the target pictures.

Procedure

The task was administered to participants in paper form. All participants were given two sample problems: one involving one relation and the other involving two relations. During the first sample problem, the experimenter gave the following instructions:

> Are you ready? We are going to play the picture game. Let me show you how it works. On every page there are two pictures like this. There is a certain pattern in the top picture, and the same pattern happens in the bottom picture, but it looks different. Let me show you what I mean on this page. See up in the top picture, there is a bigger boy and a smaller boy. This is the bigger boy, and this is the smaller boy [the experimenter pointed to each object as it was described]. Now in the bottom picture, there is a bigger bear and a smaller bear [the experimenter pointed]. See, the same pattern happens in both, but it looks different.
>
> Now, in this game, first you have to figure out what the pattern is that happens in both pictures. Okay? Then I am going to point to one thing in the top picture, and your job is to tell me what is in the same part of the pattern in the bottom picture. So on this first page, if we have a smaller boy and a bigger boy, and a smaller bear and a bigger bear, if I point to the smaller boy, which one is like the smaller boy in the bottom picture? Which one is in the same part of the pattern in the bottom picture? [the experimenter pointed to each object as it was described].

If the child responded correctly, the experimenter gave feedback and then moved to the next sample problem. If the child responded incorrectly, the experimenter gave feedback and then repeated the description of the relational objects in the top and bottom pictures. The experimenter then asked the question again. If the child again gave an incorrect answer, the experimenter pointed out the correct answer (the smaller bear) and moved to the next sample problem. The following instructions for the second sample problem cued children to the possibility that they sometimes would be required to attend to additional levels of relational complexity:

> Now sometimes the pattern will have two parts, like the one you just saw with the bigger boy and the smaller boy, and sometimes the pattern will have three parts. Let me show you what I mean. In this top picture, there is a mom reading to a girl, who is reading to a teddy bear [the experimenter pointed to each object]. Then in the bottom picture, there is a dad reading to a boy, who is reading to a doll. See, the pattern is the same in both pictures, but it looks different. Now, if I point to this girl, you can see that she has someone reading to her and she is reading to someone else. She has two things happening to her. Now, if I point to this girl, who is like her in the bottom picture? What is in the same part of the pattern in the bottom picture? [the experimenter pointed to each object as it was described].

If the child answered correctly, the experimenter responded with "Good job, perfect because this boy is the only one that both has someone reading to him and is reading to someone else. Great, let's do some more." If the child answered incorrectly, the experimenter gave feedback and then repeated the preceding instructions beginning with the description of the pattern. If the child's answer was still incorrect, the experimenter continued with this cycle a third time and then gave the answer and went on to the experimental problems.

The problems were presented in random order following the sample problems. The task was administered to the 13- and 14-year-old participants in groups; all other children were tested individually by a single experimenter. On each page, the experimenter pointed to the object with the arrow in the top picture and asked, "What is like the ____ in the bottom picture?"

If the child refused to provide an answer to a problem, the experimenter repeated the question; if the child refused repeatedly, the experimenter moved on to the next page. If more than five pages were left blank, the child's data were excluded. The data for seven 3- and 4-year-olds and two 13- and 14-year-olds were excluded for this reason, primarily because they failed to finish the task. Also, if a child failed on both sample problems with repeated opportunities and failed on the first five problems, indicating that he or she did not understand the relational instructions, that child's data were excluded. The data for four 3- and 4-year-olds and three 6- and 7-year-olds were excluded for this reason. Some of these children were later identified as having limited proficiency in English or as having attention deficit disorder.

Results

Pattern of relational responses

Fig. 2 presents the proportions of correct relational responses for each of the four picture conditions as a function of age. A 2 (Distraction) × 2 (Relational Complexity) within-subjects analysis of variance (ANOVA) was conducted with age (three levels) as a between-subjects variable. The analysis revealed main effects of age, $F(2,65) = 78.15$, $p < .001$, $h_p^2 = .71$; distraction, $F(1,65) = 26.07$, $p < .001$, $h_p^2 = .29$; and relational complexity, $F(1,65) = 24.83$, $p < .001$, $h_p^2 = .02$.

Interactions were also examined among age, relational complexity, and distraction. The Age × Distraction condition was reliable, $F(2,65) = 3.15$, $p = .05$, $h_p^2 = .09$, whereas the

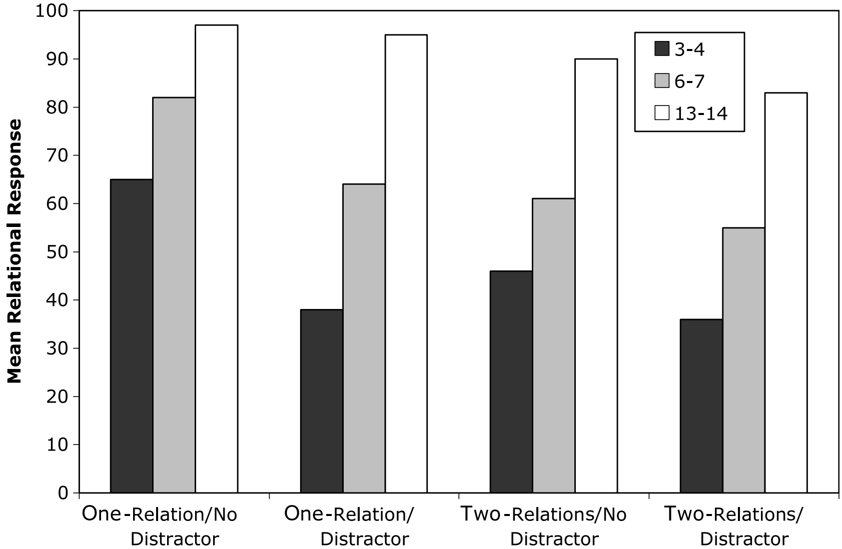

Fig. 2. Proportions correct relational responses as a function of distraction and number of relations across age groups (Experiment 1).

Age × Relational Complexity interaction was not, $F(2,65) = .57$, $p = .57$, $h^2_p = .02$. Importantly, the three-way interaction was reliable, $F(2,65) = 3.28$, $p < .05$, $h^2_p = .09$.

The pattern of three-way interaction was investigated using a 2 (Relational Complexity) × 2 (Distraction) repeated-measures ANOVA for each age group separately. Results show that for the youngest children (3- and 4-year-olds), there was a main effect of relational complexity, $F(1,21) = 4.44$, $p < .05$, $h^2_p = .18$; a main effect of distraction, $F(1,21) = 14.08$, $p < .01$, $h^2_p = .40$; and a reliable Relational Complexity × Distraction interaction, $F(1,21) = 4.21$, $p = .05$, $h^2_p = .17$. For the 6- and 7-year-olds, there was a main effect of relational complexity, $F(1,20) = 10.43$, $p < .01$, $h^2_p = .34$, and of distraction, $F(1,20) = 10.31$, $p < .01$, $h^2_p = .34$, but there was only a trend toward the Relational Complexity × Distraction interaction, $F(1,20) = 2.71$, $p = .12$, $h^2_p = .12$. Data for the 13- and 14-year-olds revealed a main effect of relational complexity, $F(1,24) = 17.66$, $p < .001$, $h^2_p = .42$, but not of distraction, $F(1,24) = 2.21$, $p = .15$, $h^2_p = .08$, nor was there a reliable interaction, $F(1,24) = 1.67$, $p = .21$, $h^2_p = .07$.

These three patterns reveal changes with age. The 3- and 4-year-olds showed strong effects of both distraction and relational complexity that interacted to reveal the highest accuracy in the one-relation/no-distractor condition and the lowest accuracy in the two-relation/distractor condition. This pattern was similar for the 6- and 7-year-olds, with main effects of both relational complexity and distraction. In contrast, the 13- and 14-year-olds showed a main effect of relational complexity but no effect of distraction.

Chance level of performance was calculated conservatively as the percentage of trials a participant would be expected to select the correct relational match within the set of reasonable choices. Here, "reasonable choices" included relational errors and featural errors but not extraneous objects. Using these criteria, chance differed by condition, reflecting the differing numbers of potential reasonable errors, ranging from 50% (two relevant possible answers) for one relation/no distractor, to 33% (three possible answers) for one relation/distractor and two relations/no distractor, to 25% (four possible answers) for two relations/distractor. Paired t tests revealed that the youngest children were above chance for the two-relation/no-distractor condition, $t(21) = 2.71$, $p < .05$, and the two-relation/no-distractor condition, $t(21) = 2.43$, $p < .05$, indicating that the 3- and 4-year-olds were understanding the analogy task and were able to reliably reason on the basis of relational similarity. Children were not above chance for the one-relation/distractor condition, $t(21) = 1.10$, $p = .29$, and although they were above chance for the two-relation/distractor condition, $t(21) = 2.35$, $p < .05$, this may be misleading because the children's mean performance was not better numerically than in the one-relation/distractor condition. The 6- and 7-year-olds and the 13- and 14-year-olds were above chance on all conditions at the .05 level, revealing that children's performance in all conditions improved with age.

Error analysis

Children's responses were categorized into four types (Table 1). Reponses were coded as either (a) relationally correct, (b) relational errors (an object in the correct relation but wrong role), (c) featural errors (the featural match in distractor conditions or an unrelated object in the corresponding spatial location in no-distractor conditions), or (d) other errors.

Relational errors

An analysis was conducted to examine the effects of distraction and relational complexity on relational errors. A 2 (Distraction) × 2 (Relational Complexity) within-subjects ANOVA was conducted with age (three levels) as a between-subjects variable, using relational errors as the dependent variable. Note that there was one possible relational error choice in the one-relation conditions and two such possible error choices in the two-relation conditions. The ANOVA revealed main effects of age, $F(2, 65) = 23.41$, $p < .001$, $h_p^2 = .76$, and relational complexity, $F(1, 65) = 59.56$, $p < .001$, $h_p^2 = .48$. There was a trend

Table 1
Proportions of each response type across age and condition (Experiment 1)

	Age (years)	One relation/ no distractor	One relation/ distractor	Two relations/no distractor	Two relations/ distractor
Correct relational response	3–4	65	38	46	36
	6–7	82	64	61	55
	13–14	97	95	90	83
Featural errors	3–4	8	46	11	46
	6–7	0	25	4	27
	13–14	0	5	0	8
Relational errors	3–4	15	9	37	15
	6–7	13	7	34	18
	13–14	2	1	8	9
Other errors	3–4	10	5	3	4
	6–7	5	5	1	0
	13–14	2	0	2	0

toward an Age × Relational Complexity interaction, $F(2,65) = 2.74$, $p = .07$, $h_p^2 = .08$. As shown in Table 1, the younger children made more relational errors overall than did the 13- and 14-year-olds. The number of relational errors increased with an additional level of relational complexity, but this result was driven primarily by errors made in the two-relation/no-distractor condition.

Importantly, there was also a main effect of distraction on children's quantity of relational errors, $F(1,65) = 21.46$, $p < .001$, $h_p^2 = .25$, and a significant Distraction × Age interaction, $F(2,65) = 5.85$, $p < .01$, $h_p^2 = .15$. The 3- and 4-year-olds made relational errors more frequently when there was no featural distractor available as an option. The 13- and 14-year-olds made fewer relational errors overall. Interestingly, the pattern of errors in the 3- and 4-year-olds and the 6- and 7-year-olds suggests that the presence of a distractor in the target picture drew errors away from relational choices, although this was less true for the 6- and 7-year-olds than for the 3- and 4-year-olds.

Featural errors

A 2 (Distraction) × 2 (Relational Complexity) within-subjects ANOVA was conducted with age (three levels) as a between-subjects variable, using featural errors as a dependent variable, to examine the relation between age and participants' featural errors across the four picture conditions. Children's choice of the featural match on the distractor conditions was compared with their choice of a nonfeatural matched object in the same spatial location for the no-distractor conditions. The main effect of age was reliable, $F(2,65) = 49.78$, $p < .001$, $h_p^2 = .61$, as was the main effect of distraction, $F(1,65) = 126.54$, $p < .001$, $h_p^2 = .66$, confirming that the featural match was an effective lure. There was also a significant Age × Distraction interaction, $F(2,65) = 20.15$, $p < .001$, $h_p^2 = .38$, such that as age increased, the likelihood that a participant would make a featural error decreased. The youngest children were most likely to make featural errors, and the oldest children rarely made these mistakes. Importantly, there was no effect of relational complexity on featural errors, $F(2,65) = 1.54$, $p = .22$, $h_p^2 = .02$.

Combined with the relational error data above, these results reveal that young children's errors were systematically dependent on condition. As shown in Table 1, 3- and 4-year-olds and 6- and 7-year-olds made more featural errors with the addition of a distractor object (one-relation/no-distractor vs. one-relation/distractor condition), and all participants made more relational errors with an increase in relational complexity (one-relation/no-distractor vs. two-relation/no-distractor condition).

In the condition where both relational complexity and distraction were present and competitive (two-relation/distractor condition), children's errors were diagnostic of which manipulation was a stronger predictor of children's attention and patterns of performance. A one-way within-subjects ANOVA was conducted to compare featural errors on two-relation/distractor problems with relational errors on these problems, with age as a between-subjects variable. There was a significant effect of error type, $F(1,65) = 13.68$, $p < .001$, $h_p^2 = .17$. There was also a significant Age × Error Type interaction, $F(2,65) = 7.49$, $p = .001$, $h_p^2 = .19$. This pattern reveals that 3- and 4-year-olds were much more likely to make featural errors than relational errors in the two-relation/distractor condition, whereas the 13- and 14-year-olds made fewer errors overall, with approximately the same number of featural errors as relational errors in this condition.

Discussion

Data from the scene analogy problems at 3 and 4, 6 and 7, and 13 and 14 years of age provide insight into the roles of relational knowledge, object similarity, and relational complexity in children's development of analogical reasoning. Patterns in participants' correct relational responses revealed main effects of age, distraction, and relational complexity, supporting the validity of the task manipulations. These main effects have several implications. First, children's above-chance performance on the one-relation/no-distractor condition provides support for the hypothesis that children can attend to and map relations by 3 or 4 years of age. The performance of 3- and 4-year-olds in this condition was 64% accurate. Although reliably above chance, this level of accuracy was not extremely high. It should be noted, however, that 15% of these children's errors for this condition were relational errors. This error pattern indicates that children were attending to the relevant relational structure but that a significant proportion of their errors were relational failures to maintain structure.

In addition, these main effects provide support for claims that children can attend to and map relations (as predicted by the relational primacy hypothesis) but are not fully able to avoid misleading object similarity or to maintain relational structure when an additional level of relational complexity is imposed.

All age groups made more relational errors in the two-relation/no-distractor condition than in the one-relation/no-distractor condition, demonstrating that additional levels of relational complexity made analogical processing more difficult, but this was particularly true for the youngest children. These results also indicate that the methodological concern that children might have used a strategy of selecting the middle relational object in the two-relation/no-distractor problems is not likely because that would have predicted a lower level of complexity, and thus higher performance, for the two-relation/no-distractor problems than for the one-relation/no-distractor problems.

Children also showed developmental patterns in their responses to the featural distraction manipulation. The youngest children made featural errors most often when there was a featural distractor present, regardless of the relational complexity of the analogy and in spite of accuracy when there was no featural distractor present. As increasingly older children were tested, they were less likely to select a featural distractor and made fewer errors overall.

The analysis of children's errors in the two-relation/distractor condition provides additional information about the nature of featural distraction and relational complexity within the development of analogical reasoning capacity. The interaction between age and error type made by children in the two-relation/distractor condition revealed that the youngest children made primarily featural errors, whereas the 13- and 14-year-olds made fewer errors overall, with as many featural errors as relational errors. Thus, when a match based on object similarity was present, 3- and 4-year-olds seemed to have limited resources to resist this lure, even though they reliably demonstrated accurate relational mapping in conditions without a featural distractor.

Although relational complexity did limit the youngest children's analogical reasoning, error patterns in the two-relation/distractor condition suggest that there was a developmental pattern such that the youngest children were more drawn to featural errors than to relational errors when both were feasible. Although this difference may reflect the relative strength of the experimental manipulations of relational complexity and featural distraction, it is important to note that this pattern disappeared for the older children. It is possible that the youngest children may have noticed and become distracted by the featural distractor before they attended to the relational structure of the pictures. It follows that as children become more able to resist featural distraction, they may actually become more vulnerable to making relational errors on problems that contain both types of difficulty. These error findings are consistent with evidence from a study of analogical reasoning in patients with damage to the prefrontal cortex (Morrison et al., 2004, Experiment 1). In that study, patients with deficits in working memory and inhibitory control showed a marked preference for featural responses over relational responses in solving the Markman and Gentner (1993) picture analogy problems.

Experiment 2

Because children's performance depended on their preliminary abstraction of relations in the sets of pictures we provided, we wanted to ensure that participants in our study were not making errors due to an inability to determine the relevant relation for comparison or due to failures to understand the relations. In Experiment 2, we used the same method but explicitly verbalized to participants the relational verb corresponding to the relevant structure for aligning the two pictures. The second change introduced in Experiment 2 was in the age groups tested. The youngest age group (3- and 4-year-olds) was tested again, but a 9- to 11-year-old group replaced the 6- and 7-year-old and 13- and 14-year-old groups. This change was made because there seemed to be a large change in performance between the 6- and 7-year-old and 13- and 14-year-old groups in Experiment 1, and we were interested in identifying more specifically the earliest age at which the effects of relational complexity and distraction are minimized on this task.

Two control conditions were also added in Experiment 2. The first was a relational knowledge control designed to ensure that children had knowledge of the 20 relations used in the scene analogy problems. The second was an instructional control conducted to verify that features of the language used in the task instructional prompts did not drive the youngest children's error patterns.

Method

Participants

The participants were 44 children (20 3- and 4-year-olds and 24 9- to 11-year-olds). These were boys and girls enrolled in preschool and elementary programs in the Los Angeles area. Although demographic data were not collected systematically, participating children were primarily from middle-class neighborhoods and were of diverse ethnic backgrounds, including Caucasian, Asian, Asian American, and Latin American descent.

Materials and design

Primary task and instructional control

The materials for the primary task and the instructional control were identical to those used in Experiment 1. The scene analogy problems were used with all participants in a 2×2 within-subjects design. Participants saw 20 picture sets in which the level of relational complexity and the presence or absence of a featural distractor were manipulated. Each participant saw five picture sets in each condition: one relation/distractor, one relation/no distractor, two relations/distractor, and two relations/no distractor. The content of the picture sets was counterbalanced across four versions as in Experiment 1.

Relational knowledge control

To ensure that 3- and 4-year-olds knew the relations and the relational words used to describe the source relations in the 20 picture pairs used in Experiment 2, a control task was designed to require children to identify a depiction of each relation. Children were given a relation and asked to identify which of two pictures showed that relation. Specifically, a task was constructed to require children to choose between two picture alternatives. Two pictures were aligned on each page with a relation written above them. The relation was not intended to be read by the children; rather, it was to be read by the experimenter.

All of the forced-choice alternatives were two-relation/distractor target pictures, which were the pictures that led to the most errors in Experiment 1. A total of 20 pairs were constructed so that each participant made judgments about all 20 relations. On a page, there was always one picture that depicted the spoken relation and one picture that was determined by undergraduates to not depict the relation. The placement of the correct picture was varied randomly between the left and right sides. Packets were assembled in which relations were sequenced in random order. Children who participated in the relational control did not participate in the experimental groups.

Procedure

Primary task

The main scene analogy problems were administered to participants in paper form. Two training problems were administered as in Experiment 1. On the test problems (presented in random order following the training problems), the experimenter verbalized the relevant relation in the source picture. The experimenter described the objects engaged in the relevant relation in the source picture, moving from the object that was only an agent of the relation (agent), to the object that was both an agent and a patient of the relation (agent–patient), to the object that was only a patient of the relation (patient). For example, in the "chasing" picture set illustrated in Figs. 1C and D, the experimenter would say, "Look, here is a dog chasing a cat that is chasing a mouse. What is like the cat in the bottom picture?"

Instructional control

A total of 12 3- and 4-year-olds participated in a control version in which several changes were made to the instructions described previously. The first change was that instead of using the word "pattern" in the instructions during the sample pictures, the experimenter said "is happening." The alternative instructions used were as follows:

> Are you ready? We are going to play the picture game. Let me show you how it works. On every page, there are two pictures like this. There is a certain thing that happens in the top picture, and the same thing happens in the bottom picture, but it looks different. Let me show you what I mean on this page. See up in the top picture,

> there is a bigger boy and a smaller boy. This is the bigger boy, and this is the smaller boy [experimenter pointed while saying this part]. Now in the bottom picture, there is a bigger bear and a smaller bear [experimenter pointed]. See, the same thing happens in both, but it looks different.
>
> Now, in this game, first you have to figure out what the thing is that happens in both pictures. Okay? Then I am going put a blue sticker on one thing in the top picture, and your job is to tell me what has the same thing happening in the bottom picture. We will put a yellow sticker on the one you point to. So on this first page, if we have a smaller boy and a bigger boy, and a smaller bear and a bigger bear, if I put a blue sticker on the smaller boy, which one has the same thing happening in the bottom picture? Where should I put a yellow sticker in the bottom picture? [the experimenter pointed at each object as it was described].

The same transitional instructions were given on the second sample problem, but the word "pattern" was replaced by "happening" and the sticker instructions were repeated:

> Now sometimes what is happening will have two parts, like the one you just saw with the bigger boy and the smaller boy, and sometimes there will be three parts to what is happening. Let me show you what I mean. In this top picture, there is a mom reading to a girl, who is reading to a teddy bear [experimenter pointed to each object]. Then in the bottom picture, there is a dad reading to a boy, who is reading to a doll. See, what is happening is the same in both pictures, but it looks different. Now, if I put a blue sticker on this girl, you can see that she has someone reading to her and she is reading to someone else. She has two things happening to her. Now, if I put the blue sticker on this girl, who has the same thing happening in the bottom picture? Where should I put the yellow sticker in the bottom picture?

The instructions were repeated if necessary to ensure the child's understanding. The second change was that instead of using the prompt "What is like the ___ in the bottom picture?" for each new picture set, the experimenter used a prompt based on sticker placement. For each picture set, the experimenter used the prompt "If I put a blue sticker on the ___ in the top picture, where should I put a yellow sticker in the bottom picture?" This prompt removes the word "like" or other linguistic comparative terms that could be interpreted by the child as indicating featural similarity. During the sample problems, the relations and relational complexity manipulation were highlighted as in Experiment 1 to ensure that the children were aware that they should be attending to the relational properties of the objects highlighted with the stickers.

Relational knowledge control

Children tested in the relational knowledge control did not participate in any of the experimental conditions. Each child was tested individually with an experimenter. The experimenter gave the instruction "Which picture shows ___ing?" for each page of the packet. If the verb required an object, an indefinite term was used. For example, for the relation *inside*, the experimenter asked, "Which picture shows something inside something else?" On the first two pages, the experimenter used a finger to underline the two pictures after asking the identification question. The order of pointing was varied so that if the left picture was pointed to first on the first page, the right picture was pointed to first on the second page. The reverse order was also used. While pointing, the experimenter asked, "This picture or this picture?" After the first two pages, this procedure was used only if the child failed to attend to the task or gave an unclear response.

Children were asked to point to the correct picture, and they did so primarily, but pointing or verbal answers were accepted as responses.

Results

Pattern of relational responses in experimental conditions

Fig. 3 presents the proportions of correct relational responses for each of the four picture conditions as a function of age. A 2 (Relational Complexity) × 2 (Distraction) within-subjects ANOVA was performed with age (two levels) as a between-subjects variable. Children's correct relational choices were used as the dependent variable. The ANOVA revealed main effects of age, $F(1,42) = 154.85$, $p < .001$, $h_p^2 = .79$; relational complexity, $F(1,42) = 6.38$, $p < .001$, $h_p^2 = .13$; and distraction, $F(1,42) = 34.32$, $p < .001$, $h_p^2 = .45$. These results establish that even though the critical relation was verbalized by the experimenter to minimize any variability due to failures to abstract the relevant relation, the scene analogy problems remain sensitive to age, the picture manipulations were effective at creating distraction and increasing relational complexity, and 3- and 4-year-olds' analogical reasoning was not robust to these challenges.

The Age × Distraction interaction was reliable, $F(1,42) = 11.20$, $p < .01$, and the Age × Relational Complexity interaction was also reliable, $F(1,42) = 9.29$, $p < .01$. Importantly, the three-way interaction was reliable as well, $F(1,42) = 4.98$, $p < .05$.

The pattern of interaction was investigated using repeated-measures ANOVAs for each age group separately. These analyses revealed that for the youngest children (3- and

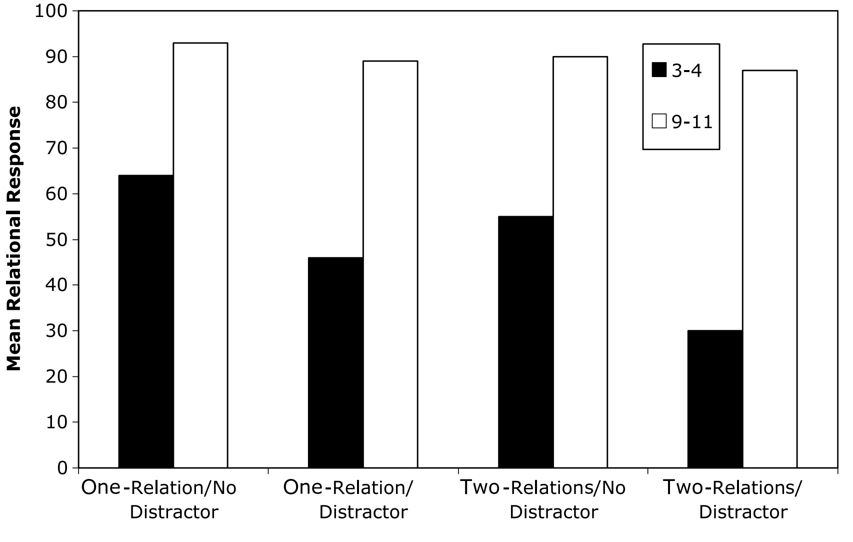

Fig. 3. Proportions correct relational responses as a function of distraction and number of relations across age groups (Experiment 2).

4-year-olds), there was a main effect of relational complexity, $F(1,19) = 4.30$, $p = .05$, and a main effect of distractor, $F(1,19) = 28.69$, $p < .01$. The interaction between relational complexity and distraction was not reliable, $F(1,19) = .71$, $p = .41$. Data for the 9- to 11-year-olds revealed no reliable differences due to either relational complexity, $F(1,19) = 1.43$, $p = .25$, or distraction, $F(1,19) = 3.29$, $p = .08$, nor was the interaction reliable, $F(1,19) = 0.04$, $p = .84$.

Paired t tests were used to test whether the young children's performance was above chance levels, defined as in Experiment 1. These analyses revealed that 3- and 4-year-olds were not above chance in the two-relation/distractor condition, $t(19) = 1.26$, $p = .22$, but they did make accurate relational judgments above chance in all other conditions: one relation/no distractor, $t(19) = 2.59$, $p < .05$; one relation/distractor, $t(19) = 2.23$, $p < .05$; and two relations/no distractor, $t(19) = 4.84$, $p < .001$. Thus, these data suggest that the children were clear on the instructions and were fairly successful on the task. However, as shown by error data and relatively low performance in the two-relation/distractor condition, their performance was impaired in predictable ways. The older children performed above chance in all conditions.

Error analysis

Children's responses were categorized into four types as in Experiment 1 (Table 2).

Featural errors

An ANOVA was performed to examine the relation between age and participants' featural errors across the four picture conditions. Children's choice of the featural match on the distractor conditions was compared with their choice of a nonfeatural matched object in the same spatial location in the no-distractor conditions. A 2 (Distraction) × 2 (Relational Complexity) repeated-measures ANOVA was performed with age (two levels) as a between-subjects factor on featural errors. The main effect of age was reliable, $F(1,42) = 105.5$, $p < .001$, $h_p^2 = .72$, as was the main effect of distraction, $F(1,42) = 112$, $p < .001$, $h_p^2 = .73$, confirming that the featural match was an effective distractor. The Age × Distraction interaction was also significant, $F(1,42) = 95.2$, $p < .001$, $h_p^2 = .69$, supporting the hypothesis that featural distraction is a developmental limiting factor on

Table 2
Proportions of each response type across age and condition (Experiment 2)

	Age (years)	One relation/ no distractor	One relation/ distractor	Two relations/ no distractor	Two relations/ distractor
Correct relational response	3–4	66	46	55	30
	9–11	93	89	90	87
Featural errors	3–4	6	42	2	50
	9–11	0	3	1	2
Relational errors	3–4	15	9	40	19
	9–11	7	8	9	12
Other errors	3–4	3	2	3	1
	9–11	0	0	0	0

analogical reasoning. Finally, there was a reliable three-way Distraction × Relational Complexity × Age interaction, $F(1, 42) = 4.64$, $p < .05$, $h_p^2 = .10$.

Relational errors

A 2 (Distraction) × 2 (Relational Complexity) repeated-measures ANOVA was performed on relational errors with age (two levels) as a between-subjects factor. Note that there was one possible relational error choice in the one-relation conditions and two such possible error choices in the two-relation conditions. The ANOVA revealed main effects of age, $F(1, 42) = 16.52$, $p < .001$, $h_p^2 = .28$, as well as relational complexity, $F(1, 42) = 19.97$, $p < .001$, $h_p^2 = .32$. Interestingly, there was also a significant main effect of distraction on relational errors, $F(1, 42) = 6.97$, $p < .05$, $h_p^2 = .14$, as well as a Distraction × Age interaction, $F(1, 42) = 11.20$, $p < .01$, $h_p^2 = .21$. As in Experiment 1, at younger ages children made relational errors more frequently when there was no featural distractor available as an option. As shown in Table 2, 3- and 4-year-olds and 6- and 7-year-olds made more featural errors with the addition of a distractor object (one-relation/no-distractor vs. one-relation/distractor condition), and all participants made more relational errors with an increase in relational complexity (one-relation/no-distractor vs. two-relation/no distractor condition).

Types of children's errors were assessed for the two-relation/distractor condition, in which relational complexity and distraction were in competition with each other as sources of young children's difficulty. A one-way repeated-measures ANOVA was performed to compare featural errors with relational errors in this condition. Age was included as a between-subjects variable. As in Experiment 1, there was a significant effect of error type, $F(1, 42) = 4.91$, $p < .05$, $h_p^2 = .11$. There was also a significant Age × Error Type interaction, $F(2, 42) = 18.7$, $p < .001$, $h_p^2 = .31$. This pattern indicates that 3- and 4-year-olds were much more likely to make featural distractor errors than to make relational errors in the two-relation/distractor condition, whereas 9- to 11-year-olds made fewer errors overall, with more relational errors than featural errors.

Instructional control

The data for 3- and 4-year-olds in the instructional control condition were compared with those for 3- and 4-year-olds in the experimental condition. A 2 (Relational Complexity) × 2 (Distraction) within-subjects ANOVA was performed with a between-subjects variable of instruction (two levels). In this combined analysis, there were main effects of distraction, $F(1, 30) = 30.67$, $p < .001$, $h_p^2 = .51$, and relational complexity, $F(1, 30) = 4.93$, $p < .05$, $h_p^2 = .14$. Importantly, however, there was no main effect of instruction, $F(1, 30) = 1.51$, $p = .23$, $h_p^2 = .05$, and no significant interactions with instruction for either distraction, $F(1, 30) = 0.22$, $p = .64$, $h_p^2 = .01$, or relational complexity, $F(1, 30) = 0.15$, $p = .70$, $h_p^2 = .01$.

In addition, as shown in Table 3, error data resemble patterns made by children in the experimental condition. These data suggest that children's errors in the experimental condition were driven not by the task instructions but rather by the distraction and relational complexity manipulations themselves.

Relational knowledge control

Data for the relational knowledge control were coded using binary codes of correct (1) and incorrect (0). A total of 15 children participated in the relational knowledge control,

Table 3
Proportions of each response type across age and condition (instructional control)

	Age (years)	One relation/no distractor	One relation/ distractor	Two relations/no distractor	Two relations/ distractor
Correct relational response	3–4	60	35	44	33
Featural errors	3–4	13	27	20	35
Relational errors	3–4	10	20	34	27
Other errors	3–4	17	18	0	5

with 1 being excluded for limited proficiency in English. Children's performance overall was 90% correct over the 20 relations, indicating that 3- and 4-year-olds were familiar with and knowledgeable about the relations used in the picture task. The task instructions required that children identify the relations in the two-relation/distractor condition. Therefore, the results indicate that in this most difficult condition, children were able to identify the relevant relations.

Discussion

The results of Experiment 2 were consistent with those of Experiment 1. There was a main effect of age on accuracy of participants' responses as well as main effects of relational complexity and distraction. These findings demonstrate that the manipulations of relational complexity and distraction limited children's performance despite an experimenter providing participants with an explicit verbalization of the relevant relations. This indicates that children's failure to understand the pictures could not drive the effects. These data support the hypothesis that both relational complexity and featural distraction limit children's analogical reasoning capacity.

In Experiment 2, as in Experiment 1, children's accuracy on the one-relation/no-distractor condition revealed their capability to be successful on the basic analogy task and their comprehension of the critical relations. The 3- and 4-year-olds were correct 64% of the time on picture sets in the one-relation/no-distractor condition, more than double their performance once an additional level of relational complexity was imposed and a featural distractor was added (30% correct). In the one-relation/no-distractor condition, 3- and 4-year-olds made relational errors 15% of the time, suggesting that children were attempting to map on the basis of the relevant relations in many of the cases where they were not correct. Because most of the developmental analogy studies have used formal problems in the format "A is to B as C is to D" in which the child selects an object for D, this type of failure to maintain correct relational roles has not been identified previously.

The three-way interaction among age, distraction, and relational complexity indicated a pattern of results consistent with the hypothesis that age-related maturational factors constrain analogical reasoning in spite of children's knowledge of the relations. The 3- and 4-year-olds were strongly affected by relational complexity and featural distraction, such that their performance fell toward chance with the imposition of a second level of relational complexity or the presence of a featural distractor.

The 9- to 11-year-olds were less affected by either relational complexity or distraction on this task and overall were highly accurate on the task in all conditions. In fact, whereas the 13- and 14-year-olds in Experiment 1 showed an effect of relational complexity on accuracy, the 9- to 11-year-olds in Experiment 2 did not. There may be two reasons for this difference. The first is that the 13- and 14-year-olds were tested in a group, whereas the 9- to 11-year-olds were tested individually with an experimenter. As a result, motivation levels may have been higher for the 9- to 11-year-olds than for the 13- and 14-year-olds. There was also an academic difference between the participant groups, such that the 9- to 11-year-olds attended an academically higher achieving school than did the 13- and 14-year-olds, suggesting that the older group may have had less experience with analogy tasks or more generally with abstract reasoning tasks.

An analysis of children's errors allowed a closer examination of children's performance patterns across the four conditions. First, participants of all ages were more likely to select the featural match than to select a nonsimilar object in the same spatial location, validating the distraction manipulation within the task. This pattern interacted with age, such that the 3- and 4-year-olds were more likely to make a featural error than were the 9- to 11-year-olds. Second, relational errors were made regularly by both 3- and 4-year-olds and 9- to 11-year-olds, even after hearing the relations verbalized for the source pictures. Children in the younger age group made these errors more frequently than did children in the older age group. In particular, 3- and 4-year-olds made relational errors most often in the two-relation/no-distractor condition, providing evidence that relational complexity is an important constraint on young children's analogical reasoning. Third, despite directions to select a relational match and explicit verbalization of the relevant relations in each picture set, 3- and 4-year-olds were more likely to make featural errors than to make relational errors when both error types were available as choices. This result was obtained even when there were two levels of relational complexity as well as a distractor present. These results replicate the pattern found in Experiment 1.

The control conditions in Experiment 2 provided evidence that even children as young as 3 or 4 years were familiar and knowledgeable about the relations used in the scene analogy problems. Despite knowledge of the relevant relations, these young participants were often unable to avoid making featural or relational errors. Furthermore, the instructional control supported the claim that participants' patterns of performance and errors were not driven by the instructions.

In summary, data from Experiment 2 replicated the basic findings of Experiment 1 and lend further support to the hypothesis that the difficulty encountered by young children on the task is due not to difficulty with interpreting the relations depicted in the pictures but rather to the demands of relational complexity and avoiding featural distraction.

General discussion

Data from two experiments support the importance of both relational complexity and featural distraction as factors in children's development of analogical reasoning and suggest the developmental progression of how these processing limits interact to constrain analogical reasoning. These data go beyond prior studies in which relational complexity and featural distraction were observed separately to impair children's analogical reasoning.

The current findings indicate that accretion of relational knowledge, although doubtless necessary for analogical development, cannot fully explain children's featural errors on the scene analogy task. If knowledge of relations were the only mechanism critical to analogy performance, children's performance should have been equal across all conditions (because the critical relation was equated). Instead, the performance of the 3- and 4-year-olds was much more accurate in the one-relation/no-distractor condition than in any of the other three conditions. The youngest children accurately selected a relational object (the correct analog or a relational error) when there was no distractor present (one relation/no distractor and two relations/no distractor) but reliably made featural errors when there was a distractor present (one relation/distractor and two relations/distractor). Furthermore, the control for relational knowledge in Experiment 2 demonstrated that 3- and 4-year-olds know and are able to identify the relevant relations in the pictures used.

It is nonetheless possible that relational knowledge affected children's performance in more subtle ways. It is certainly feasible that young children may have experienced some difficulty in interpreting the relation and in understanding the role of each object in the relations when they selected an item from one picture to map to another item in another picture. Nonetheless, the relations used in the scene analogy problems are based on verbs of the type known to be familiar to children by 3 or 4 years of age (Golinkoff et al., 1995, 1996); thus, describing the relations verbally (Experiment 2) should have minimized variability in their performance due to failure to understand the pictures. Instead, even with the aid of the experimenter's verbalization, the youngest children performed poorly on two-relation problems even when no distractor was present.

The current findings fit more closely with the hypothesis that maturational limitations constrain children's performance on analogical reasoning tasks. The strength of the featural distraction manipulations can be explained by maturational patterns in children's inhibitory control. Even the youngest children understood the task and selected a relational match regularly when no additional featural distractions were imposed, as demonstrated by their relatively high level of accuracy on the one-relation/no-distractor picture sets. However, when a featural distractor was available, young children were very likely to select that object. This tendency diminished with age until the 9- to 11-year-olds and 13- and 14-year-olds made few featural errors. The current data support the existence of the phenomenon identified by Gentner and Rattermann (1991) as the relational shift, but our findings suggest that knowledge acquisition is not the sole mechanism underlying the developmental progression.

Data from both experiments also support prior studies indicating that children's ability to reason analogically at higher levels of relational complexity develops with age and related increases in working memory capacity (Halford, 1993). In the neurally plausible computational account of analogical reasoning, the LISA model (Hummel & Holyoak, 1997, 2003), analogical mapping depends critically on the systematic activation and inhibition of information in working memory. Studies of analogical reasoning based on populations with compromised working memory functions, including patients with prefrontal damage (Morrison et al., 2004) and older adults (Viskontas et al., 2004), have provided evidence that both adequate relational knowledge and intact working memory are necessary for the full capacity to reason analogically. As in the current study, Viskontas and colleagues (2004) found a three-way interaction among age, relational complexity, and distraction in analogical reasoning across the age range from young to older adults. This interaction was modeled by LISA as a change in inhibitory control.

The relation between the two sources for failure in children's development of analogical reasoning has not been examined systematically previously. The theoretical relationship between relational complexity and featural distraction remains to be established. Andrews, Halford, Bunch, Bowden, and Jones (2003) proposed that some effects attributed to inhibitory control could be explained by relational complexity. Alternatively, inhibitory control may be a separable factor that operates in addition to the capacity to represent and integrate multiple relations.

In sum, the results of the current study help to clarify the relations among alternative hypotheses regarding the development of children's analogical reasoning. By examining the relations between these constraints, we were able to separate the roles of relational complexity and featural distraction from deficits in domain knowledge. Our findings reveal that young children have difficulties in analogical reasoning due to featural distraction and relational complexity even when they understand the critical relations. This insight is important, because analogical reasoning provides a basic cognitive tool that children may use to approach novel phenomena and transfer across contexts. Revealing the mechanisms underlying development of this capacity provides a window into children's broader cognitive development.

Acknowledgments

The authors thank the Spencer Foundation (dissertation fellowship to L. E. Richland), the National Institute of Mental Health (NRSA fellowship MH-64244-01A1 to R. G. Morrison), Xunesis, and the Institute of Education Sciences (R305H030141 to K. J. Holyoak) for their generous support. We also thank Ann Fink for drawing the pictures used in the scene analogy problems. The scene analogy problems are available from the first author on request. A preliminary report of Experiment 1 was presented at the 2004 meeting of the Cognitive Science Society in Chicago.

References

Andrews, G., & Halford, G. S. (2002). A cognitive complexity metric applied to cognitive development. *Cognitive Psychology, 45*, 153–219.

Andrews, G., Halford, G. S., Bunch, K. M., Bowden, D., & Jones, T. (2003). Theory of mind and relational complexity. *Child Development, 74*, 1476–1499.

Chen, Z., Sanchez, R., & Campbell, T. (1997). From beyond to within their grasp: Analogical problem solving in 10- and 13-month-olds. *Developmental Psychology, 33*, 790–801.

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J. K., Holyoak, K. J., et al. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *NeuroImage, 14*, 1136–1149.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 466–503). London: Oxford University Press.

Diamond, A., Kirkham, N., & Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Developmental Psychology, 38*, 352–362.

Frye, D., & Zelazo, P. D. (1998). Complexity: From formal analysis to final action. *Behavioral and Brain Sciences, 21*, 836–837.

Frye, D., Zelazo, P. D., Brooks, P. J., & Samuels, M. C. (1996). Inference and action in early causal reasoning. *Developmental Psychology, 32*, 120–131.

Gentner, D. (1977). 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. (1978). On relational meaning: The acquisition of verb meaning. *Child Development, 49*, 988–998.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155–170.

Gentner, D. (1988). Analogical inference and analogical access. In A. Prieditis (Ed.), *Analogica* (pp. 63–88). Los Altos, CA: Morgan Kaufmann.

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., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10*, 277–300.

Gick, M. L., & Holyoak, K. L. (1980). Analogical problem solving. *Cognitive Psychology, 15*, 306–355.

Golinkoff, R. M., Hirsh-Pasek, K., Mervis, C. B., Frawley, W., & Parillo, M. (1995). Lexical principles can be extended to the acquisition of verbs. In M. Tomasello & W. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs* (pp. 185–222). Hillsdale, NJ: Lawrence Erlbaum.

Golinkoff, R. M., Jacquet, R., Hirsh-Pasek, K., & Nandakumar, R. (1996). Lexical principles may underlie the learning of verbs. *Child Development, 67*, 3101–3119.

Goswami, U. (1992). *Analogical reasoning in children*. Hillsdale, NJ: Lawrence Erlbaum.

Goswami, U. (2001). Analogical reasoning in children. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 437–470). Cambridge, MA: MIT Press.

Goswami, U., & Brown, A. L. (1989). Melting chocolate and melting snowmen: Analogical reasoning and causal relations. *Cognition, 35*, 69–95.

Goswami, U., Leevers, H., Pressley, S., & Wheelwright, S. (1998). Causal reasoning about pairs of relations and analogical reasoning in young children. *British Journal of Developmental Psychology, 16*, 553–569.

Halford, G. S. (1993). *Children's understanding: The development of mental models*. Hillsdale, NJ: Lawrence Erlbaum.

Halford, G. S., Andrews, G., Dalton, C., Boag, C., & Zielinski, T. (2002). Young children's performance on the balance scale: The influence of relational complexity. *Journal of Experimental Child Psychology, 81*, 383–416.

Holyoak, K. J., Junn, E. N., & Billman, D. (1984). Development of analogical problem-solving skill. *Child Development, 55*, 2042–2055.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104*, 427–466.

Hummel, J. E., & Holyoak, K. J. (2003). A symbolic–connectionist theory of relational inference and generalization. *Psychological Review, 110*, 220–264.

Kroger, J. K., Sabb, F. W., Fales, C. L., Bookheimer, S. Y., Cohen, M. S., & Holyoak, K. J. (2002). Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: A parametric study of relational complexity. *Cerebral Cortex, 12*, 477–485.

Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology, 25*, 431–467.

Morrison, R. G. (2005). Thinking in working memory. In K. J. Holyoak & R. G. Morrison (Eds.), *Cambridge handbook of thinking and reasoning* (pp. 457–473). Cambridge, UK: Cambridge University Press.

Morrison, R. G., Holyoak, K. J., & Truong, B. (2001). Working memory modularity in analogical reasoning. In *Proceedings of the twenty-fourth annual conference of the cognitive science society* (pp. 663–668). Mahwah, NJ: Lawrence Erlbaum.

Morrison, R. G., Krawczyk, D., Holyoak, K. J., Hummel, J. E., Chow, T., Miller, B., et al. (2004). A neurocomputational model of analogical reasoning and its breakdown in frontotemporal dementia. *Journal of Cognitive Neuroscience, 16*, 260–271.

Piaget, J., Montangero, J., & Billeter, J. (1977). La formation des correlats. In J. Piaget (Ed.), *Recherches sur l'abstraction reflechissante I* (pp. 115–129). Paris: Presses Universitaires de France.

Prabhakaran, V., Smith, J. A. L., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1997). Neural substrates of fluid reasoning: An fMRI study of neocortical activation during performance of the Raven's Progressive Matrices Test. *Cognitive Psychology, 33*, 43–63.

Rattermann, M. J., & Gentner, D. (1998). More evidence for a relational shift in the development of analogy: Children's performance on a causal-mapping task. *Cognitive Development, 13*, 453–478.

Viskontas, I. V., Morrison, R. G., Holyoak, K. J., Hummel, J. E., & Knowlton, B. J. (2004). Relational integration, inhibition, and analogical reasoning in older adults. *Psychology and Aging, 19*, 581–591.

Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Mishkin, F. S., de Menezes Santos, M., et al. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological Science, 10*, 119–125.

Waltz, J. A., Lau, A., Grewal, S. K., & Holyoak, K. J. (2000). The role of working memory in analogical mapping. *Memory & Cognition, 28*, 1205–1212.
Zelazo, P. D., & Frye, D. (1998). Cognitive complexity and control: The development of executive function. *Current Directions in Psychological Science, 7*, 121–126.
Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development, 11*, 37–63.
Zelazo, P. D., & Müller, U. (2002). The balance beam in the balance: Reflections on rules, relational complexity, and developmental processes. *Journal of Experimental Child Psychology, 81*, 458–465.