Development of Semantic Knowledge and Its Role in the Development of Category-Based Reasoning

Karrie E. Godwin (kegodwin@andrew.cmu.edu)
Carnegie Mellon University, Department of Psychology, 5000 Forbes Avenue, Pittsburgh, PA 15213 USA

Bryan J. Matlen (bmatlen@cmu.edu)
Carnegie Mellon University, Department of Psychology, 5000 Forbes Avenue, Pittsburgh, PA 15213 USA

Anna V. Fisher, (fisher49@andrew.cmu.edu)
Carnegie Mellon University, Department of Psychology, 5000 Forbes Avenue, Pittsburgh, PA 15213 USA

Abstract

Prior research indicates a protracted developmental course in category-based reasoning. One possible explanation for the development of this ability is the gradual reorganization of semantic knowledge. To measure development of semantic knowledge we developed a new paradigm, the Semantic Space task, which uses distance in a two-dimensional space to infer semantic similarities between two objects. Using this paradigm we examined development of semantic knowledge in young children (preschoolers, kindergarteners, and first-grade children) and in adults. We also examined whether conceptual organization as measured by the Semantic Space task is predictive of children's scores on a category-based reasoning task. The findings point to the possibility that development of semantic knowledge plays an important role in the development of category-based reasoning.

Keywords: Semantic Space; Concepts; Cognitive Development

Introduction

Category-based reasoning is a critical cognitive ability that enables an individual to generalize from the known to the unknown (Hayes, Heit, & Swendsen, 2010; Proffitt, Coley, & Medin, 2000). For example, upon learning that a chicken has 39 pairs of chromosomes, one may infer that a dove also has 39 pairs of chromosomes because chickens and doves are the same kind of animal (i.e., birds). Prior research indicates that children's category-based reasoning undergoes a protracted developmental course (e.g., Badger & Shapiro, 2012; Fisher, Matlen, & Godwin, 2011; Godwin, Matlen, & Fisher, in press; Fisher, 2010; Fisher & Sloutsky, 2005; Sloutsky, Kloos, & Fisher, 2007), with marked improvements in category-based reasoning apparent between 4 and 6 years of age. However, it is not clear what factors underlie this improvement. One possible explanation for the development of category-based reasoning is representational change: the gradual accretion and reorganization of domain-specific knowledge (Goswami & Brown, 1989, 1990; Rattermann & Gentner, 1998).

Representational change has been identified as a factor that fosters cognitive development in a wide array of domains such as numerical development (e.g., Opfer & Siegler, 2007), problem solving (e.g., Karmiloff-Smith, 1986), and analogical reasoning (e.g., Gentner, Rattermann, Markman, & Kotovsky, 1995).

One of the few studies to examine the relationship between semantic development and category-based reasoning was conducted by Chi, Hutchinson, and Robin (1989). Chi et al. classified 6-year-old children as either dinosaur experts or novices, based on their pre-test performance. Subsequently, the children completed an inference task about dinosaurs. The stimuli were digitally modified in order to create novel dinosaurs for both experts and novices. Chi et al. found that children who were classified as domain experts tended to make category-based inferences about the novel dinosaurs (e.g., "he is probably a good swimmer … cause duckbills are good swimmers", p. 48). In contrast, children who were classified as dinosaur novices tended to make inferences based on a salient attribute (e.g., [the dinosaur] "could walk real fast cause he has giant legs", p. 49). These findings can be taken to suggest that category-based induction is a function of one's domain knowledge (also see Gobbo & Chi, 1986).

There is also converging evidence suggesting that representational change may play a role in semantic development. First, multidimensional scaling studies have investigated people's ability to classify familiar objects. These studies provide evidence of advancement in classification from initially classifying objects according to more concrete characteristics to utilizing more abstract features when making groupings (e.g., Howard & Howard, 1977; Saltz, Seller, & Sigel, 1972). For example, preschool-age children are likely to classify familiar animals on the dimension of size, whereas school-age children are more likely to classify animals along the dimensions of domesticity and predativity.

Second, studies on the development of priming suggest that associative priming (e.g., faster responding to the word 'banana' after 'monkey' is presented) appears early in development whereas semantic priming (e.g., faster responding to the word 'banana' after 'cherry' is presented) develops during the school years (McCauley, Weil, & Sperber, 1976; Arias-Trejo & Plunkett, 2011).

Finally, work in cognitive modeling points to a gradual developmental progression in conceptual organization from relatively undifferentiated (e.g., groupings including a *penguin*, a *trout*, and an *alligator*) to more differentiated groupings (Kemp & Tenenbaum 2008; Rogers &

McClelland 2004). However, these predictions are yet to be confirmed by empirical studies.

The present study investigates how young children organize knowledge and whether the organizational structure changes over the course of development. In particular, we are interested in examining whether semantic similarity influences how knowledge is organized. In order to examine this issue, we developed a task in which participants are asked to represent the semantic similarity of animal dyads in two-dimensional space. The distance between animal pairs is taken as a measure of how closely the participants represented the concepts. The use of physical distance as an indicator of representational similarity has been successfully used in prior studies (e.g., Goldstone, 1994; Howard & Howard, 1977).

Unlike multidimensional studies, in which children are free to arrange the items along any desired dimension, we explicitly asked children to put animals of similar kind close together on the game board. Therefore, this paradigm allowed us to examine whether knowledge of semantic similarity changes over the course of development in 4- to 7-year-old children. We also assessed whether children's semantic organization scores are predictive of their tendency to engage in spontaneous category-based reasoning.

Method

Participants

Participants were preschool children (*N*=43, *M*age=4.32 years, *SD*=0.28 years), kindergarteners (*N*=22, *M*age=5.41 years, *SD*= 0.30 years), and first-grade children (*N*=23, *M*age=6.96 years, *SD*=0.32 years) attending local preschools and elementary schools. The preschool children were also part of a longitudinal study examining the development of inductive reasoning (see Godwin, Matlen, Fisher, 2012). Adult participants were undergraduate students (*N*=20, *M*age=20.38 years *SD*=1.22 years) from a local university who received partial course credit for participation.

Design & Procedures

Children were tested individually in a quiet room adjacent to their classroom. Adult participants were tested in a laboratory located on campus. Tasks were administered by hypothesis-blind experimenters.

Semantic Space Task

This task was designed to assess children's semantic organization. Visual stimuli entailed a game board consisting of a 9x9 grid (see Figure 1). Two 1" wooden blocks were used as game pieces. The wooden blocks were used as game pieces instead of pictures so that children would use their knowledge about kinds rather than rely on perceptual similarity. A similar approach has been successfully used in prior studies (e.g., Howard & Howard, 1977).

Linguistic stimuli included 24 pairs of animal names. The stimuli could be classified into one of four categories: (1) semantically-similar dyads (e.g., *lamb-sheep*), (2) dyads that share a common setting or habitat (e.g., *lamb-horse*), (3) unrelated dyads (la*mb-swan*), and (4) filler items. During the game, the target item was paired with three different test items (i.e., category-choice, setting/habitat match, and unrelated item). Linguistic stimuli is provided in Table 1.

Figure 1: Schematic depiction of the Semantic Space game board. Squares highlighted in red indicate the location of the critical trials. Squares highlighted in yellow mark the location of the filler trials. In the experiment proper, the location of the critical and filler trials was not marked and all squares on the board were white.

Table 1: Linguistic Stimuli for the Semantic Space Task

Critical Trials			
Target	**Category - Choice**	**Setting/ Habitat**	**Unrelated**
Crocodile	Alligator	Fish	Grasshopper
Chick	Hen	Goat	Goldfish
Lamb	Sheep	Horse	Swan
Whale	Dolphin	Octopus	Elephant
Monkey	Gorilla	Parrot	Chipmunk
Mouse	Rat	Pig	Hippo
Filler Pairs			
1. Zebra/Turkey; 2. Bear/Snake; 3. Panther/ Turtle; 4. Tiger/Butterfly; 5. Frog/Lion; 6. Giraffe/Seal			

In the Semantic Space task, participants were asked to help *Zibbo* the zookeeper organize his zoo by placing animals of the same kind close together. At the beginning of the task, the experimenter introduced the game and provided the participants with two examples (Example 1: a *bunny* and a *rabbit* were placed on adjacent squares on the game board and the experimenter explained that they should be placed close together because they are the same kind of thing; Example 2: a *dog* and a *shark* were placed far apart on the game board and the experimenter explained that they should be placed far apart because they are not the same kind of thing). On each test trial, the experimenter showed the participant where Zibbo put the target animal (e.g., the experimenter placed the first game piece on one of the squares marked in red in Figure 1 and said, *"The zookeeper put the mouse here"*). Then, the experimenter handed the participant the second game piece and asked him or her to identify where the test item should be placed (e.g., *"Where*

do you think the hippo should go?"). The participant's response was recorded in order to calculate the distance between the placement of the target and test item. After each trial was administered, the game board was cleared before the experimenter presented the next pair.

Placement of the 18 critical trials (i.e., semantically-similar dyads, common habitat/setting dyads, and unrelated items) was pseudo randomized to eight central squares (marked in red in Figure 1). The central squares were utilized for the critical trials in order to equalize the maximum possible distance from the square where the experimenter placed the target. Each of the eight squares was utilized at least twice and no more than three times. The six filler trials were randomly assigned to one of the remaining 24 squares in order to encourage participants to use the entire game board. The animal dyads were presented in one of two pseudo randomized orders. The following stipulations were used when creating the presentation orders: one filler trial was presented after every three critical trials, at least three trials were required in between target repeats, and at least two trials were presented in between semantically-similar dyads. The presentation order was counterbalanced across participants.

Participants' responses were scored in the following way: Raw scores were calculated for each trial by adding the number of squares occupied by the game pieces plus the number of squares between the target and test item (the distance was based on the shortest route between the two game pieces barring diagonal movement). A composite score for non semantically-similar dyads was created by averaging together participants' raw scores for common setting/habitat dyads and unrelated items. A Semantic Space Difference score was calculated by subtracting the average score for semantically-similar dyads from the non semantically-similar composite score. Positive difference scores indicate that participants put semantically-similar dyads closer together and non semantically-similar dyads farther apart. Difference scores approaching zero indicate that participants did not reliably discriminate between semantically-similar dyads and non semantically-similar dyads.

Category-Based Reasoning Task

The Category-Based Reasoning task is a property-induction task in which children are presented with triads of objects and asked to generalize a novel property from the target to one of the test items. Each triad included a target, a category-choice, and a lure (e.g., *lamb-sheep-frog*). Nine label triads were administered: 3 triads referring to artifacts, 3 triads referring to inanimate natural kinds, and 3 triads referring to animate natural kinds (see Table 2 for the complete list of linguistic stimuli). The 3 animate natural kind triads were also included in the Semantic Space task.

Visual stimuli were presented on the computer and consisted of sets of three identical doors; see Figure 2. The objects remained hidden behind the doors in order to encourage children to rely on the category information

conveyed by the labels. This procedure has been utilized successfully in prior work (e.g., Fisher et al., 2011; Godwin et al., in press).

On each test trial children were told where each object was hiding. Children were told that the target item had a novel property and asked to generalize this property to either the category-choice or the unrelated lure. Concerns regarding the working memory demands of the task are mitigated based on Fisher et al.'s (2011) findings in which children recalled, with high accuracy, which objects were hiding behind each door at the end of each test trial.

All properties were two-syllable blank predicates. Two presentation orders were created: In order 1 all trials were randomized and the presentation order was reversed for order 2. Presentation order was counterbalanced across participants.

Table 2: Category-Based Reasoning Task Linguistic Stimuli

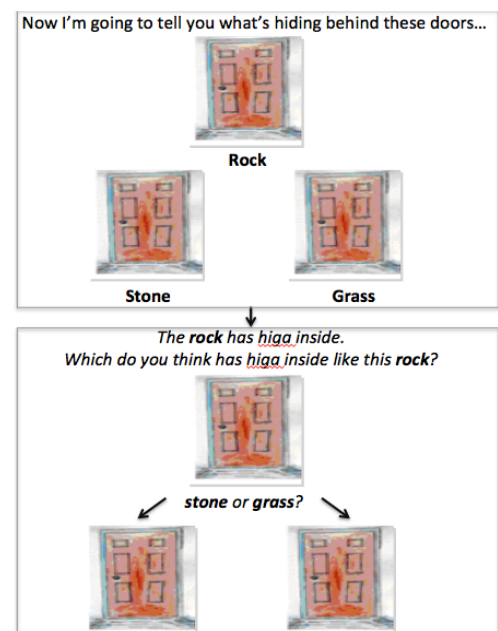| Target	Category Choice	Lure	Property
Artifacts			
Rug	Carpet	Window	Koski
Sofa	Couch	Cup	Creighan
Shoe	Boot	Car	Troxel
Inanimate Natural Kinds			
Sea	Ocean	Apple	Manchin
Hill	Mountain	Flower	Erwin
Rock	Stone	Grass	Higa
Animate Natural Kinds			
Alligator	Crocodile	Butterfly	Omat
Rat	Mouse	Fish	Lignin
Lamb	Sheep	Frog	Matlen

Figure 2: Schematic depiction of the Category-Based Reasoning task. All instructions were given verbally by the experimenter.

Picture Identification Task

The picture identification task served to assess children's familiarity with the labels utilized in the Category-Based Reasoning task. The picture identification task is a computer-based task akin to the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997). Stimuli included 27 labels and 108 pictures. On every trial, children were presented with 4 pictorial response options (the target object and 3 lures). Children were asked to point to the target object. The trials were presented in one of two orders. The presentation order was counterbalanced across participants. The task was administered immediately following the Category-Based Reasoning task.

The Category-Based Reasoning task and the picture identification task were not administered to adults; only preschoolers, kindergartners, and first-grade children completed this portion of the experiment. Additionally, because preschool children were also participating in a related longitudinal study, they participated in the Category-Based Reasoning task twice, with approximately one week between the two testing sessions. Repeated testing was administered to obtain a more stable estimate of young children's performance. As children's scores on both testing sessions were within 3% (adjusted means $M_{time1}=63\%$, $M_{time2}=66\%$), we averaged the scores across the repeated administrations of the task. The analyses reported below are based on these average scores.

Results

Semantic Space Task Performance

Preschool children exhibited considerable variability in their performance on the Semantic Space task, with Difference scores ranging from -2.58 to 5.67, and an average Difference score of 1.37 (SD=1.88). Kindergarten children's performance was also highly variable, with Difference scores ranging from -0.92 to 7.00 and an average Difference score of 2.44 (SD=2.48). The Difference scores of first-grade children ranged from -0.17 to 8.25 and their average difference score was 4.23 (SD=2.08). The Difference scores of adult participants ranged from 2.50 to 6.50 and their average Difference score was 4.99 (SD=1.13); See Table 3.

Participants' Difference scores were analyzed in a one-way ANOVA with age as the between-subject factor. This analysis revealed a significant effect of age, $F(3, 104)=20.41$, $p<0.0001$. This effect was further explored through planned comparisons.

Performance on the Semantic Space task was found to improve with age. In general, preschoolers exhibited greater difficulties discriminating between semantically-similar and non semantically-similar dyads compared to the other age groups. Post-hoc Tukey tests revealed no significant difference between mean Difference scores of preschoolers (M=1.37) and kindergartners (M=2.44), p=0.163. At the same time, both first graders (M=4.23) and adults (M=4.99) exhibited superior performance compared to preschoolers (M=1.37), both $ps<0.0001$. A marked improvement in

performance on the Semantic Space task was observed between kindergarten (M=2.44) and first-grade (M=4.23), p=0.014. There was no significant difference between the Difference scores of first-graders and adults (p=0.583), providing preliminary evidence that semantic differentiation for certain animal categories may begin to reach adult levels by 6 to 7 years of age. Taken together, the pattern of results suggests that the ability to reliably discriminate between semantically-similar and non semantically-similar dyads improves with age; see Figure 3.
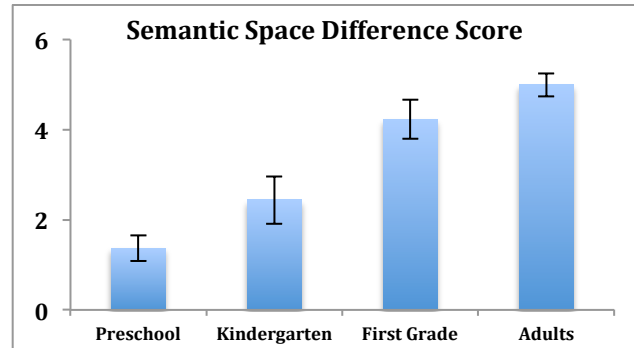

Figure 3: Mean difference scores by age group. Semantic Space Difference scores were calculated by subtracting the average score for semantically-similar dyads from the non semantically-similar composite score. Error-bars represent the standard errors of the mean.

Table 3: Semantic Space mean scores by item type and age group.

Age Group	Mean (SD)
Semantically-Similar Dyads	
Preschool	4.37 (1.35)
Kindergarten	4.09 (1.63)
First-Grade	2.91 (0.87)
Adults	2.32 (0.42)
Non Semantically-Similar Dyads	
Preschool	5.74 (1.51)
Kindergarten	6.53 (1.93)
First-Grade	7.14 (1.79)
Adults	7.31 (0.99)
Common Setting/Habitat Dyads	
Preschool	5.83 (1.60)
Kindergarten	6.45 (2.06)
First-Grade	6.66 (1.87)
Adults	6.13 (1.14)
Unrelated Dyads	
Preschool	5.64 (1.76)
Kindergarten	6.61 (1.95)
First-Grade	7.62 (2.08)
Adults	8.48 (1.05)

Picture Identification

The results from the picture identification task indicated that children were familiar with the labels used in the Category-Based Reasoning task (Preschoolers: $M=.92$, $SD=.14$, Kindergarteners: $M=.99$, $SD=.01$, First-Graders: $M=.99$, $SD=.01$). As an additional precaution, for the preschool group the Category-Based Reasoning scores were adjusted for their vocabulary knowledge to ensure that children possessed the pre-requisite knowledge to perform category-based induction. Thus, if a child missed an item on the picture identification task, this trial was removed from this child's Category-Based Reasoning score. This adjustment resulted in the increase of mean Category-Based Reasoning scores in preschoolers from $M=.62$ to $M=.64$. Because the picture identification scores of the other groups of participants were nearly at ceiling, no adjustments to the induction scores were made in the older age groups.

Category-Based Reasoning Performance

Participants' reasoning scores were submitted to a one-way ANOVA with age as the between-subject factor. This analysis revealed a significant effect of age, $F(2, 82)=16.49$, $p<0.0001$. This effect was further explored through planned comparisons.

Category-Based Reasoning performance improved as a function of age; see Figure 4. Posthoc Tukey tests revealed that both kindergarten and first-grade children exhibited superior performance on the Category-Based Reasoning task compared to the preschoolers; all $ps<.0001$. However, there was no significant difference in performance on the Category-Based Reasoning task between the kindergarten children and the first-grade children; $p=.90$.
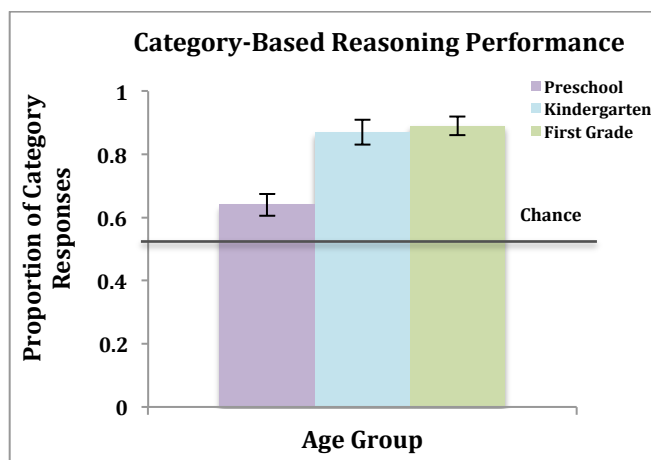

Category-Based Reasoning Performance

Figure 4: Proportion of category-based responses by age group. Error-bars represent standard errors of the means. Line indicates chance performance.

The final analysis compared the mean Category-Based Reasoning scores to chance (.50) using single sample t-tests. Participants in all age groups exhibited Category-Based Reasoning performance that was significantly above chance; Preschoolers: $M=.64$, $SD=.22$; Kindergartners: $M=.87$,

$SD=.18$; First-graders: $M=.89$, $SD=.14$; all $ts>3.67$, all $ps<0.0001$. The rate of category-based responding in preschool-age children was somewhat higher than in our prior studies ($M=.54$ across Fisher et al., 2011; Godwin et al., in press, Matlen, Fisher & Godwin, under review). However, it should be noted that in the present study the sample of preschool children was recruited entirely from a laboratory campus school at a private university and our prior research utilized more diverse community-based samples.

Is Category-based Reasoning Related to Children's Semantic Space Organization?

A correlational analysis was conducted to examine the potential relationship between children's Category-Based Reasoning performance and their performance on the Semantic Space task. This analysis revealed a significant positive correlation between the Semantic Space Difference scores and Category-Based Reasoning scores when scores were aggregated across preschoolers, kindergarteners, and first-graders, $r=.484$, $p<0.0001$ (see Figure 5). When separated by age group, there was a significant correlation in the preschool group $(r=.473$; $p=0.002)$, a marginally significant correlation in the kindergarten group $(r=.34$, $p=0.12)$, and no correlation among first-graders $(r=.10$, $p=0.66)$.

It is perhaps not surprising that the magnitude of the correlation between the Semantic Space Difference scores and Category-Based Reasoning scores decreased with age, as children's performance on both tasks improved and variability in performance decreased (e.g., many children achieved ceiling scores on the category-based reasoning task by first-grade). However, it is noteworthy that in preschool-age children, who exhibited a high degree of variability on both tasks, there was a fairly strong relationship between children's category-based reasoning and semantic organization.
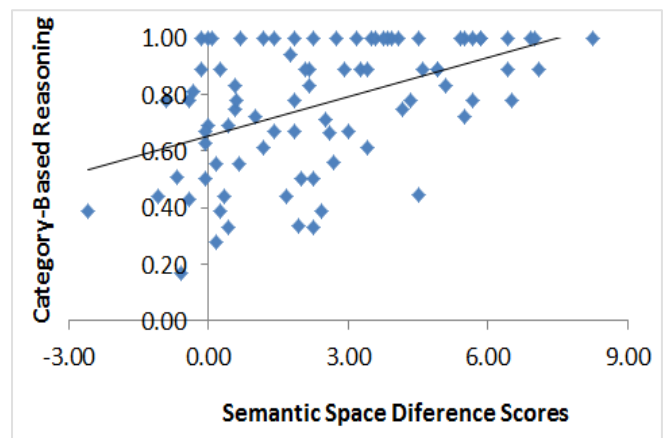

Figure 5: A scatterplot of children's Category-Based Reasoning scores and their Difference scores on the Semantic Space task.

Discussion

The results from the present study point to several novel findings. First, the new paradigm designed to measure development of semantic organization successfully captured increased differentiation among the animal concepts during the preschool and early school years. The gradual increase in the Semantic Space Difference scores in preschoolers, kindergarteners, and first-graders suggests that children increasingly become more sensitive to semantic similarity.

Second, the present findings provide preliminary evidence that individual differences in knowledge organization, as measured by the Semantic Space task, may be related to developmental differences in category-based reasoning.

In conclusion, these findings indicate that children's semantic knowledge undergoes gradual reorganization across development. Additionally, performance on this measure was found to predict preschoolers' inductive generalization performance. This latter finding suggests that the ability to make inductive inferences based on categories may be related to improvements in semantic organization.

Acknowledgments

References

Arias-Trejo, N. & Plunkett, K. (2009). Lexical-semantic priming effects during infancy. *Philosophical Transactions of the Royal Society Biological Sciences,* 364, 3633-3647.

Badger, J. R. & Shapiro, L. R. (2012). Evidence of a transition from perceptual to category induction in 3- to 9-year-old children. *Journal of Experimental Child Psychology,* 113, 131-146.

Chi, M. T. H., Hutchinson, J., & Robin, A. F. (1989). How inferences about novel domain-related concepts can be constrained by structured knowledge. *Merrill-Palmer Quarterly,* 35: 27-62.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test-Third Edition: Manual.* Circle Pines, MN: American Guidance Services.

Fisher, A.V. (2010). What's in the name? Or how rocks and stones are different from dogs and puppies. *Journal of Experimental Child Psychology,* 105, 198-212.

Fisher, A.V., Matlen, B.J., & Godwin, K.E. (2011). Semantic similarity of labels and inductive generalization: Taking a second look. *Cognition,* 118, 432-438.

Fisher, A. V. & Sloutsky, V. M. (2005). When induction meets memory: Evidence for gradual transition from similarity-based to category-based induction. *Child Development,* 76(3), 583-597.

Gentner, D., Rattermann, M. J., Markman, A. B., & Kotovsky, L. (1995). Two forces in the development of relational similarity. In T. J. Simon & G. S. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 263-313). Hillsdale, NJ: LEA.

Gobbo, C., & Chi, M. T. H. (1986). How knowledge is structured and used by expert and novice children. *Cognitive Development,* 1: 221-237.

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

Goswami, U. & Brown, A.L. (1990). Higher-order structure and relational reasoning: Contrasting analogical and thematic relations. *Cognition,* 36, 207-226.

Godwin, K., Fisher, A., & Matlen, B. (In press). Development of category-based reasoning in 4- to 7-year-old children: The influence of label co-occurrence and kinship knowledge. *Journal of Experimental Child Psychology.*

Godwin, K., Matlen, B., & Fisher, A. (2012). Development of Category-Based Reasoning in Preschool-Age Children: Preliminary Results of a Longitudinal Study. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 1608-1613). Austin, TX: Cognitive Science Society.

Goldstone, R. (1994). An efficient method for obtaining similarity data. *Behavior Research Methods, Instruments, & Computers,* 26, 381-386.

Hayes, B., Heit, E., & Swendsen, H. (2010). Inductive reasoning. *Wiley Interdisciplinary Reviews: Cognitive Science, 1,* 278-292.

Howard, D.V. & Howard, J.H., Jr. (1977). A multidimensional scaling analysis of the development of animal names. *Developmental Psychology,* 13, 108-113.

Karmiloff-Smith, A. (1984). "Children's problem solving". In M.E.Lamb, A.L.Brown & B.Rogoff (Eds.) *Advances in Developmental Psychology,* Vol.III. New Jersey: Erlbaum, 39-90.

Kemp, C. & Tenenbaum, J. B. (2008). Structured models of semantic cognition. *Behavioral and Brain Sciences.* 31(6), 717-718.

McCauley, C., Weil, C. M., & Sperber, R. D. (1976). The development of memory structures as reflected by semantic-priming effects. *Journal of Experimental Child Psychology,* 22, 511-518.

Opfer, J. E., & Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cognitive Psychology, 55,* 169-195.

Proffitt, J.B. Coley, J.D., & Medin, D. (2000). Expertise and category-based induction. *Journal of Experimental Psychology: Learning Memory, and Cognition,* 26(4), 811-828.

Rattermann, M. J., & Gentner, G. (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.

Rogers, T. T. & McClelland, J. L. (2004). *Semantic Cognition: A Parallel Distributed Processing Approach.* Cambridge, MA: MIT Press.

Saltz, E., Seller, E., & Sigel, I. E., (1972). The development of natural language concepts. *Child Development,* 43, 1191-1202.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. *Psychological Science, 14,* 237-243.

Sloutsky, V. M., Kloos, H., & Fisher, A. V. (2007). When looks are everything: Appearance similarity versus kind information in early induction. *Psychological Science,* 18(2), 179-185.