Title
Development of Category-Based Reasoning: Results from a Longitudinal Study

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 35(35)

ISSN
1069-7977

Authors
Godwin, Karrie
Fisher, Anna
Matlen, Bryan

Publication Date
2013

Peer reviewed

Development of Category-Based Reasoning: Results from a Longitudinal Study

Karrie E. Godwin (kegodwin@andrew.cmu.edu)
Anna V. Fisher, (fisher49@andrew.cmu.edu)
Bryan J. Matlen (bmatlen@cmu.edu)
Carnegie Mellon University, Department of Psychology, 5000 Forbes Avenue, Pittsburgh, PA 15213 USA

Abstract

Prior research on the development of category-based reasoning indicates a protracted developmental course of this ability as well as a high degree of individual variability. However, the sources of this individual variability as well as the sources of developmental change remain unclear. The present study aimed to examine these issues, with a focus on the role of representational change and executive function development. Across two time points spaced approximately 7 months apart, children's category-based reasoning was assessed along with a battery of executive function and representational change measures. Results replicated prior work in that only a small proportion of children exhibited spontaneous category-based reasoning at Time1, and this proportion increased with development. In addition, both executive function and representational change were found to predict the development of category-based reasoning.

Keywords: Category-based reasoning; inductive reasoning

Introduction

Category-based reasoning is central to mature cognition and underlies much of our learning and functioning in the world (e.g., Osherson et al., 1990; Sloman, 1993). Despite early reports that even very young children spontaneously engage in category-based reasoning (e.g., Gelman & Markman, 1986; Gelman & Coley, 1990; Welder & Graham, 2001), recent evidence suggests that development of category-based reasoning follows a relatively protracted developmental course (e.g., Badger & Shapiro, 2012; Godwin, Matlen, & Fisher, in press; Fisher, Matlen, & Godwin, 2011; Fisher, 2010; Fisher & Sloutsky, 2005).

One of the hallmarks of category-based reasoning is one's ability to make inferences based on the knowledge that two (or more) items belong to similar kinds in the absence of supporting perceptual information. For example, if one is shown a picture of a rock, a sponge, and another rock and asked to predict which two items have properties in common, one could rely on perceptual similarity to make an inference. Similarly, if the pictures are ambiguous (or not presented) and labels are used to indicate category membership, one could base their inference on matching labels (e.g., rock–rock), not necessarily because one understands that labels refer to kinds, but because the labels are perceptually identical (Sloutsky & Fisher, 2004). However, one's ability to rely on semantically-similar labels (e.g., rock-stone) to make inferences is commonly interpreted as an index of category-based reasoning (e.g., Gelman & Markman, 1986).

Several studies have documented that the ability to spontaneously engage in category-based reasoning appears between 4 and 6 years of age (e.g., Badger & Shapiro, 2012; Godwin, et al., in press; Fisher, et al., 2011; Fisher 2010; Fisher & Sloutsky, 2005). However, it remains unclear what leads to the development of spontaneous category-based reasoning during the preschool years. Two classes of explanations have been put forth to explain changes in various areas of cognitive development, namely Representational Change and Executive Function development. We briefly discuss both explanations below.

Representational Change

Representational change is "reorganization of existing knowledge or a difference in the utilization of information, rather than the acquisition of new information" (Nelson, 1977, p. 109). Representational change has been implicated as an explanatory factor in several areas of cognitive development, including analogical reasoning (e.g., Gentner et al., 1995), problem solving (e.g., Karmiloff-Smith, 1984), and numerical development (e.g., Opfer & Siegler, 2007).

With regards to semantic development, there are several compelling sources of evidence pointing to representational change in the multidimensional scaling literature (e.g., Howard & Howard, 1977), free association studies (e.g., Brown & Berko, 1960), and development of semantic priming (e.g., McCauley, Weil, & Sperber, 1976).

Furthermore, different approaches to modeling semantic cognition suggest that early conceptual organization is fairly undifferentiated (such that *penguin*, *trout*, and *alligator* may start out as belonging to the same cluster) with greater differentiation emerging with development (Kemp & Tenenbaum 2008; Rogers & McClelland 2004). At present, there is no direct empirical evidence testing these predictions, although Carey's (1985) seminal work is largely consistent with these developmental profiles.

Executive Functions

Executive Functions (EF) are psychological processes thought to control other (typically, higher-order) psychological processes such as planning, reasoning, and problem-solving. Most researchers distinguish the following EF processes: set shifting, active maintenance of representations (sometimes referred to as working memory), and inhibitory control (Bunge et al., 2002; Carlson et al., 2002).

The EF system is traditionally associated with prefrontal cortex, which is believed to be one of the slowest brain regions to mature (e.g., Diamond, 2002). Development of EF has been implicated in developmental accounts of category learning (Sloutsky, 2010), and there is evidence that representation maintenance and inhibitory control play a role in the development of analogical reasoning (Morrison et al., 2011). With regards to semantic cognition, regions of the PFC (specifically, ventrolateral PFC or VLPFC) have been shown to be engaged in controlled semantic access, for instance in classification or category generation tasks. While it remains unclear whether the role of VLPFC is to bias retrieval of task-relevant semantic information through maintaining task representations or to select task-relevant representations among competing activated representations (e.g., Wagner, 2002; Kan & Thompson-Schill, 2004), prefrontal cortex clearly is important for controlled semantic access.

The Present Study

Prior research on the development of category-based reasoning indicates not only a protracted developmental course of this ability, but also a high degree of individual variability. Specifically, results aggregated across several studies suggest that approximately 20% of 4-year-olds spontaneously make category-based inferences with semantically-similar labels, and this proportion increases to approximately 40% and 65% among 5- and 6-year-olds, respectively (Fisher, 2010; Fisher et al., 2011; Godwin et al., in press). However, the sources of this individual variability as well as sources of developmental change remain unclear. The goal of the present study was to begin the exploration of these questions, with a focus on the putative role of representational change and executive functions. A battery of assessments was administered over the course of one school year. At Time1 (Fall) we collected measures of children's category-based reasoning, verbal working memory, IQ, and semantic knowledge organization. At Time2 (Spring) we collected measures of children's category-based reasoning, semantic knowledge organization, inhibitory control, non-verbal working memory, sustained attention, and category generation.

Method

Participants

Participants in this study were 43 four-year-old children from a local preschool (*Mage*=4.32 years, *SD*=0.28 years, 20 females, 23 males).

Design and Procedure

Each child participated in 13 sessions over the course of the school year (6 sessions at Time1 and 7 sessions at Time2). Children were tested individually in a quiet room adjacent to their classroom by a trained research assistant. A brief description of the task battery is provided below.

Category-Based Reasoning Task

This task included 9 label triads, 3 of which referred to artifacts, 3 to inanimate natural kinds, and 3 to animate natural kinds (see Table 1). All triads contained a target item, category-choice, and an unrelated lure (e.g., *rat-mouse-fish*). Visual stimuli consisted of sets of three identical doors. Children were told that objects were hiding behind doors. The objects were never revealed in order to encourage children to rely on the category information conveyed by the labels.

Table 1: Category-Based Reasoning Task Linguistic Stimuli

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

Children were first told what objects were hiding behind the doors and then told about a novel property of the target item (e.g., "*The rock has higa inside*"). Children were asked to generalize the novel property from the target item to either the category-choice or the unrelated lure. The task was administered four times (twice within each time point) in order to obtain a more stable estimate of children's performance. The delay between task administrations within a time point was one to two weeks. The trials were administered in one of two counter-balanced orders.

Picture Identification Task

The picture identification task is similar in format to the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997). It assessed children's familiarity with the linguistic stimuli utilized in the category-based reasoning task. The task was administered at Time 1 and Time 2.

Intelligence Test

The *Wechsler Preschool and Primary Scale of Intelligence* (WPPSI) provided a measure of children's general intelligence (Full-scale IQ or FSIQ), as well as an index of children's Verbal IQ (VIQ), Performance IQ (PIQ), and Processing Speed Quotient (PSQ). The WPPSI was administered at Time 1 only.

Semantic Space Task

The semantic space task served as a measure of children's semantic organization. The stimuli entailed a game board (9x9 grid; see Figure 1) and 2 game pieces (1" wooden cubes). Verbal stimuli included 24 animal pairs: 6 semantically-similar dyads (e.g., *chick-hen*), 6 dyads that share a common setting or habitat (e.g., *chick-goat*), 6 dyads that are unrelated (e.g., *chick-goldfish*), and 6 filler dyads (see Table 2). Thus, throughout the task, the target animal was paired with 3 different test items (i.e., the category-choice, setting/habitat match, and the unrelated item). The 3 animal trials from the Category-Based reasoning task were included in the Semantic Space task.

Children were told that they were helping Zibbo the zookeeper organize his zoo. Children were instructed to put animals of the same kind close together on the board. For each trial, the experimenter put one of the game pieces on a predetermined square on the game board and told the child that the specified location was where Zibbo put the target animal (e.g., "*The zookeeper put the chick here*"). The child was then asked to identify where the test item (i.e., the second game piece) should be placed (e.g., "*Where do you think the goat should go?*"). After each trial the child's response was recorded in order to calculate the distance between the target and the placement of the test item. Prior to playing the game, children were given two examples where the experimenter demonstrated that animals of similar kind (e.g., *bunny* and *rabbit*) should be placed close together and animals of different kind (e.g., *dog* and *shark*) should be placed far apart.

Table 2: List of 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			

Figure 1: Schematic depiction of the semantic space game board. Red squares mark the location of critical trials; yellow squares mark the location of filler trials. Note that none of the locations were colored when the task proper was administered.

Children's score on the semantic space task was calculated by averaging habitat dyads and unrelated dyads together in order to create a composite score for non-semantically-similar dyads. Next, the average score for semantically-similar dyads was subtracted from the non semantically-similar composite score to obtain a difference score. Difference scores approaching zero indicate that children did not differentiate the placement of semantically-similar and dissimilar dyads. Difference scores above zero indicate that children placed semantically-similar dyads closer than dissimilar dyads. The semantic space task was administered at Time1 and Time2.

EF Measures

Verbal Working Memory Tasks Forward and Backward word-span tasks were administered to assess children's verbal working memory capacity. Verbal stimuli consisted of 60 common count nouns selected from the MacArthur Communicative Development Inventory (Dale & Fenson, 1996). In the Forward-word span task, children were asked to recite the words in the same order in which they were presented; in the Backward-word span task children were asked to repeat the words in the opposite order. If a child made a mistake, they were given another opportunity to recite a different list of the same word length. Children's score was determined by the longest list length the child was able to recite correctly.

Inhibitory Control Measures Two common measures of response inhibition were included in the assessment battery: the Day-Night task (Gerstadt et al., 1994) and the Flanker task (Rueda et al., 2004). In the Day-Night Task, children were shown a set of cards depicting the sun and the moon. Children were asked to provide a verbal response that conflicts with the presented image (e.g., if the child was shown a picture of the sun, the correct response would be "*night*". Conversely if the child was shown a picture of the moon, the correct response would be "*day*"). The task consisted of 16 trials (the moon and the sun were presented 8 times each). Two presentation orders were created: The trials were randomized for order 1 and the sequence was reversed for order 2.

We used the version of the Flanker Task adapted for use with young children (Rueda et al., 2004). In this version children are presented with arrays of fish on a computer screen. Children are asked to feed the center fish by pressing either the left or right button. The correct response is dependent upon the direction the center fish is facing. The center fish is surrounded by four other fish (two on each side). The surrounding fish may be congruent (e.g., swimming in the same direction as the center fish) or incongruent (e.g., swimming in the opposite direction as the center fish). Neutral trials were also presented in which the central fish appears in isolation (i.e., not flanked by other fish). A total of 48 trials were administered: 16 neutral trials, 16 incongruent trials, and 16 congruent trials. For the purposes of the analyses reported below, we used the Flanker Accuracy Difference score (calculated by subtracting each child's accuracy for the Incongruent trials from the Neutral trials) and Flanker RT Difference score (calculated by subtracting each child's reaction time for the Incongruent trials from the Neutral trials).

Non-Verbal Working Memory & Sustained Attention The Track-It Task (Fisher et al., 2012) was used as an index of non-verbal working memory; this task also provided a measure of sustained attention. In this task children watched a set of moving objects: six distractors and one target. The objects moved randomly across a computer screen for 10 seconds, and then disappeared. On each trial, children were asked to select the location where the target object disappeared; the location questions provided a measure of sustained attention. Upon answering the location question, children were shown a laminated card that contained an array of 9 objects (the target object and 8 lures). Children were asked to point to the target object that they had been tracking; children's responses to this question provided a measure of non-verbal working memory (WM). The Track-it task included 10 experimental trials and one practice trial.

Results

Picture Identification The results from the Picture Identification task suggested that children were familiar with the labels used in the category-based reasoning task: Children's accuracy on this task approached ceiling levels (*M*=92%, *SD*=14% and *M*=96%, SD=8% for Time 1 and 2 respectively). As an additional precaution, children's 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 their category-based reasoning score.

Category-Based Reasoning Task Mean category-based reasoning scores at Time 1a and 1b were very similar (adjusted means: 0.63 and 0.66, respectively) and significantly correlated (*r*=.483, *p*=0.001). Mean category-based reasoning scores at Time 2a and 2b were also similar (adjusted means: 0.73 and 0.80, respectively) and significantly correlated (*r*=.689, *p*=0.0001). Consequently, induction scores were averaged across Time 1a and 1b and across Time 2a and 2b to yield average category-based reasoning scores for Time1 (*M*=0.64, *SD*=0.22) and Time2 (*M*=0.76, *SD*=0.21).

The rate of category-based responding at Time1 (*M*=.64) was above chance ($t(40)=4.08$, $p<.001$) and somewhat higher than in our prior studies (*M*=.54 across Fisher et al., 2011; Godwin et al., in press; Matlen et al., under review). However, it should be noted that in the present study the sample consisted entirely of children enrolled in a laboratory campus school at a private university, and prior studies utilized more diverse community-based samples.

The proportion of category-based responding at Time2 (*M*=.76) was also above chance, ($t(41)=8.01$, $p<.001$) and higher than at Time1, paired-samples $t(39)=3.53$, $p<.001$, Cohen's d = .56. Note that the latter finding cannot be attributed simply to children having experience with performing the same task, as scores were not significantly different at Time 1a (.63) and 1b (.66). Therefore, this finding points to a developmental increase in the propensity towards category-based reasoning.

To investigate individual patterns of responses, participants were classified as either category-based or non-category-based responders. A category-based responder was defined as a participant who gave a category-based response on at least 7 out of 9 (78%) trials (binomial *p*=0.09). Analysis of the individual patterns revealed that only a small percentage of children were classified as category-based responders at Time1 (27%). In contrast, the majority of children were classified as category-based responders at Time 2 (67%). The association between responder type and testing point (Time1 vs. Time2) was significant, McNemar's $\chi^2(1)=7.22$, $p<.005$.

Predicting Category-Based Reasoning We performed linear stepwise regression to identify the best predictors of category-based responding at Time1 and 2. Only predictors that were significantly correlated with category-based responding were entered into the model. Thus, three predictors were included: Semantic Space scores Time1, FSIQ, and Non-Verbal WM score. Overall, the model significantly predicted children's responses on the category-based reasoning task, $R^2=.211$, $F(1)=11.42$, $p=.002$. However, only one predictor was found to be significantly related to children's induction performance at Time 1: Semantic Space scores ($\beta=.481$, $t(1)=3.38$, $p=.002$).

For predicting category-based responding at Time2, only predictors that were significantly correlated with induction performance at Time2 were included in the model. The following predictors were entered into the model: Semantic Space Time2 scores, Forward Word-Span, Backward Word-Span, FSIQ, Non-Verbal WM score, Sustained Attention score, and Day/Night score (VIQ was excluded from the analysis due to concerns regarding colinearity based on its high correlation with FSIQ). Overall, the two-predictor model significantly predicted children's responses on the category-based reasoning task, $R^2=.474$, $F(2)=19.00$, $p<.001$. However, only two predictors were found to be significantly related to children's induction performance at Time 2: Non-verbal WM ($\beta=.522$, $t(2)=4.37$, $p<.0001$) and Day/Night scores ($\beta=.352$, $t(2)=2.95$, $p=.005$).

Category-Based Reasoning: What Develops? What factors play a role in the development of category-based reasoning? Since several children performed at nearly ceiling level on the category-based reasoning task at Time 1, it was not possible to address this question using gain scores from Time1 to Time2. Therefore, to address this question we split the sample into three groups based on the children's performance on the category-based reasoning task at Time 1 and Time2 (see Figure 2). Group 1 included children who were already category-based responders at Time1 (27% of the sample); all of these children remained category-based responders at Time2. Group 2 included children who were not yet category-based responders at Time1 but became category-based responders at Time2 (40% of the sample). Group 3 included children who were not yet category-based responders at either Time1 or Time2 (32.5% of the sample). Splitting the sample in this manner allowed for analyses examining potential factors that may differentiate Groups 2 and 3 (i.e., children who became category-based responders at Time2 from children who were not yet category-based). Three children were missing scores for either Time 1 or 2 and were omitted from this analysis.

Importantly, children in Groups 2 and 3 obtained comparable FSIQ scores (*M*=108, *M*=100 respectively), $t(27)=1.41$, *ns*. This finding suggests that performance differences between the two groups on the category-based reasoning task were not simply a result of disparities in children's general intelligence. Children in Group 1 obtained FSIQ scores (*M*=118) that were over one standard deviation above the population mean (*M*=100, *SD*=15;

Wechsler, 2002). FSIQ scores of children in Group 1 were also significantly higher than children in Group 3, $t(21)=3.04$, $p<.01$, and marginally higher than those of children in Group 2, $t(25)=1.78$, $p=.088$. Based on IQ scores, the group of children who were already category-based responders on the induction task at Time1, were cognitively advanced. However, it is important to note that children in Groups 2 and 3 were not lagging behind as they exhibited average intelligence compared to the general population.

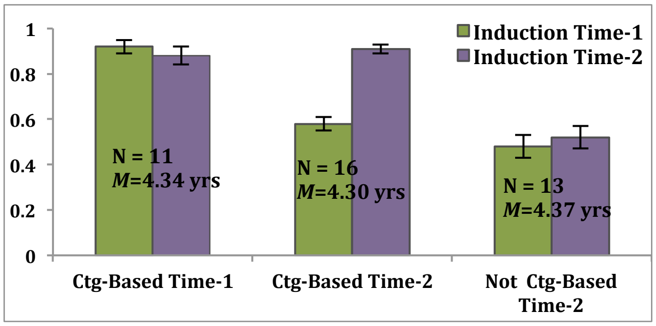

Figure 2. Proportion of category-based responses at Time 1 and Time 2 by group.

Due to space limitations we cannot describe the performance patterns for these three groups on all administered tasks. However, it should be noted that on some measures children's performance was equivalent across groups. For example, children's Flanker Accuracy Difference scores were comparable in all three groups ($M_{Group1}=.35$, $M_{Group2}=.38$, $M_{Group3}=.30$, all *ts*<.75, *ns*). Similarly, Flanker RT Difference scores were comparable in all three groups ($M_{Group1}=-127.37$, $M_{Group2}=-116.24$, $M_{Group3}=-106.48$, all *ts*<.21, *ns*). The biggest performance differences were found on three measures: Semantic Space, Non-Verbal WM, and the Day/Night task.

Recall that the Semantic Space task was administered twice, once at Time1 and again at Time2. Therefore, we were able to compare children's performance on this task across time. As can be seen in Figure 3, children in Group 1 exhibited equivalently high performance on the Semantic Space task at both Time1 and Time2 ($M_{T1}=2.41$, $M_{T2}=3.04$, paired-sample $t(10)=.73$, *ns*). In contrast, children in Group 2 significantly improved in their performance on the Semantic Space task from Time1 to Time2 ($M_{T1}=1.4$, $M_{T2}=3.01$, paired-sample $t(15)=2.38$, $p=.03$). Children in Group 3 exhibited relatively low performance on the Semantic Space task at both Time1 and Time2 ($M_{T1}=.88$, $M_{T2}=1.15$, $t(12)=.59$, *ns*).

Overall, these findings suggest that children who showed consistently high performance on the Semantic Space task also showed consistently high performance on the category-based reasoning task, and children who showed consistently low performance on the Semantic Space task also showed consistently low performance on the category-based reasoning task. Only those children who showed improved performance on the Semantic Space task also showed improved category-based reasoning.

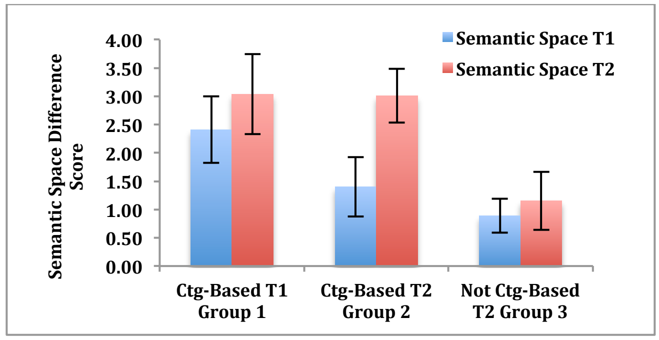

Figure 3: Mean performance on the Semantic Space task at Time 1 and Time 2 by group.

A stark difference was observed in children's non-verbal WM performance. Children in Groups 1 and 2 demonstrated similar levels of performance on the non-verbal WM task ($M=.70$ and $M=.62$ respectively; $t(25)=.67$, *ns*) with children in both groups demonstrating better non-verbal WM than children in Group 3 ($M=.18$); all *ts*>4.66, $p<.0001$). A similar pattern of results was obtained for the Day/Night task as the mean accuracy rate was superior for Groups 1 and 2 ($M=.71$ and $M=.79$ respectively) compared to Group 3 ($M=.47$); all *ts*>1.84, $p<.08$.

Taken together these findings corroborate the results obtained from the regression models suggesting that developmental improvement in working memory, inhibitory control, and semantic differentiation underlies the development of category-based reasoning.

Discussion

One potential limitation of the present study is that the assessment battery did not include a direct measure of vocabulary size. Arguably, vocabulary may be a precursor to category-based reasoning. Nevertheless, there is reason to believe that the high degree of individual variability observed in preschool children's category-based reasoning performance is unlikely to be explained by differences in vocabulary. It is important to note that FSIQ is a composite measure that includes an index of children's verbal ability. Recall that FSIQ scores were comparable between children who became category-based responders at Time 2 and children who were not yet category-based. Additionally, children's high accuracy rates on the Picture Identification task suggests that children are familiar with the labels that were utilized in the category-based reasoning task.

In line with prior work (Fisher et al., 2011; Godwin et al. in press), the present study provides additional evidence demonstrating that only a small percentage of preschool-age children spontaneously engage in category-based reasoning. Additionally, this work implicates three cognitive factors in the development of young children's category-based reasoning: representational change, working memory, and inhibitory control.

First, the strong relationship between improvements in semantic differentiation scores and induction scores suggests that representational change may be one underlying mechanism in the development of children's ability to engage in spontaneous category-based reasoning. Second, working memory may be an important cognitive factor in the development of category-based reasoning as children need to maintain and manipulate the task-relevant information in working memory. Finally, sufficiently developed inhibitory control may be required to select task-relevant representations among competing activated representations. These findings indicate that both general cognitive advances (EF and working memory) and changes in domain-specific knowledge (representational change) contribute to the advancement of category-based reasoning.

Acknowledgments

We thank Malika Sinha, Laura Pacilio, Alyssa Montanaro, Anna Loiterstein, Like Li, Kayoung Joung, and Rachel Walsh for their help collecting data. We thank the children, parents, and teachers for making this work possible. This work was supported by a Graduate Training Grant awarded to Carnegie Mellon University by the Department of Education (R305B090023 & R305B040063).

References

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.

Brown, R. & Berko, J. (1960). Word association and the acquisition of grammar. *Child Development*, 31, 1-14.

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. E. (2002). Immature frontal lobe contributions to cognitive control in children: Evidence from fMRI. *Neuron, 33*, 301-311.

Carey, S. (1985). *Conceptual Change in Childhood*. Cambridge, MA: Bradford Books, MIT Press.

Carlson, S, M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development, 11*, 73 – 92.

Dale, P.S., & Fenson, L. (1996). Lexical development norms for young children. Behavior Research Methods, Instruments, & Computers, 28, 125-127.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D.T. Stuss & R.T. Knight (eds.), *Principles of Frontal Lobe Function* (p. 466-503). London, UK: Oxford University Press.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test–Third Edition*. Circle Pines, MN: AGS Publishing.

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.

Fisher, A. V., Thiessen, E., Godwin, K., Kloos, H., & Dickerson, J. P. (2012). Mechanisms of focused attention in 3- to 5-year-old children: Evidence from a new object tracking task. *Journal of Experimental Child Psychology,* 114(2), 275-294.

Gelman, S.A. & Coley, J.D. (1990). The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology, 26*, 796-804.

Gelman, S.A., & Markman, E. (1986). Categories and induction in young children. *Cognition*, *23*, 183-209.

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.

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of 3.5 – 7 years old on a stroop-like day-night test. *Cognition, 53*, 129-153.

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*.

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

Kan, I. P. & Thompson-Schill, S. L. (2004). Selection from perceptual and conceptual representations. *Cognitive, Affective, & Behavioral Neuroscience, 4,* 466-482.

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.

Matlen, B.J., Fisher, A.V., & Godwin, K.E. (under review). Mechanisms of inductive generalization: The influence of co-occurrence probability.

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

Morrison, R. G., Doumas, L. A. A., & Richland, L. E. (2011). A computational account of children's analogical reasoning: Balancing inhibitory control in working memory and relational representation. *Developmental Science, 14,* 516-529.

Nelson, K. (1977). The syntagmatic-paradigmatic shift revisited: A review of research and theory. *Psychological Bulletin*, 84, 93-116.

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

Osherson, D., Smith, E., Wilkie, O., Lopez, A. & Shafir, A. (1990). *Category-based induction. Psychological Review, 97,* 185 – 200.

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

Sloman, S.A. (1993). Feature-based induction. *Cognitive Psychology, 25*, 231 – 280.

Sloutsky, V.M. (2010). From perceptual categories to concepts: What develops? *Cognitive Science, 34*, 1244–1286.

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.

Wagner, A.D. (2002). Cognitive control and episodic memory: Contributions from prefrontal cortex. L. R. Squire & D. L. Schacter (Eds.). *Neuropsychology of Memory (3rd ed.),* pp. 174-192. New York: Guilford Press.

Welder, A.N., & Graham, S.A. (2001). The influences of shape similarity and shared labels on infants' inductive inferences about nonobvious object properties. *Child Development*, *72*, 1653-1673.

Wechsler, D. (2002). WPPAI-III Technical Manual. San Antonio, TX: Pearson, Inc.