Developmental Psychology
2013, Vol. 49, No. 5, 848–860

0012-1649/13/$12.00 DOI: 10.1037/a0028891

The Relationship Between Memory and Inductive Reasoning: Does It Develop?

Brett K. Hayes and Kristina Fritz
University of New South Wales

Evan Heit
University of California, Merced

In 2 studies, the authors examined the development of the relationship between inductive reasoning and visual recognition memory. In both studies, 5– to 6-year-old children and adults were shown instances of a basic-level category (dogs) followed by a test set containing old and new category members that varied in their similarity to study items. Participants were given either recognition instructions (memorize study items and discriminate between old and new test items) or induction instructions (learn about a novel property shared by the study items and decide whether it generalizes to test items). Across both tasks, children made a greater number of positive responses than did adults. Across both age groups, a greater number of positive responses were made in induction than in recognition. The application of a mathematical model, called GEN-EX for generalization from examples, showed that both memory and reasoning data could be explained by a single exemplar-based process that assumes task and age differences in generalization gradients. These results show considerable developmental continuity in the cognitive processes that underlie memory and inductive reasoning.

Keywords: inductive reasoning, recognition memory, quantitative modeling

Although there is little doubt that memory and reasoning are central to cognitive development, in previous work, they have generally been treated as discrete components. Each has been studied with different experimental paradigms, with the result that developmental change in reasoning and memory is currently addressed by separate theories, and accounts linking the development of these processes are rare (for exceptions, see Brainerd & Reyna, 1993; Kalish, 2010).

This disconnect is unfortunate, especially when one considers the potential overlap between memory and inductive reasoning. Inductive reasoning involves extending knowledge from known to novel instances. For example, after learning that tigers and giraffes have some novel anatomical property, both adults (Osherson, Smith, Wilkie, Lopez, & Shafir, 1990; Sloman, 1993) and children (Hayes, 2007; López, Gelman, Gutheil, & Smith, 1992) might be reasonably confident that this property is shared by other mammals.

Even a cursory consideration of induction suggests a central role for memory. Being able to retrieve the similarities (and differences) between tigers and giraffes seems central to explaining how a property shared by these categories will be generalized. Memory for feature co-occurrences within categories has also been linked to accurate feature prediction in 7-year-olds and adults (Kalish, 2010). A more specific point of overlap is the central role accorded in both recognition and induction to the similarity between familiar and novel exemplars. In recognition, the probability that an item is recognized as "old" is a positive function of its total similarity to previously studied items (Hintzman, 1988; Jones & Heit, 1993; Ratcliff, 1990). The role of total exemplar similarity in inductive reasoning has not been examined in as much detail, but it seems safe to assume that the probability that a novel item is judged to have a property is related to its similarity to known instances of items having that property (cf. Osherson et al., 1990; Sloman, 1993).

The possibility that an assessment of total exemplar similarity could be a common component in children's reasoning and memory is consistent with evidence of the early emergence of exemplar-based representations. The specific similarity between previously studied items and a novel probe has been shown to affect probe recognition in both infants (Oakes, Kovack-Lesh, & Horst, 2009) and young children (Brainerd, Reyna, Wright, & Mojardin, 2003; Ghetti & Angelini, 2008). Research on the development of categorization has established that infants (Hayne, 1996; Hayne, Rovee-Collier, & Perris, 1987) and preschoolers (Krascum & Andrews, 1993) often represent visual categories as sets of individual exemplars. When a novel probe item is presented, samples of exemplars from familiar categories are retrieved and categorization is determined by the relative similarity of the probe to these instances. The specific similarity between experienced and novel exemplars has also been shown to influence the inductive generalization of novel properties by infants (Furrer & Younger, 2008; Welder & Graham, 2001).

In the current studies, we aimed to take the study of the effects of exemplar-based similarity on children's cognition in an important new direction; namely, to examine the extent to which such

This article was published Online First June 11, 2012.

Brett K. Hayes and Kristina Fritz, School of Psychology, University of New South Wales, Kensington, New South Wales, Australia; Evan Heit, School of Social Sciences, Humanities and Arts, University of California, Merced.

This work was supported by Australian Research Council Discovery Grant DP1093197 to Brett K. Hayes and Evan Heit and National Science Foundation Grant BCS-0616979 to Evan Heit. We thank Ann Martin, Helen Paton, and Melissa Lim for their assistance with this research.

Correspondence concerning this article should be addressed to Brett K. Hayes, School of Psychology, University of New South Wales, Sydney, New South Wales 2052, Australia. E-mail: b.hayes@unsw.edu.au

exemplar-based processing represents a common component in children's recognition memory and inductive reasoning. In addition, we examine the implications of the recognition–induction relationship for the current debate about the developmental course of similarity in induction (cf. Hayes, McKinnon, & Sweller, 2008; Sloutsky & Fisher, 2004).

A Common Paradigm for Studying Memory and Reasoning

As noted above, one reason why links between the development of memory and reasoning have generally been overlooked is because very different methods have been used to study each process. The current work addresses this issue through an approach that allows for a direct comparison of induction and recognition judgments within and between different age groups. In brief, this approach involved presenting 5- to 6-year-old children and adults with a common set of study and test instances under either inductive reasoning ("learn which items share a common property") or recognition memory instructions ("memorize these items"). Critically, the items in the test set varied in their total similarity to study items. This allowed us to examine the contribution of total exemplar similarity to both induction and recognition judgments and to examine developmental change in this process.

Heit and Hayes (2011) piloted this method in a series of studies with adults. This work revealed some differences between patterns of responding in recognition and induction. Adults were generally more likely to make a positive response to novel animals in induction than recognition. This likely reflects the different emphases in induction and recognition instructions; the former invites generalization beyond the given instances, whereas positive recognition responses should be limited to identity matches.

Nevertheless, Heit and Hayes (2011) found evidence of a common cognitive process underlying induction and recognition. A central component determining test responding in both tasks was the assessment of the total similarity between a novel item and previously experienced category instances. This was reflected in a very strong empirical relationship between test responses in induction and recognition (mean correlation = .86 across eight studies). Moreover, a computational model, called GEN-EX for generalization from examples, which assumes that both recognition and induction are based on the total similarity between a test item and previously experienced study exemplars, was able to fit empirical data from both tasks.

Recognition–Induction Relationship and Role of Similarity in Inductive Development

A second reason for examining developmental stability and change in the relationship between induction and memory is that such data have potentially important implications for the current debate on the role of similarity in children's cognition. Early accounts of conceptual development (e.g., Inhelder & Piaget, 1964; Werner & Kaplan, 1963) argued that young children's categories are structured around the perceptual similarity between instances. An understanding of the role of more abstract relations (e.g., taxonomy, causal relations) was thought to emerge much later in development. This *similarity-first* account has recently been revived and extended to children's inductive reasoning and word learning (Jones & Smith, 2002; Rakison & Lupyan, 2008; Sloutsky & Fisher, 2004; Sloutsky, Kloos, & Fisher, 2007; also see Mandler, 2004, for a related account). According to this approach, the similarity between old and new instances is the main driver of categorization and inductive generalization in children below the age of 7 years. Categorization and induction in older children and adults, however, is thought to be dominated by more complex relations, such as category membership.

In contrast with the similarity-first approach, a number of accounts emphasize developmental continuity in some of the processes involved in categorization and induction (e.g., Gelman, 2003; Hayes et al., 2008; Keil, Smith, Simons, & Levin, 1998). Gelman (2003) and Hayes et al. (2008), for example, have shown that like adults, young children consider both perceptual similarity and category membership when generalizing novel properties from familiar to novel instances. Work on children's list memory (for a review, see Bjorklund & Coyle, 1995) has also found that under certain encoding conditions, young children show sensitivity to categorical relations, clustering related items in recall. Moreover, many have argued that conceptual development is better characterized as a gradual and continuous elaboration of knowledge about the ways in which objects are similar (and different) rather than an qualitative shift from the use of perceptual similarity to conceptual similarity (e.g., Murphy, 2004; Needham, Dueker, & Lockhead, 2005). In terms of the development of memory and induction, this *continuous similarity* hypothesis suggests that the processing of exemplar similarity is likely to play a key role in induction and memory in both children and adults.

A third possible developmental trajectory is that, under certain circumstances, children may show less sensitivity to exemplar similarity in induction than adults. There is good evidence that young children are more likely than older children and adults to assume that the members of basic-level categories are homogeneous in terms of their nonobvious properties (Gelman, 1988; Rhodes & Brickman, 2010). When taught novel nonobvious features of a category exemplar, young children are more likely than older children to assume that other category members share this property, even when there are perceptual differences in the surface appearance of category members (Rhodes & Gelman, 2008). This is particularly true for biological categories where young children often believe that all members of the same species share the same unobservable biological features (Gutheil, Vera, & Keil, 1998). This approach suggests that children may show less reliance on exemplar similarity than do adults when generalizing properties between members of the same category.

Experiment 1

To test these competing accounts, we adapted the common induction and recognition paradigm first described by Heit and Hayes (2011) for use with children. In the induction condition, participants were asked to learn about instances from a single category (large dogs) that shared a novel anatomical property (e.g., "has sarca inside"), whereas those in the recognition condition were asked to memorize the same instances. Both groups were then shown a common test set that contained both old instances and new instances that varied in similarity to old items (i.e., new large dogs that functioned as lures as well as new medium-sized and small dogs). In the induction condition, children and adults were instructed to respond "yes" if they thought a test item had the target property. In the recognition condition, they responded "yes" if they thought a test item had been presented during the study phase.

A considerable body of evidence suggests a developmental increase in recognition memory accuracy over the elementary school years (e.g., Hayes & Hennessey, 1996; Sophian & Stigler, 1981). Hence, one general prediction was that adults should be more accurate than children in discriminating old from new and lure test items in recognition. A second general prediction based on Heit and Hayes (2011) was that induction instructions would lead children and adults to a higher level of positive responding to novel items (i.e., would show broader generalization of knowledge acquired during study) than would recognition instructions.

The competing accounts of inductive development described above make different predictions about (a) the level of overlap between induction and recognition responding and (b) developmental changes in induction responding. The similarity-first account suggests that children should use overall similarity as a basis for responding in both recognition and induction conditions. Children's positive responses on the induction and recognition tests should therefore be highly correlated. Adults should use similarity as a basis for recognition (Hintzman, 1988; Ratcliff, 1990) but use other principles, like category membership, as the basis for induction judgments (Sloutsky & Fisher, 2004). Because the test items are all members of the same basic level category, this account suggests that adults should generalize the anatomical property more broadly across test items than children. The relationship between recognition and induction responding at test should therefore be weaker for adults than for children.

In contrast, the continuous similarity account suggests that total similarity will be used as a basis for responding in both induction and recognition by both age groups. According to this account, once age differences in generalization gradients are factored out, adults and children should show parallel patterns of responding in recognition and induction. The empirical relation between memory and reasoning should be strong for adults and possibly for children as well.

Previous work on children's beliefs about category homogeneity (e.g., Rhodes & Brickman, 2010) predicts a third possible pattern of results. Although exemplar similarity should influence recognition judgments in both age groups, this work suggests that children, compared with adults, should show broader inductive generalization of a novel property to new test items. Note that although the similarity-first and development of homogeneity accounts both predict a stronger correlation between recognition and inductive judgments in adults than in children, they do so for very different reasons. In the case of similarity first, it is because children doing induction will be more reliant than adults on exemplar similarity. In the case of the homogeneity account, it is because children's induction will be less influenced than adults' induction by variations in exemplar similarity within a category.

These predictions were tested in three ways. First, for each age group, we examined the proportion of positive responses to old and new test stimuli in the induction and recognition conditions and used signal detection methods to compare sensitivity in discriminating between old and new test items. Second, for each age group, we computed the itemwise correlation between the probability of making a positive response at test under induction and recognition conditions. Third, we fitted the exemplar-based model GEN-EX to induction and recognition data in each age group. This allowed us to examine whether task and age-related changes in test responding could be captured by a single cognitive model based on total exemplar similarity.

Method

Participants. Sixty first-grade children (M_{age} = 6 years 3 months; range = 58–85 months; 31 girls, 29 boys) were recruited from private elementary schools in a metropolitan area. Sixty psychology undergraduates (M_{age} = 18 years 11 months; 40 women, 20 men) participated for course credit. Within each age group, equal numbers were randomly assigned to a recognition or an induction condition.

Materials. The stimuli were 10-cm^2 color photographs of dogs, 280 pixels square, adapted from a compendium of dog breeds (American Kennel Club, 2006) and other Internet sources. Each photograph showed a dog in a canonical left-facing side view. The same stimulus set was used for both task conditions. The study list consisted of 10 pictures of large dogs. The test list consisted of 45 pictures of dogs, 10 of which were old items (the large dogs originally studied), 15 of which were lure items (other large dogs, not previously studied), and 20 of which were new items (10 small dogs and 10 medium dogs).

Procedure. Participants were tested individually in university laboratories or in a quiet room at their school. All experimental stimuli were presented on a computer using Revolution 4.0 software. In the recognition condition, participants were instructed to memorize study pictures for a subsequent recognition test. They were shown the 10 pictures on the study list in a random order, each presented for 5 s, with a 0.5-s interstimulus interval during which the screen was blank. There was a 60-s unfilled retention interval before the test phase during which participants were asked to count backward. Children counted out loud with the experimenter. The study phase for the induction condition was similar to that of the recognition condition, except for the experimental instructions. Before commencing study, participants were told their task was to learn about animals "who have sarca, which is something inside their body." All of the animals presented during study were said to have sarca.

During the test phase for both groups, 45 test pictures (10 old dogs, 10 new medium dogs, 10 new small dogs, and 15 lures of new large dogs) were shown sequentially, in a different random order for each participant. Those in the recognition condition were instructed to respond "yes" if they believed that the picture had been presented at study or "no" if the picture was novel. Those in the induction condition were instructed to respond "yes" if they believed the animal had sarca and "no" if it did not. Adults responded by clicking their mouse on either a "yes" or a "no" button on the computer screen. Children responded by verbally stating "yes" or "no" and the experimenter clicked the mouse on the corresponding button. Each test item remained on the screen until a response was made. After each response was a 0.5-s interstimulus interval during which the screen was blank. Note that no mention of the category label *dogs* was made at any point in the procedure. This avoids the problem, noted by Sloutsky and Lo (1999), that applying a common verbal label to different instances may lead children to see them as more similar.

Results

Probability of responding "yes." The probability of making a positive response to test items under recognition and induction conditions is shown in Table 1. What is most immediately striking about the results is that children show worse discrimination between old and new items than do adults; indeed, children seem to be completely insensitive to the distinction between old and new items for induction. These response data were entered into a series of univariate 2 (age) $\times$ 2 (task) analyses of variance. For old items, a larger number of positive responses (i.e., hits) were made under induction than recognition conditions, $F(1, 116) = 10.74$, $p = .001$, but responding was not affected by age ($F < 2.1$). For both age groups, the probability of responding positively to new small and medium dogs (i.e., false alarms) was higher in induction than in recognition conditions, $F(1, 116) = 75.99$, $p < .001$. Positive responding to lures was also higher in the induction condition than the recognition condition, $F(1, 116) = 74.73$, $p < .001$. Across both tasks, children were more likely than adults to respond positively to both new items, $F(1, 116) = 37.07$, $p < .001$, and to lures, $F(1, 116) = 22.63$, $p < .001$. No significant Age $\times$ Task interactions were found in any of these analyses (all Fs < 1.2).

Overall, there were a larger number of positive responses to medium than small new dogs, $F(1, 116) = 8.70$, $p = .04$, but this effect interacted with age, $F(1, 116) = 12.71$, $p = .001$. Table 1 shows that adults made a greater number of positive responses to medium than to small dogs, but children were equally likely to respond positively to each type of new item.

To further examine patterns of generalization in the recognition and induction conditions, we calculated a d' measure of sensitivity for each participant using individual hit rates and false alarm rates for new (small and medium) dogs and lure items, respectively. Individual hit and false alarm rates of zero or one were corrected using the procedure outlined by Macmillan and Kaplan (1985), whereby a value of $1/(2N)$ was added to rates of 0 and subtracted from rates of 1, where N is the number of stimuli of a given type presented at test. This meant that a defined value of d' could be calculated for all participants. This correction method was also used in Experiment 2.

The mean sensitivity values for each condition are shown in Figure 1. Adults were more sensitive in discriminating between old and new items (i.e., showed higher ds) than were children, $F(1, 116) = 60.47$, $p < .001$. Sensitivity in old–new discrimination was higher for recognition than induction, $F(1, 116) = 55.84$, $p < .001$. There was no significant Age $\times$ Task interaction ($F < 1.0$). A similar pattern was found in the discrimination between old and lure items, with sensitivity higher in recognition than induction, $F(1, 116) = 57.78$, $p < .001$, and higher sensitivity for adults than children, $F(1, 116) = 44.58$, $p < .001$. Again, the Age $\times$ Condition interaction was not significant ($F < 3.3$).[1]

Relationship between reasoning and memory. The proportion of "yes" responses for each of the 45 test items was averaged across participants within each of the recognition and induction conditions, and the correlation between responses in these conditions was computed separately for each age group. The correlation between the recognition and the induction conditions for adults was .90 ($p < .001$). This relation is illustrated in Figure 2A, showing a scatterplot of recognition responses versus induction responses, for each type of test item (note that some data points overlap so closely that all 45 data points may not be discriminable). This scatterplot reinforces the finding that there was a greater level of generalization of positive responding for induction than for recognition. The correlation between the recognition and induction responding for children was .28 ($p = .057$). The scatterplot in Figure 2B suggests this lower correlation for children was mainly due to the fact that children's rate of "yes" responding in induction was approximately the same for all types of test items.

Modeling. The GEN-EX model is embodied by two equations. Equation 1 shows the familiarity rule. The familiarity of each test stimulus, *fam*, equals its summed similarity to n studied items. Similarity is assumed to be a negative exponential function of distance, *dist*, between the test and study items, calculated according to the standard Euclidean formula. The free parameter c reflects specificity of responding to test items; lower values of c correspond to broader generalization, whereas higher values correspond to narrower generalization gradients. Put another way, the c parameter reflects "overall discriminability in the psychological space" (Nosofsky, 1986, p. 41), with higher values indicating a greater level of discrimination and less generalization.

$$fam\,(\text{test}) = \sum_{i=1}^{n} \exp\,[-c\; dist\,(\text{test}, \text{study}_i)] \tag{1}$$

$$resp\,(\text{test}) = \frac{fam\,(\text{test})}{fam\,(\text{test}) + \beta} \tag{2}$$

The response rule is shown in Equation 2. Essentially, the probability of a positive response is a monotonic function of a test item's familiarity. The response rule has a single scaling parameter, β. A lower value of β corresponds to a greater overall tendency to respond positively.

We used similarity ratings that had been collected from adults by Heit and Hayes (2011) for each pairing of study and test items. It was assumed that similarity would be a negative exponential function of psychological distance (e.g., Nosofsky, 1986), as illustrated by Equation 3.

$$sim(x, y) = \exp\,[-c\; dist\,(x, y)] \tag{3}$$

The 1–7 similarity ratings had been normalized to the range of 0 to 1, by subtracting 1 then dividing by 6. Equation 3 was used for two purposes. First, Equation 3 was used to convert the similarity ratings to hypothetical distances in stimulus space; that is, this equation was solved for distance as a logarithmic function of similarity, to calculate distances as a function of similarity. For this purpose, the c parameter was arbitrarily set at 1 (allowing c to vary did not systematically improve model fit).

[1] For both experiments, the same pattern of results was found when a nonparametric measure of sensitivity (A′) was used as the unit of analysis (i.e., there were main effects of age and condition but no interaction). In Experiment 1, the d' and A′ results remained unchanged when three children over the age of 7 years were excluded. We also computed the C (criterion) measure of response bias (Macmillan & Creelman, 2005) for all participants. For old–new and old–lure discriminations, children showed a more liberal response bias than did adults (ps $< .001$), and responding was generally more liberal under induction than recognition instructions (ps $< .01$).

Table 1
Mean Proportion of "Yes" Responses and d′ (With Standard Deviations) and GEN-EX (Generalization From Examples) Model Predictions in Experiment 1

	Old	New small	New medium	All new	Lure	d' (old–new)	d' (old–lure)
Data							
Adults memory	.79 (.14)	.07 (.34)	.15 (.14)	.11 (.10)	.18 (.15)	2.23 (0.73)	1.90 (0.73)
Adults reasoning	.85 (.16)	.44 (.39)	.54 (.38)	.49 (.37)	.60 (.34)	1.03 (1.14)	0.66 (0.97)
Children memory	.67 (.24)	.34 (.33)	.38 (.36)	.36 (.33)	.41 (.29)	1.01 (0.86)	0.80 (0.62)
Children reasoning	.86 (.23)	.88 (.18)	.83 (.24)	.84 (.20)	.83 (.24)	−0.12 (0.55)	0.01 (0.50)
Model							
Adults memory	.79	.07	.15	.11	.18	2.06	1.73
Adults reasoning	.85	.46	.53	.49	.61	1.05	0.75
Children memory	.69	.36	.40	.38	.44	0.78	0.64
Children reasoning	.86	.87	.82	.85	.82	0.13	0.04

Second and more important, Equation 3, when substituted into Experiment 1, was used to predict underlying similarity for various experimental conditions. For this purpose, the *c* parameter was allowed to vary for recognition versus for induction judgments and for adults versus children. Hence, we were able to investigate whether children show poorer memory discrimination (and greater generalization) than adults and whether, like the adults in Heit and Hayes (2011), they show poorer discrimination (and greater generalization) for induction than for recognition.

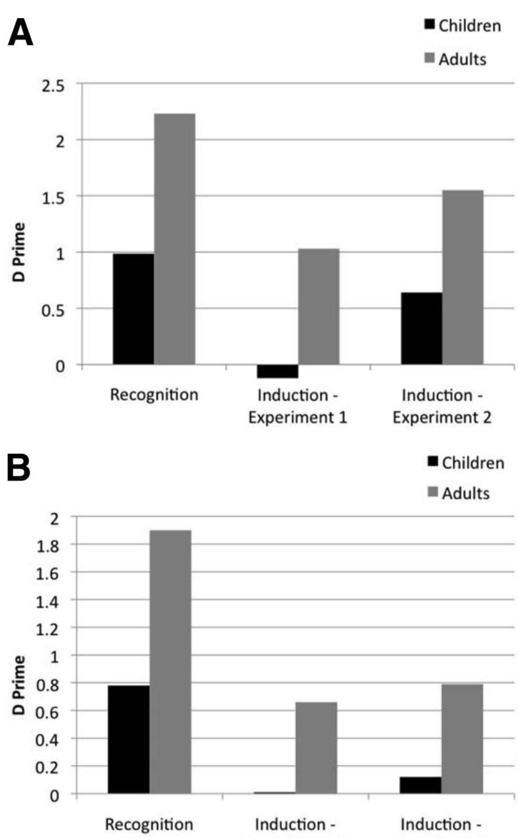

Figure 1. Sensitivity (d') of old–new (A) and old–lure (B) discrimination.

The best fitting parameter estimates for GEN-EX, for adults, were as follows. The *c* parameter for recognition judgments was 13.56, much higher than *c* for induction judgments, which was 5.13. The β parameters for recognition and induction were similar,

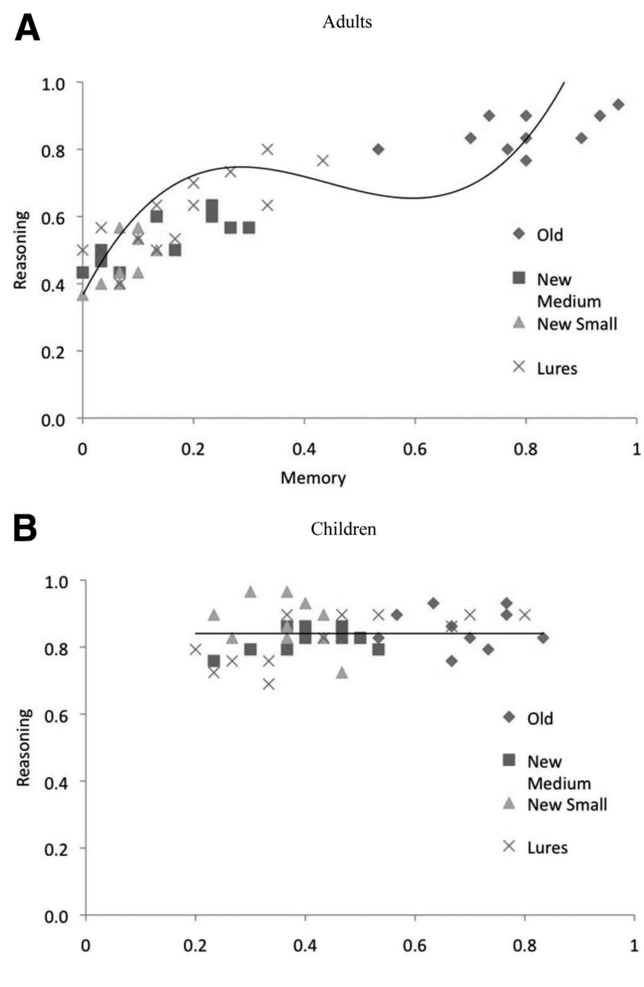

Figure 2. Scatterplots with trend lines for positive responses to test items for adults (A) and children (B) in Experiment 1.

0.29 and 0.55, respectively. With these four free parameter estimates, the model fit the data well overall, with a correlation between model predictions and response proportions, across the 90 data points, of .9329. The root-mean-squared error (RMSE) of prediction was 0.1067. The best fitting parameter estimates for children were as follows. The c parameter for recognition judgments was 4.62, much higher than the c for induction judgments, which was estimated to be 0.00, that is, zero discriminability in memory. The β parameters for recognition and induction were similar, 1.28 and 1.90, respectively. With these four free parameter estimates, the model fit the data reasonably well, with a correlation between model predictions and response proportions, across the 90 data points, of .9568. (Given that all of the children's induction judgments were approximately the same, it would be difficult to improve on this correlation). The RMSE of prediction was 0.1031.

We also fit a more restricted version of the model, attempting to use the same four free parameters for both the adult and the child data. Using Borowiak's (1989) method for model comparison (as described in Heit & Hayes, 2011), we found that fitting all 180 data points with eight free parameters led to a significantly better fit than fitting the data points with just four free parameters, after taking account of the number of free parameters, $\chi^2(4) = 172.57$, $p < .001$. In effect, we found that it is more appropriate to describe the adult and child data with different parameters than with the same parameters.

Table 1 shows average predictions of the GEN-EX model for key types of stimuli. The table shows that the main trends in the data have been captured, such as differences between adults and children; between memory and reasoning conditions; and between old, lure, new medium, and new small items. Likewise, the predicted d' measures are close to the original results.

In Figures 2A and 2B, model predictions are shown as a trend line, which was derived by fitting a third-degree polynomial function to the model's predictions on individual items. The purpose of doing so is to assess whether the model predictions fall in the region of the data points for different types of items. Overall, the trend predicted by the model well captures the trend in the data points (especially for adults), although there is some scatter in the data points around the model predictions.

Discussion

In this study, we examined the development of recognition and inductive judgments about a common set of visual stimuli. In both recognition and induction conditions, children made a greater number of positive responses to novel items, and hence showed lower sensitivity in old–new discrimination, than did adults. We also found a robust main effect of task; adults and children doing induction were more likely than those doing recognition to make positive test responses (and hence showed lower sensitivity). Induction instructions promoted broader generalization of the novel property across the dog category, whereas recognition instructions discouraged positive responses to test pictures that differed from study items.

Notably, the effects of age and task were additive. No interactions between these factors were found in analyses of positive responding or d' sensitivity. In other words, in both recognition and induction conditions, children showed broader generalization of positive responding than did adults. Once this general age difference was accounted for, children showed the same pattern of task differences between recognition and induction as adults.

These conclusions about age and task effects on generalization were backed up by the parameter estimates from fitting the GEN-EX model, which led to several important conclusions. Replicating Heit and Hayes (2011), adults showed a greater level of generalization for induction than for recognition. An important new finding was that children showed greater generalization, or poorer memory discrimination, than did adults. Like adults, children also showed a different level of generalization for induction; however, this level of generalization was so great that all test items essentially received the same positive response regardless of type. Finally, the modeling showed that GEN-EX fits significantly better when different parameters are estimated for adults and children. Indeed, a major aim in applying the models was to show differences in the c parameter for the different conditions. The overall fit of the GEN-EX model was satisfactory, especially considering that there were 22.5 data points per free parameter.

We now turn to the deeper issues of the processes underlying recognition and induction and whether these processes undergo developmental change. Our adult data suggest that induction and recognition differ in terms of the generalization of positive responding. Notwithstanding this difference, we found that the total similarity between a target item and previously studied items played an important role in both adult recognition and induction. This was shown by the high itemwise correlation between positive test responses in induction and recognition and through the good fit to the data provided by the GEN-EX model, which focuses on exemplar similarity.

For children, the conclusions about underlying processes were more complicated. On the one hand, the absence of Age $\times$ Task interactions in analyses of test responding suggests there were no qualitative differences in the way that adults and children approached induction and recognition. Moreover, GEN-EX was able to fit the child data for both tasks using only parameters reflecting the effects of total exemplar similarity by assuming zero generalization for the induction test.

On the other hand, we need to be cautious about interpreting the model fit for children's induction because there was so little systematic variance in children's test responses on this task. Moreover, the itemwise correlation between children's induction and recognition responses was very low.

Despite these complexities, it is clear that these results are at odds with the similarity-first account. There was no evidence that children were more reliant than adults on exemplar similarity for either induction or recognition. Rather, as suggested by continuity accounts, the data and modeling suggest that exemplar similarity influenced children's induction and recognition judgments, with the main developmental change being an age-related increase the gradient of generalization in both tasks.

Children's relatively flat pattern of inductive responding across test items is consistent with previous work showing that children doing induction are more likely to assume homogeneity in the unobservable features of basic-category members (e.g., Rhodes & Brickman, 2010). In the current circumstance, this apparent homogeneity effect reduced variation in children's test responding in induction, making made it difficult to assess the contribution of exemplar similarity to children's induction. In Experiment 2, we therefore modified the induction procedure with the aim of directing children's attention to the within-category variation in study stimuli and promoting a steeper gradient of generalization in children's induction responses.

Experiment 2

Although young children show broad generalization of internal properties within biological categories, they are sensitive to form–function relations in property induction (e.g., Kelemen, Widdowson, Posner, Brown, & Casler, 2003; McCarrell & Callanan, 1995). For example, Kelemen et al. (2003) presented preschool children with triads where a target stimulus had a higher overall similarity to one study item but had a body part (e.g., a birdlike bill) that was similar to a different study item. Children generalized specific biological properties (e.g., the sort of food eaten by the animals) on the basis of the similarity of relevant body parts rather than overall similarity.

Such findings suggest that children in our induction condition would pay more attention to the dimension of size within the dog category if this dimension was plausibly related to property being generalized. Hence, in this experiment, we ran a new induction condition in which children and adults were told that all of the study items (large dogs) shared the property "has a strong bite." It was expected that children's background knowledge of the relation between size and strength or ferocity would lead them to consider the size of test-phase dogs when generalizing this property; with the bite property more likely to generalize to lures (new large dogs) than to medium and small dogs.

This modification was important for two reasons. First, it was expected to produce more variation in children's test-phase responding in induction, allowing for a stronger test of the empirical relationship between induction and recognition in children. Second, such variation in responding also provides better conditions for testing the fit of the GEN-EX model.

The exemplar-based account of induction and recognition predicts that properties that draw attention to the specific similarities and differences between category members should lead to an increased contribution of total similarity to inductive responding. This means that we should see an increase in the strength of the relationship between children's responses in induction and recognition tasks (note that child and adult data from the new induction–bite condition were compared with the recognition data from Experiment 1). It also means that GEN-EX should show an even better fit to children's induction and recognition data than was found in Experiment 1.

Method

Participants. Thirty 1st-grade school children from a metropolitan private school were recruited (M = 6 years 1 month; range = 62–83 months; 11 girls, 19 boys). The same number of first-year undergraduates (M = 18 years 11 months; 23 women, seven men) participated for course credit.

Procedure. The procedure was the same as the induction condition in Experiment 1, except that in the study phase, participants were told that they would be shown a set of animals "who all have a strong bite." After a 60-s delay during which participants counted aloud, they were shown the 45 test items. Participants were instructed to respond "yes" if they thought a test animal had a strong bite and "no" if they did not. The mode of responding for adults and children was the same as in Experiment 1.

Results and Discussion

Probability of responding "yes." The probability of making a positive response to test items in the induction–bite condition is shown in Table 2. Positive responses from the recognition condition in Experiment 1 and the data from the induction–bite condition were entered into a series of univariate 2 (age) × 2 (task) analyses of variance. The rate of positive responding to old items was higher in induction than in recognition, $F(1, 116) = 16.26$, $p < .001$, and higher for adults than children, $F(1, 116) = 4.61$, $p = .03$. There was a higher rate of positive responding in induction than recognition for both new items, $F(1, 116) = 24.81$, $p < .001$, and lures, $F(1, 116) = 86.15$, $p < .001$. Children were more likely than adults to make a positive response to both new items, $F(1, 116) = 26.47$, $p < .001$, and lures, $F(1, 116) = 23.19$, $p < .001$. No significant Age × Task interactions were found in any of these analyses (Fs < 1.0).

Overall, there were more positive responses to medium than to small dogs, $F(1, 116) = 44.4$, $p = .04$, but this effect interacted with task, $F(1, 116) = 7.04$, $p = .009$, and age, $F(1, 116) = 5.29$, $p = .023$. Table 2 shows that the difference in responding to small and large dogs was greater in the induction–bite condition than in recognition and larger in adults than in children. Follow-up tests confirmed that children in the induction–bite condition made more positive responses to lures than to medium dogs, $F(1, 29) = 18.63$, $p < .001$, and more positive responses to medium than to small dogs, $F(1, 29) = 5.65$, $p = .024$.

Sensitivity in the discrimination between old and new items (see Figure 3A) was higher in the recognition than in the induction–bite condition, $F(1, 116) = 10.60$, $p = .001$, and higher for adults than

Table 2
Mean Proportion of "Yes" Responses and d′ (With Standard Deviations) for Induction Condition and GEN-EX (Generalization From Examples) Model Predictions in Experiment 2

	Old	New small	New medium	All new	Lure	d' (old–new)	d' (old–lure)
Data							
Adults reasoning	.87 (.13)	.20 (.26)	.37 (.24)	.35 (.28)	.60 (.26)	1.55 (1.07)	0.79 (0.77)
Children reasoning	.85 (.17)	.56 (.34)	.65 (.30)	.60 (.31)	.79 (.22)	0.64 (0.74)	0.12 (0.43)
Model							
Adults reasoning	.86	.17	.33	.25	.58	1.76	0.88
Children reasoning	.85	.56	.65	.61	.79	0.78	0.24

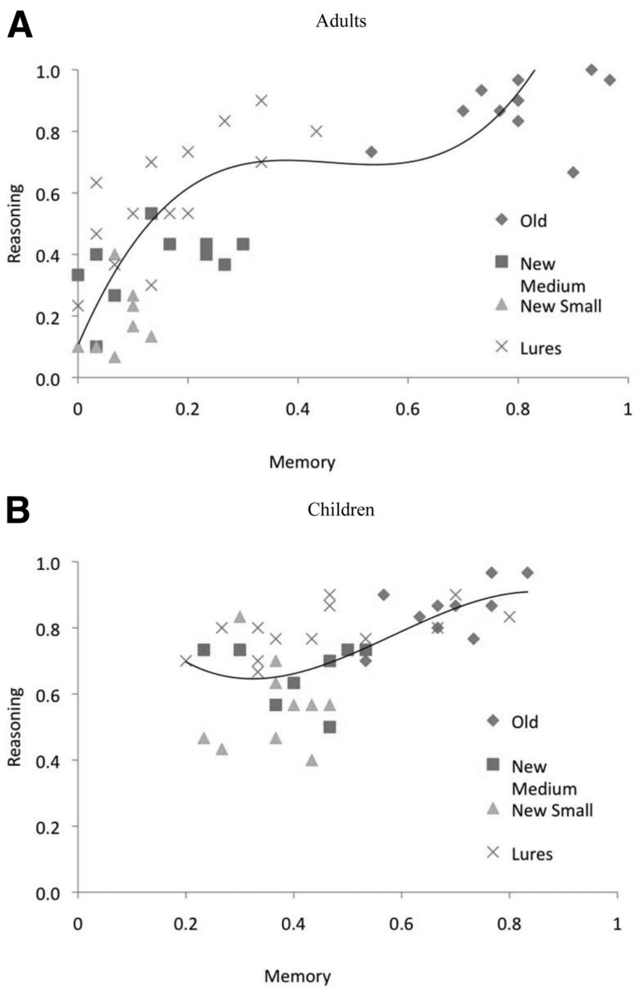

Figure 3. Scatterplots with trend lines for positive responses to test items for adults (A) and children (B) in Experiment 2.

for children, $F(1, 116) = 46.90$, $p < .001$. There was no significant Age × Task interaction ($F < 1.5$). A similar pattern was found for sensitivity in the discrimination between old and lure items (see Figure 3B). Sensitivity was higher in recognition than induction, $F(1, 116) = 54.81$, $p < .001$, and higher for adults than children, $F(1, 116) = 56.69$, $p < .001$, but again there was no significant interaction ($F < 3.5$).

Relationship between reasoning and memory. The proportion of "yes" responses for each of the 45 test items was averaged across participants within each of the two task conditions, and the correlation between responses in these conditions was computed for each age group. The correlation between the recognition condition and the induction condition for adults was .80 ($p < .001$). This relation is illustrated in Figure 3A. As in Experiment 1, there was a greater level of generalization for induction than for recognition. The correlation between the recognition condition and the induction condition for children was .57 ($p < .001$). This correlation was significantly lower than the corresponding correlation for adults, $z = 2.41$, $p = .02$. Nevertheless, the scatterplot of children's responses in Figure 3B suggests that despite children's higher rate of positive responding, there was a roughly linear relationship between children's probability of saying "yes" to test items in induction and recognition.

Modeling. The best fitting parameter estimates for GEN-EX for adults were as follows. The c parameter for induction judgments was 9.41, higher than c for induction judgments in Experiment 1 but lower than c for Experiment 1 recognition judgments. The β parameter was .19. With these two free parameter estimates, the model fit the data well overall, with a correlation between model predictions and response proportions across the 45 data points of .8954. The RMSE of prediction was 0.1310. The best fitting parameter estimates for children were as follows. The c parameter for induction judgments was 4.47, much higher than c for induction judgments in Experiment 1 but approximately the same as for recognition judgments in that experiment. The β parameter for induction was 0.40. With these two free parameter estimates, the model fit the data reasonably well, with a correlation between model predictions and response proportions across the 45 data points of .8231. The RMSE of prediction was 0.0806.

We also fit a more restricted version of the model, attempting to use the same two free parameters for both the adult and the child data. Using Borowiak's (1989) method, we found that fitting all 90 data points with four free parameters led to a significantly better fit than fitting the data points with just two free parameters after taking account of the number of free parameters, $\chi^2(2) = 32.24$, $p < .001$. Again, we found that it is more appropriate to describe the adult and child data with different parameters than with the same parameters.

Table 2 shows average predictions of the GEN-EX model, for key types of stimuli. The table shows that the main trends in the data have been captured, such as differences between adults and children and differences between old, lure, new medium, and new small items. Likewise, the predicted d' measures are close to the original results. In Figures 3A and 3B, model predictions are shown as a trend line. Overall, the trend predicted by the model well captures the trend in the data points, although there is some scatter in the data points not predicted by the model.

Hence, GEN-EX was again able to capture the essential patterns in both child and adult responses. Here, the key difference in induction responses between children and adults can be explained in terms of greater generalization (or poorer memory discrimination) for children. Unlike Experiment 1, one does not need to assume zero memory discrimination for children in the induction condition.

In summary, the use of the "bite" property had relatively little effect on adult induction. Adults continued to show a strong effect of the similarity between study and test items on inductive generalization, and the relationship between induction and recognition remained strong for this group.

Notably, the use of this more concrete property led children to rely more on the similarity between study and test items as a basis for induction than was the case in Experiment 1. For example, children in this study were more likely to generalize a strong bite to new large dogs (lures) than to medium-sized dogs, which in turn were more likely to be judged to have the property than small dogs. This change also affected the relationship between children's induction and recognition responses to the common test set. In this study, we found a moderate to strong positive correlation between children's memory and reasoning judgments, although this relationship was still weaker than the corresponding relation for adults.

The results of the GEN-EX modeling suggested that memory and reasoning in both children and adults could be explained using a common exemplar-based framework with developmental differences reflecting an age-related narrowing in the breadth of generalization.

General Discussion

The current experiments represent a first step toward addressing the disconnect between the study of memory development and the development of inductive reasoning. To examine the similarities and differences between memory and inductive judgments, we administered a common study and test set to adults and children under either recognition or induction instructions. For adults, a key finding was that the similarity between a given test item and previously experienced study items was a critical determinant of both recognition memory and inductive judgments. Adult induction was associated with broader generalization than recognition (i.e., in both studies, adults made a greater number of positive responses to test items under induction than recognition instructions). Nevertheless, there was a strong correlation between the probability of responding "yes" at test in induction and recognition. These results reinforce the view that for adults, there is a common similarity-based component in inductive reasoning and recognition memory, but that induction is associated with broader generalization from learned information (cf. Heit & Hayes, 2011).

The key developmental question was whether this same kind of close relationship between memory and reasoning exists for 5- to 6-year-old children. On the one hand, we did find some developmental differences. Both experiments found that children were more likely than adults to make positive test responses (and so were poorer at discriminating old from new items) in recognition and induction. On the other hand, there were some important developmental continuities. In both experiments, children (like adults) showed a robust effect of old–new similarity on their recognition responses. When a more concrete property was used (as in Experiment 2), this effect of similarity was also found in judgments about inductive generalization. This was reflected in a moderate to strong correlation between children's memory and induction responses. In effect, Experiment 2 showed that we could predict a substantial amount of variance in children's induction judgments from their recognition responses to the same items.

Further insights into the relationship between recognition and induction responses came from our GEN-EX modeling. This exemplar-based model successfully captured the key elements of recognition and induction data for both age groups in each study. The model was able to account for task differences in the probability of making a positive test response by allowing for variation in the c parameter, which reflects the gradient of generalization around the instances encoded during study. Variations in the same parameter could be used to explain developmental differences in responding; for both tasks, children were assumed to have a greater tendency to generalize around a given stimulus set (i.e., had a lower value of c) than adults.

Although the fit of the GEN-EX model was satisfactory in absolute terms, it is worth noting that Heit and Hayes (2011) also fitted versions of GEN-EX with additional components, namely, deterministic responding (say "yes" to old items and "no" to new items) and subtyping (say "yes" to large dogs and "no" to other items), to adult recognition and induction data. We also fit these extended models to the present developmental data. In general, the core of GEN-EX, based on the total similarity between test items and individual study items, was crucial to the success of the model in explaining child (and adult) recognition and induction, including the pattern of generalization across items. Adding a component for deterministic responding did improve the fit beyond exemplar similarity alone, notably for the recognition task, as in Heit and Hayes (2011). However, exemplar similarity contributed more overall to accounting for variance in test phase responding. In contrast, adding a subtyping component did not significantly improve the fit beyond exemplar similarity alone, and, in relative terms, the subtyping component gave a poor account of the results on its own.

In sum, our common induction and recognition paradigm highlighted one important difference, namely, in generalization gradients, and many similarities between induction and recognition judgments. For adults (Experiments 1 and 2) and children (Experiment 2), there was a strong positive relationship between the probability of making a positive response to test items under induction and recognition instructions. Test responding in both conditions and both age groups could be explained by a processing model based on an assessment of the total similarity between test and study exemplars. Children showed broader generalization in recognition and induction than adults but otherwise also showed evidence of a common process of exemplar similarity influencing their responses on both tasks. These data suggest that a capacity for exemplar-based representations of learned information emerges relatively early in development and that such representations influence responding on a range of tasks involving the generalization of learned knowledge.

Implications for the Debate About Inductive Development

These data also have implications for the current debate about developmental change in the role of similarity in inductive reasoning. They reinforce the important role played by exemplar similarity in children's induction (Hayes et al., 2008). They also challenge the view that children always rely more than adults on similarity as a basis for induction (cf. Jones & Smith, 2002; Rakison & Lupyan, 2008; Sloutsky & Fisher, 2004). Adults relied extensively on exemplar similarity as a basis for generalizing both unobservable anatomical properties (Experiment 1) and more familiar behavioral properties (Experiment 2). If anything, adults showed more evidence for similarity-based induction among members of a basic-level category than did children.

The results suggest a qualified developmental continuity view of the role of exemplar similarity in induction. Our data support a strong role for exemplar similarity as a basis for inductive inference in both young children and adults. The one specific way in which the age groups appeared to differ in their inductive responding was in their generalization of completely abstract properties across members of the target category (Experiment 1). The broad pattern of inductive generalization shown by children in that study suggests a strong assumption of category homogeneity (cf. Rhodes & Brickman, 2010) that was not shared by the adults.

Comparison With Other Models of Inductive Development

Although similarity has been long been acknowledged as a key component in models of induction (e.g., Osherson et al., 1990) and recognition (e.g., Hintzman, 1988), our approach is novel in that it attempts to explain the development of recognition and induction within a single theoretical framework. Moreover, previous models of induction (e.g., Osherson et al., 1990) have defined similarity in terms of the relations between different categories (e.g., lions, horses, mammals). The current work makes a novel contribution by highlighting the importance of total similarity between known and novel category exemplars in children's property induction.

The model that most closely resembles GEN-EX is Sloutsky and Fisher's (2004). SINC—which stands for similarity, induction, and categorization—was developed to explain how item similarity determines children's categorization and induction judgments. Although SINC was never applied to recognition data, it has been shown elsewhere (Heit & Hayes, 2005) that the model can be recast to generate recognition predictions by adding a component that assesses the total similarity of novel items to familiar items. The GEN-EX model, however, extends this idea in three ways. First, the computation of similarity has been refined to make it more consistent with well-supported models of categorization such as the generalized context model (GCM). Second, GEN-EX explicitly assumes that recognition and induction judgments differ in the breadth of generalization around study items. Third, SINC was only used to explain induction in young children, whereas GEN-EX is proposed as a general account of the relationship between induction and recognition. One of the key contributions of the current work is to show that GEN-EX can account for child and adult responses on both tasks.

Relationship to Other Work on the Development of Memory and Reasoning

Although it is generally true that the study of memory development and the development of reasoning have followed separate courses, an important exception is the body of work inspired by fuzzy trace theory (FTT; Brainerd & Reyna, 1993, 2004; Brainerd et al., 2003). One of the key motivations for FTT was to provide an account of memory–reasoning relationships and how they develop. According to FTT, novel information is encoded in two parallel formats. Verbatim representations are complete records of studied stimuli that include perceptual details. Gist representations involve more abstract summaries of semantic content. An important assumption of FTT is that verbatim traces are usually accessed in tasks such as recognition and recall, whereas gist traces are used for reasoning. Hence, memory for the details of inputs into reasoning (e.g., the initial premises in a reasoning task) can be statistically independent of reasoning performance (for a review of relevant evidence, see Brainerd & Reyna, 1993).

At one level, our results seem inconsistent with this prediction. In Experiment 2, we found a modest but significant positive correlation between children's recognition and inductive judgments for the same items. Moreover, a single process of exemplar similarity, albeit with flexible levels of generalization around exemplars, was able to account for both recognition and induction data in children as well as adults.

It is important to note, however, that FTT also allows for shifts in the kinds of traces that are stored in different types of memory and reasoning tasks (Brainerd & Reyna, 2004; Brainerd et al., 2003). In the current experiments, the use of study and test sets in which all items were drawn from a single basic-level category may have prompted the formation of gist-based traces in recognition as well as induction, leading to a positive relationship between the tasks.

It is also important to note that much of the previous evidence for stochastic independence between memory and reasoning in FTT has involved tasks that focus on deductive reasoning (e.g., transitive inference, class inclusion), where children have to apply the rules of logic to determine whether a conclusion necessarily follows from a set of premises. By contrast, our work focused on the development of the relationship between memory and inductive reasoning. Other evidence suggests that different processing principles are used in induction and deduction (e.g., Heit & Rotello, 2010; Markovits & Handley, 2005; Rotello & Heit, 2009). The current work may be seen as further evidence of a developmental dissociation between the two types of reasoning. Deductive reasoning skills are known to develop slowly, with children below the age of 7 to 8 years generally performing poorly on many deductive tasks (Markovits & Barrouillet, 2002). Moreover, the evidence from FTT shows that deductive ability is not dependent on the development of verbatim memory. Inductive reasoning, however, appears to follow a different developmental course. An ability to make inductive inferences emerges early in development (e.g., Welder & Graham, 2001). The current work suggests that, unlike deduction, the development of inductive reasoning is closely linked to the way children generalize other kinds of knowledge, such as knowledge of identity in recognition.

Caveats and Challenges for Future Research

The current experiments differed in one important respect from many previous studies of children's induction in that they focused on inferences between members of the same basic-level category (dogs). Many other studies have examined how children of various ages project novel properties between basic level categories that do (or do not) belong to the same superordinate category (e.g., from dogs to mammals; e.g., Gelman, 1988; López et al., 1992). Basic and superordinate categories differ in a number of important respects, not the least of which is the degree of perceptual variability among category members. Such variability is thought to be an important factor affecting inductive generalization, with increased variability associated with broader generalization of learned properties (Loose & Mareschal, 1999).

It remains an open question whether we would find the same close links between induction and recognition when members of different basic-level categories are presented during study and test. On the one hand, it seems reasonable to expect that a child would generalize a novel property from a dog to a horse but that they are unlikely to falsely recognize the horse. On the other hand, the members of different taxonomic categories will vary in their mean similarity to a known target (e.g., on the whole, cats are more similar to dogs than are horses), so it is reasonable to predict that we would still find that induction and recognition judgments vary as function of specific similarity.

Another possible concern about the current findings is that model fitting was based on similarity ratings between study and test items made by adults. Previous work suggests that the way that people make perceptual similarity judgments undergoes significant developmental change (Sloutsky & Lo, 1999; Smith, 1989). However, this means the use of adult similarity estimates should, if anything, have reduced the accuracy of GEN-EX predictions for children's induction and recognition. Our current modeling, therefore, may actually underestimate the contribution of exemplar-based similarity to children's induction and recognition.

A final possible concern about our conclusions regarding the relationship between reasoning and memory is that performance on these tasks was measured between subjects. This contrasts with the approach taken by researchers such as Brainerd and Reyna (1993), who have argued that it is possible for task manipulations to have similar effects on memory and reasoning but for performance on the tasks to be stochastically independent within subjects. This raises the question of whether we would have obtained a weaker relationship between children's memory and inductive reasoning performance if we had used a within-subjects task manipulation. Previous results suggest that the answer is no. When we (Heit & Hayes, 2011) asked adults to complete both recognition and inductive reasoning tasks, we found that the relationship between the tasks remained strong ($r = .84$).

The current results and the GEN-EX model itself highlight a commonality between memory and inductive reasoning, namely, that they are both heavily influenced by comparisons between the features of familiar and novel exemplars. However, we do not dispute the notion that other processes contribute to children's memory and reasoning. In memory research with both adults (Yonelinas, 2002) and children (Ghetti & Angelini, 2008; Holliday & Hayes, 2000), there is evidence for a slower, more controlled, recollective process that also contributes to recognition, especially in older children. Likewise, as they get older, elementary school children are more likely to use more abstract causal or thematic relations as a basis for property induction (Hayes & Thompson, 2007). An intriguing prospect for future work is to examine the links between the development of these more controlled and complex processes in inductive reasoning and memory.

An important related question concerns the source of the age difference in generalization gradients that we found in both recognition and induction. One possible answer would be that this reflects an age-related change in the ability to use recollective processes to reject similar lures. This has been repeatedly demonstrated in work on memory development (e.g., Brainerd et al., 2003; Ghetti & Angelini, 2008), but the same effect may also explain children's broader pattern of inductive generalization; that is, children's automatic tendency to generalize properties to similar instances is not moderated by a recollective appreciation of the differences between studied and novel items.[2] Our modeling efforts shed some light on this issue. Simply varying the c parameter for discriminability, from children to adults and from induction to recognition, was sufficient to account for the whole pattern of results in these experiments. However, adding a component for deterministic responding, as in Heit and Hayes (2011), further improved model fits. Hence, although we do not rule out developmental changes in recollection, there appears to be a crucial role for age changes in the ability to discriminate between memory traces.

A final question is to what extent the obtained developmental differences are due to changes in response bias (cf. Brainerd & Reyna, 2004). First, it should be pointed out that children's induction results for Experiment 1, where d' was essentially zero, cannot possibly be explained in terms of a response bias difference from the other conditions. There must be a real difference in sensitivity in that condition compared with other conditions. More generally, we did observe a greater proportion of positive responses for children than for adults, consistent with a more liberal response bias for children. However, our model-based analyses targeted the developmental change in terms of the c parameter for generalization. Age variations in the β parameter, corresponding most directly to response bias within the modeling framework itself, were quite small. In sum, we do not rule out developmental changes in response bias, but such changes cannot account for the whole pattern of results (for further discussion of response bias measurement in memory, including analyses in terms of receiver operating characteristic curves, see Heit, Rotello, & Hayes, 2012).

Conclusion

The current work makes a number of novel contributions to the study of inductive reasoning in children. First, it represents the first evidence of a close link between the processes involved in children's induction and memory. Second, although induction involved broader generalization of responding to novel items, we found clear evidence of a shared exemplar-based process in children's induction and visual recognition. A third important finding was that exemplar-based similarity remained central to both reasoning and memory into adulthood. Although older children and adults are likely to be aware of more abstract relations that can guide induction, this does not mean that similarity-based induction is abandoned. Finally, the link between children's inductive reasoning and memory offers a number of interesting prospects for further research. A range of empirical factors are known to affect children's memory performance (Ornstein & Hayden, 2009). Our work suggests that these same factors may also impact children's inductive reasoning.

[2] We thank an anonymous reviewer for this suggestion.

References

American Kennel Club. (2006). AKC recognized breeds. Retrieved from http://www.akc.org/breeds/breeds_a.cfm

Bjorklund, D. F., & Coyle, T. R. (1995). Utilization deficiencies in the development of memory strategies. In F. E. Weinert & W. Schneider, *Memory performance and competencies: Issues in growth and development* (pp. 161–180). Mahwah, NJ: Erlbaum.

Borowiak, D. S. (1989). *Model discrimination for nonlinear regression models.* New York, NY: Dekker.

Brainerd, C. J., & Reyna, V. F. (1993). Memory independence and memory interference in cognitive development. *Psychological Review, 100,* 42–67. doi:10.1037/0033-295X.100.1.42

Brainerd, C. J., & Reyna, V. F. (2004). Fuzzy-trace theory and memory development. *Developmental Review, 24,* 396–439. doi:10.1016/j.dr.2004.08.005

Brainerd, C. J., Reyna, V. F., Wright, R., & Mojardin, A. H. (2003). Recollection rejection: False-memory editing in children and adults.

Psychological Review, 110, 762–784. doi:10.1037/0033-295X.110.4.762

Furrer, S. D., & Younger, B. A. (2008). The impact of specific prior experiences on infants' extension of animal properties. *Developmental Science, 11,* 712–721. doi:10.1111/j.1467-7687.2008.00721.x

Gelman, S. A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology, 20,* 65–95. doi:10.1016/0010-0285(88)90025-4

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought.* New York, NY: Oxford University Press.

Ghetti, S., & Angelini, L. (2008). The development of recollection and familiarity in childhood and adolescence: Evidence from the dual-process signal detection model. *Child Development, 79,* 339–358. doi:10.1111/j.1467-8624.2007.01129.x

Gutheil, G., Vera, A., & Keil, F. C. (1998). Do houseflies think? Patterns of induction and biological beliefs in development. *Cognition, 66,* 33–49. doi:10.1016/S0010-0277(97)00049-8

Hayes, B. K. (2007). The development of inductive reasoning. In A. Feeney & E. Heit (Eds.), *Inductive reasoning: Experimental, developmental and computational approaches* (pp. 25–54). Cambridge, England: Cambridge University Press.

Hayes, B. K., & Hennessy, R. (1996). The nature and development of nonverbal implicit memory. *Journal of Experimental Child Psychology, 63,* 22–43. doi:10.1006/jecp.1996.0041

Hayes, B. K., McKinnon, R., & Sweller, N. (2008). The development of category-based induction: Reexamining conclusions from the induction then recognition (ITR) paradigm. *Developmental Psychology, 44,* 1430–1441. doi:10.1037/0012-1649.44.5.1430

Hayes, B. K., & Thompson, S. P. (2007). Causal relations and feature similarity in children's inductive reasoning. *Journal of Experimental Psychology: General, 136,* 470–484. doi:10.1037/0096-3445.136.3.470

Hayne, H. (1996). Categorization in infancy. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research* (Vol. 10, pp. 79–120). Norwood, NJ: Ablex.

Hayne, H., Rovee-Collier, C., & Perris, E. E. (1987). Categorization and memory retrieval by three-month-olds. *Child Development, 58,* 750–767. doi:10.2307/1130212

Heit, E., & Hayes, B. K. (2005). Relations between categorization, induction, recognition, and similarity: Comment on Sloutsky and Fisher (2004). *Journal of Experimental Psychology: General, 134,* 596–605. doi:10.1037/0096-3445.134.4.596

Heit, E., & Hayes, B. K. (2011). Predicting reasoning from memory. *Journal of Experimental Psychology: General, 140,* 76–101. doi:10.1037/a0021488

Heit, E., & Rotello, C. M. (2010). Relations between inductive reasoning and deductive reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 805–812. doi:10.1037/a0018784

Heit, E., Rotello, C. M., & Hayes, B. K. (2012). Relations between memory and reasoning. In B. H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 57, pp. 58–101). San Diego, CA: Academic Press.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95,* 528–551. doi:10.1037/0033-295X.95.4.528

Holliday, R. E., & Hayes, B. K. (2000). Dissociating automatic and intentional processes in children's eyewitness memory. *Journal of Experimental Child Psychology, 75,* 1–42. doi:10.1006/jecp.1999.2521

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

Jones, C. M., & Heit, E. (1993). An evaluation of the total similarity principle: Effects of similarity on frequency judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 799–812. doi:10.1037/0278-7393.19.4.799

Jones, S. S., & Smith, L. B. (2002). How children know the relevant properties for generalizing object names. *Developmental Science, 5,* 219–232. doi:10.1111/1467-7687.00224

Kalish, C. W. (2010). How children use examples to make conditional predictions. *Cognition, 116,* 1–14. doi:10.1016/j.cognition.2010.03.008

Keil, F. C., Smith, W. C., Simons, D. J., & Levin, D. T. (1998). Two dogmas of conceptual empiricism: Implications for hybrid models of the structure of knowledge. *Cognition, 65,* 103–135. doi:10.1016/S0010-0277(97)00041-3

Kelemen, D., Widdowson, D., Posner, T., Brown, A. L., & Casler, K. (2003). Teleo-functional constraints on preschool children's reasoning about living things. *Developmental Science, 6,* 329–345. doi:10.1111/1467-7687.00288

Krascum, R. M., & Andrews, S. (1993). Feature-based versus exemplar-based strategies in preschoolers' category learning. *Journal of Experimental Child Psychology, 56,* 1–48. doi:10.1006/jecp.1993.1024

Loose, J., & Mareschal, D. (1999). Inductive reasoning revisited: Children's reliance on category labels and appearance. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the Twenty-First Annual Conference of the Cognitive Science Society* (pp. 320–325). London, England: Psychology Press.

López, A., Gelman, S. A., Gutheil, G., & Smith, E. E. (1992). The development of category-based induction. *Child Development, 63,* 1070–1090. doi:10.2307/1131519

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). New York, NY: Psychology Press.

Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false alarm rates. *Psychological Bulletin, 98,* 185–199. doi:10.1037/0033-2909.98.1.185

Mandler, J. M. (2004). The foundations of mind: Origins of conceptual thought. Oxford, England: Oxford University Press.

Markovits, H., & Barrouillet, P. (2002). The development of conditional reasoning: A mental model account. *Developmental Review, 22,* 5–36. doi:10.1006/drev.2000.0533

Markovits, H., & Handley, S. (2005). Is inferential reasoning just probabilistic reasoning in disguise? *Memory & Cognition, 33,* 1315–1323. doi:10.3758/BF03193231

McCarrell, N. S., & Callanan, M. A. (1995). Form–function correspondences in children's inference. *Child Development, 66,* 532–546. doi:10.2307/1131595

Murphy, G. L. (2004). On the conceptual–perceptual divide in early concepts. *Developmental Science, 7,* 513–515. doi:10.1111/j.1467-7687.2004.00373.x

Needham, A., Dueker, G., & Lockhead, G. (2005). Infants' formation and use of categories to segregate objects. *Cognition, 94,* 215–240. doi:10.1016/j.cognition.2004.02.002

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General, 115,* 39–57. doi:10.1037/0096-3445.115.1.39

Oakes, L. M., Kovack-Lesh, K. A., & Horst, J. S. (2009). Two are better than one: Comparison influences infants' visual recognition memory. *Journal of Experimental Child Psychology, 104,* 124–131. doi:10.1016/j.jecp.2008.09.001

Ornstein, P. A., & Haden, C. A. (2009). Developments in the study of the development of memory. In M. Courage & N. Cowan (Eds.), *Development of memory in infancy and childhood* (pp. 367–386). Hove, England: Psychology Press.

Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review, 97,* 185–200. doi:10.1037/0033-295X.97.2.185

Rakison, D. H., & Lupyan, G. (2008). Developing object concepts in infancy: An associative learning perspective. *Monographs of the Society for Research in Child Development, 73*(1).

Ratcliff, R. (1990). Connectionist models of recognition memory: Constraints imposed by learning and forgetting functions. *Psychological Review, 97,* 285–308. doi:10.1037/0033-295X.97.2.285

Rhodes, M., & Brickman, D. (2010). The role of within-category variability in category-based induction: A developmental study. *Cognitive Science: A Multidisciplinary Journal, 34,* 1561–1573. doi:10.1111/j.1551-6709.2010.01137.x

Rhodes, M., & Gelman, S. A. (2008). Categories influence predictions about individual consistency. *Child Development, 79,* 1270–1287. doi:10.1111/j.1467-8624.2008.01188.x

Rotello, C. M., & Heit, E. (2009). Modeling the effects of argument length and validity on inductive and deductive reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 1317–1330. doi:10.1037/a0016648

Sloman, S. A. (1993). Feature-based induction. *Cognitive Psychology, 25,* 231–280. doi:10.1006/cogp.1993.1006

Sloutsky, V. M., & Fisher, A. V. (2004). Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology: General, 133,* 166–188. doi:10.1037/0096-3445.133.2.166

Sloutsky, V. M., Kloos, H., & Fisher, A. V. (2007). When looks are everything: Appearance similarity versus kind information in early induction. *Psychological Science, 18,* 179–185. doi:10.1111/j.1467-9280.2007.01869.x

Sloutsky, V. M., & Lo, Y.-F. (1999). How much does a shared name make things similar? Part 1. Linguistic labels and the development of similarity judgment. *Developmental Psychology, 35,* 1478–1492. doi:10.1037/0012-1649.35.6.1478

Smith, L. B. (1989). A model of perceptual classification in children and adults. *Psychological Review, 96,* 125–144. doi:10.1037/0033-295X.96.1.125

Sophian, C., & Stigler, J. W. (1981). Does recognition memory improve with age? *Journal of Experimental Child Psychology, 32,* 343–353. doi:10.1016/0022-0965(81)90085-0

Welder, A. N., & Graham, S. A. (2001). The influence of shape similarity and shared labels in infants' inductive inferences about nonobvious object properties. *Child Development, 72,* 1653–1673. doi:10.1111/1467-8624.00371

Werner, H., & Kaplan, B. (1963). *Symbol formation: An organismic-developmental approach to language and the expression of thought.* New York, NY: Wiley.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46,* 441–517. doi:10.1006/jmla.2002.2864

Received October 25, 2011
Revision received February 3, 2012
Accepted April 20, 2012 ■