Inductive Generalization in Early Childhood: The Contribution of Perceptual and Representational Similarity

Karrie E. Godwin (kegodwin@andrew.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

Inductive generalization is ubiquitous in human cognition; however, the factors underpinning this ability early in development remain contested. Two alternative perspectives have been proposed for how children make inductive inferences: a naïve theory account (Gelman & Markman, 1986; Markman, 1990) and a similarity-based account (Sloutsky & Fisher, 2004; 2012). Although both theories claim considerable empirical support, the debate is ongoing and results of extant studies are often deemed inconclusive. We report an experiment designed to evaluate the predictions of each account. In this study, 2- to 5-year-old children were asked to make inferences about highly familiar object categories. The reported findings are not fully consistent with either the naïve theory or the similarity-based approach. Therefore, we propose a revised version of the similarity-based account, which can account for the reported findings.

Keywords: inductive reasoning; categories; representations; cognitive development.

Introduction

The ability to generalize from the known to the unknown is a critical component of human cognition. For example, by 10 months of age infants are able to generalize observable object properties (e.g., Baldwin, Markman & Melartin, 1993), by 24 months of age children can generalize labels to novel objects (e.g., Booth & Waxman, 2002; Jones & Smith, 1998; Smith, Jones, & Landau, 1996), and during the preschool years children begin to make inductive inferences about unobservable properties (e.g., Fisher, Matlen, & Godwin, 2011; Gelman & Markman, 1986; Sloutsky & Fisher, 2004). Despite general agreement regarding the importance of inductive generalization for human cognition, there is little agreement regarding the developmental origins of this ability.

Two alternative perspectives have been proposed for how children make inductive inferences: a naïve theory account (Gelman & Markman, 1986; Markman, 1990) and a similarity-based account (Sloutsky & Fisher, 2004; 2012). According to the naïve theory approach, from very early in development people first identify *category membership* of items under consideration and then generalize a known property to items of the same kind: "by 2 ½ years, children expect categories to promote rich inductive inferences… and they can overlook conflicting perceptual appearances in doing so" (Gelman & Coley, 1990, p. 802). Furthermore, it has been suggested that the ability to make category-based inferences is not a product of development and learning. Instead, children are "initially biased" to recognize that labels denote categories and make inferences on the basis of shared category membership (Gelman & Markman, 1986, p. 207), an idea that has been highly influential in the literature (e.g., Booth & Waxman, 2002; Gelman & Coley, 1990; Jaswal & Markman, 2007; Kalish, 2006; Keil, 1989).

In contrast to the two-step account of inductive inference suggested by the naïve theory approach, Sloutsky and Fisher (2004) proposed a one-step similarity-based account called SINC (Similarity, Induction, Naming, and Categorization). According to SINC, children make inferences on the basis of the overall similarity of presented entities computed over all perceived object features. Within this approach, labels are considered to be object features (rather than category markers) that contribute to the overall perceptual similarity. Therefore, according to SINC an inference can be label-based without necessarily being category-based. Several findings suggest that children rely primarily on perceptual features of objects (but not category membership information) to make inferences well beyond the preschool years, possibly until 7 to 9 years of age (e.g., Badger & Shapiro, 2012; Fisher & Sloutsky, 2005; Sloutsky, Kloos, & Fisher, 2007).

Evidence in support of the naïve theory of inductive generalization stems from the seminal study by Gelman and Markman (1986). In this study researchers asked preschool-age children and college students to make inferences about natural kinds when perceptual information was ambiguous or conflicted with category membership (cf. Sloutsky & Fisher, 2004). Labels were used to communicate category information; for instance, participants were asked whether a *rock* shared a non-obvious property with a *stone* or *chalk*. The overall rate of category match choices was above chance, both in preschool children and college students. These findings were taken as evidence that even young children hold a belief (or a naïve theory) that natural kind objects share a number of unobservable properties if they belong to the same category, and make inductive inferences on the basis of this belief. Subsequent studies reported similar findings in younger children and infants (e.g., Gelman & Coley, 1990; Graham, Kilbreath, & Welder, 2004).

The similarity-based approach explains these findings through the contribution of the similarity of auditory features (i.e., linguistic labels in this case) to inductive

inference. A mathematical model based on the SINC account successfully captured the pattern of findings reported by Gelman and Markman (1986). Specifically, when visual features of the stimuli were ambiguous (e.g., the target matched one of the test objects on the shape dimension and the other test object on the texture and color dimensions; for details see Fisher, 2007), identical auditory features (such as linguistic labels) dramatically increased the perceptual similarity between pairs of objects. Thus, the same set of findings can have very different interpretations, which contributes to the current theoretical stalemate (Fisher, 2007; see Smith & Samuelson, 2006 for related arguments). However, this debate can be advanced (if not resolved) by removing linguistic labels from the paradigm. Specifically, if highly familiar and readily identifiable objects are used as stimuli, labels are not necessary to communicate object kind. The present study was designed to implement this solution.

Experiment 1

Method

Participants

Participants were 18 five-year-olds (*Mage* = 5.46 years, *SD* = 0.34 years, 9 females, 9 males), 21 four-year-olds (*Mage* = 4.44 years, *SD* = 0.26 years, 10 females, 11 males), 18 three-year-olds (*Mage* = 3.65 years, *SD* = 0.28 years, 10 females, 8 males), and 6[1] two-year-olds (*Mage* = 2.59 years, *SD* = 0.12 years, 3 females, 3 males). Participants were recruited from local schools, preschools, or the Phipps conservatory in Pittsburgh, Pennsylvania. Children were tested individually by trained research assistants.

Design and Procedure

Visual Stimuli. The visual stimuli included 14 triads displayed on a computer screen: 7 triads referred to artifacts and the remaining 7 triads referred to animals (see Figure 1). Item selection was based on a Familiarity Calibration (described below), which ensured that children of this age group could readily label the stimuli using common basic level labels[2].

Familiarity Calibration. A separate group of preschool children (N = 10, M_{age} = 4.58 years, SD = 0.69) participated in the calibration. The calibration consisted of a basic naming task: Participants were presented with a series of pictures displayed individually on a computer screen and asked to identify the object in the picture.

Mean accuracy for correctly labeling the pictures selected for the induction task approached ceiling (M = 0.91, SD = 0.07). Children's high accuracy on the calibration suggests that the stimuli chosen for this study were highly familiar to preschool-age children, to the point that children could spontaneously label the objects correctly.

All triads utilized in the Property Induction Task consisted of a target item, category match, and a perceptual match (e.g., *bird-bird-bat*). The triads were designed such that category membership was in conflict with perceptual similarity. To ensure the triads used in the Property Induction task contained strong conflict a Similarity Calibration (described below) was also conducted.

Similarity Calibration. A separate sample of 4-year-old children (N = 20, M_{age} = 4.43 years, SD = 0.26, 10 Males, 10 Females) participated in the calibration study to ensure the triads used in the Property Induction task contained strong conflict between perceptual similarity and object kind. The calibration study used the same visual stimuli as the Property Induction task (see Figure 1). Children were randomly assigned to one of two experimental conditions (*Kind* vs. *Perceptual similarity*). In both conditions, a simple matching task was administered in which children were asked to match the target object to one of the test items based on object kind or perceptual similarity (according to the child's condition assignment).

In both conditions, the experimenter introduced the game to the children by explaining that the child's task was to identify which objects "*go together*." Then the experimenter delineated the matching rule based on the child's condition assignment. For example, in the Kind condition children were told that the rule of the game was as follows: *"objects that are the **same kind** of thing go together."* The children were then provided with an example: "*…this is a lemon and this is a lemon slice; they go together because they are the same kind of thing"*. In the Perceptual similarity condition an analogous procedure was followed; however, the matching rule and example were modified accordingly: "*The rule of the game is that objects that **look similar** go together. For example, this lemon is yellow and round and this tennis ball is yellow and round, so they go together, because they look similar."* Akin to the Property Induction Task, no labels were utilized in the matching tasks.

Children's accuracy on the matching tasks was first compared to chance (0.50). In both conditions performance did not differ significantly from chance: M_{Kind} = 0.59 (SD = 0.25), $M_{Perceptual\ Similarity}$ = 0.65 (SD = 0.24); both $ps > 0.075$. Next, accuracy rates were compared across the two conditions and were found to be statistically equivalent, independent-samples $t(18) = 0.51$, $p = 0.61$. The calibration results confirm that the triads selected for the Property Induction task successfully placed perceptual similarity in conflict with object kind.

Property Induction Task. In the Property Induction Task children were presented with 14 triads. Each triad included a target, category match, and perceptual match (see Figure 1). Category membership was communicated solely through detailed color photographs and no labels were used (cf. Smith & Heise, 1992).

[1] Data collection with the 2-year-olds is currently in progress.
[2] 78% of the stimuli were included in this calibration, 19% of the labels were calibrated in prior work (see Fisher, 2011).

Figure1. Visual stimuli used in the Property Induction Task. No labels were presented during the task. Each triad includes a target (center), category match (left) and a perceptual match (right). *Note. *'Monkey'* is a common label that children apply to this item and was therefore counted as a 'correct' response.

On every trial children were told that the target object had a particular property. All properties were one-syllable blank predicates chosen from the NOUN database (e.g., *fisp*, *wilp*, etc.; Horst, 2009). Then, the children were asked to generalize the target property to one of the test items (i.e., the category match or the perceptual match). For the animal triads children were told that the target item possessed an internal pseudo-biological property (e.g., *"This one has fisp cells inside"*) and they were asked to generalize the property to one of the test items. For the artifact triads children were told what the target object was made of (e.g., *"This one is made of fupp"*) and asked to generalize the property to one of the test items. The screen location of the test items was counterbalanced and the trials were presented in one of two orders: In Order 1 the trials were randomized. For Order 2 the presentation order was simply reversed. Presentation order was counterbalanced across participants.

Naming Task. After the Property Induction task, all participants completed a Naming Task to ensure that participants were familiar with all of the stimuli. The naming task was identical to the procedure utilized in the Familiarity Calibration: Participants were presented with a series of 42 pictures displayed individually on a computer screen. Children were asked to identify the object in the picture. Two presentation orders were created. In Order 1 the items were pseudo randomized with the following constraints: for any given triad, the target, category match, and perceptual match could not occur on successive trials. For Order 2 the items were administered in reverse order. The presentation order was counterbalanced across participants.

Results

Mean induction scores by age group and trial type are displayed in Table 1. Children's induction scores were submitted to a mixed ANOVA with *age* (5-, 4-, 3-, and 2-year-olds) as the between-subject factor and *trial type* (Animals, Artifacts) as the within-subject factor. The effect of *trial type* was not significant, $F(1, 59) = 2.31$, $p = 0.13$. The interaction between *trial type* and *age* was also not significant, $F(3, 59) = 0.44$, $p = 0.72$.

Table 1. Mean induction scores by age group and trial type

Age Group	Animals	Artifacts
5-year-olds	0.81 (0.22)	0.80 (0.22)
4-year-olds	0.60 (0.26)	0.58 (0.26)
3-year-olds	0.49 (0.23)	0.41 (0.21)
2-year-olds	0.52 (0.14)	0.41 (0.25)

A significant effect of age was found, $F(3, 59) = 11.15$, $p < 0.0001$. There was no significant difference in the induction performance between 2-, 3-, and 4-year-old children; Post-hoc Tukey: all *ps* > 0.141. Induction performance of the 5-year-old children was significantly higher than all other age groups; Post-hoc Tukey: all *ps* < 0.005. Additionally, only the 5-year-old children selected category match items at above chance (0.50) level, one-sample $t(17) = 7.02$, $p < 0.0001$. The rate of choices of category match items in 4-year-old children approached significance, one-sample $t(20) = 1.81$, $p = 0.09$; whereas performance of 2- and 3-year-old children did not differ from chance, one-sample *ts* < 1.10, *ps* > 0.29 (See Figure 2).

Preschool children's difficulty on the Property Induction task was clearly not due to lack of familiarity with the stimuli as evidenced by children's ability to label the stimuli with high accuracy. For 3-, 4-, and 5-year-old children, performance on the Naming Task[3] approached ceiling levels ($M_{3\text{-year-olds}} = 0.87$, $SD = 0.10$; $M_{4\text{-year-olds}} = 0.85$, $SD = 0.26$; $M_{5\text{-year-olds}} = 0.95$, $SD = 0.06$). Although the youngest participants (2-year-olds) accuracy on the naming task was

[3] Two children did not provide a verbal response and thus did not contribute any Naming task data. Across Experiments 1 and 2, 0.3% of the Naming task trials were excluded due to experimenter error.

lower than the accuracy rates of the older children, they were able to accurately label the majority of the stimuli ($M_{2\text{-}year\text{-}olds}$ = 0.71 *SD* = 0.15). Critically, there was no significant difference in the naming accuracy between the 5-year-olds and 4-year-olds, independent-samples $t(37) = 1.51$, $p = 0.14$, despite dramatic differences in their induction performance (the effect size on the difference in induction performance was large, Cohen's $d = 1.08$). Therefore, even though children were highly familiar with the categories of objects used in the current study, only kindergarten-age children were able to resolve the conflict between different sources of information in favor of object kind, whereas preschoolers were not able to do so and performed at chance.

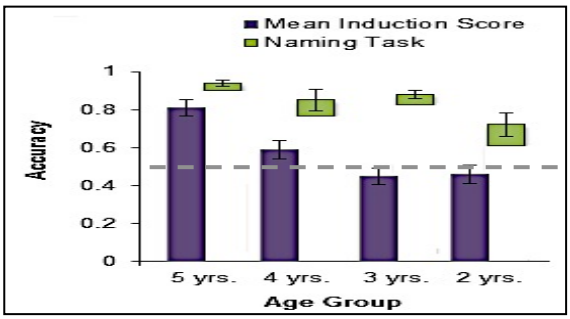

Figure 2. Summary of children's performance on the Property Induction Task and Naming Task across age groups. Error bars represent the standard error of the means. Line indicates chance performance (0.50)

Importantly, the results of Experiment 1 are not fully consistent with either the naïve theory approach or the similarity-based approach. Specifically, the naïve theory approach predicts little to no developmental trend and above chance category-based induction in all age groups; neither of these predictions was supported by the findings. The similarity-based approach predicts a developmental increase in category-based induction, which was observed in the reported data. However, the similarity-based approach also predicts that perceptual similarity should have a larger influence on performance than object kind information, and this prediction was not supported for any of the age groups tested in this study.

Experiment 2

To ensure that children's poor performance on the Property Induction Task stemmed from the presence of conflict and not from children's inability to identify the object kind of the chosen stimuli, we re-paired the items to remove the conflict and presented children with a simple matching task in which children were asked to match the target object to one of the test items based on object kind or perceptual similarity.

Method

Participants

In this study participants included: 5-year-olds ($N = 34$, $M_{age} = 5.50$, $SD = 0.38$, 16 Males, 18 Females); 4-year-olds ($N = 32$, $M_{age} = 4.56$, $SD = 0.31$, 15 Males, 17 Females); 3-year-olds ($N = 36$, $M_{age} = 3.63$, $SD = 0.27$, 17 Males, 19 Females); and 2-year-olds ($N = 16$, $M_{age} = 2.64$, $SD = 0.34$, 9 Males, 7 Females). None of the children from Experiment 1 participated in Experiment 2.

Design and Procedure

Matching Tasks: Kind and Perceptual Similarity. Children were randomly assigned to either the No Conflict Kind Matching condition or the No Conflict Perceptual Similarity Matching condition. Visual stimuli were identical to those utilized in the Property Induction task in Experiment 1. However, for every triad the lures were repaired to remove conflict between category membership and perceptual similarity. Lures were repaired based on the following constraint: lures from the animal triads were repaired with other animal triads and lures from artifact triads were repaired with other artifact triads. In other words, the repairing of lures did not result in crossing ontological boundaries. A full list of the repaired stimuli utilized in the experiment is provided in Table 2.

The task instructions were identical to the instructions utilized in the Similarity Calibration described above. The children were asked to match the target object to one of the test items based either on object kind or perceptual similarity (according to the child's condition assignment). The location of the correct response and lure were counterbalanced across trials. Two presentation orders were used; the presentation orders were identical to those used in the Property Induction Task in Experiment 1. Presentation order was counterbalanced across participants.

Naming Task. After the experiment proper, all participants completed the same Naming Task used in Experiment 1. The Naming Task served to ensure that the children in Experiment 2 were familiar with all of the stimuli. Participants were presented with a series of 42 pictures displayed individually on a computer screen. Participants were asked to identify the object in the picture.

Results

Performance on the Matching Tasks. First, we compared children's performance to chance (0.50). In the No conflict Kind Matching condition all age groups performed significantly above chance, all one-sample $ts > 4.41$, $ps < 0.003$. Similarly, 3-, 4-, and 5-year-old children performed above chance in the No Conflict Similarity Matching condition, all one-sample $ts > 5.00$, $ps < 0.0001$; however, 2-year-olds performance in this condition was not significantly different from chance, one-sample $t(7) < 1$, *ns.* Therefore, with the conflict removed, young children were largely successful in matching the same pairs of objects according to object kind and perceptual similarity (see Figure 3).

Children's scores on the matching task were submitted to a two-way ANOVA with *age* (5-, 4-, 3-, and 2-year-olds) and *condition* (No Conflict Kind Matching vs. No Conflict Perceptual Similarity) as between-subject factors. The effect

of *condition* was not significant indicating that children were equally accurate at matching the stimuli according to kind relations as they were by perceptual similarity; $F(1, 110) = 2.12$, $p = 0.15$. A main effect of age was found ($F(3, 110) = 14.79$, $p < 0.0001$). Five-year-olds' scores were significantly higher than the 2- and 3-year-old children (post hoc Tukey: $ps < 0.002$), while there was no significant difference in the performance of the 4- and 5-year-olds (post hoc Tukey: $p = 0.92$). However, the interaction between *condition* and *age* was not significant, $F(3, 110) = 0.14$, $p = 0.93$.

Table 2. Stimuli for the matching tasks in which conflict between category-membership and perceptual similarity was removed.

No Conflict Kind Stimuli		
Target	Category Match	Lure
Bunny	Bunny	Gorilla
Bird	Bird	Dog
Cat	Cat	Bat
Pig	Pig	Cat
Bear	Bear	Cow
Monkey	Monkey	Raccoon
Dog	Dog	Squirrel
Lights	Light	Drum
Book	Book	Candy cane
Umbrella	Umbrella	Plate
Cake	Cake	Microphone
Balloon	Balloon	Necklace
Clock	Clock	Lollipop
Flashlight	Flashlight	Present
No Conflict Perceptual Similarity Stimuli		
Target	Perceptual Match	Lure
Bunny	Squirrel	Pig
Bird	Bat	Dog
Cat	Raccoon	Bird
Pig	Dog	Monkey
Bear	Gorilla	Cat
Monkey	Cat	Bear
Dog	Cow	Bunny
Lights	Necklace	Cake
Book	Present	Balloon
Umbrella	Candy cane	Flashlight
Cake	Drum	Light
Balloon	Lollipop	Clock
Clock	Plate	Umbrella
Flashlight	Microphone	Book

These findings suggest that results of Experiment 1 cannot be attributed to children's inability to identify object kind and perceptual similarity matches per se. Instead, preschoolers' difficulty in making systematic (i.e., different from chance) inferences in Experiment 1 can be attributed to their inability to resolve conflict between perceptual similarity and object kind.

In line with the results obtained in Experiment 1, children in Experiment 2 also exhibited highly accurate performance on the Naming Task ($M_{5\text{-year-olds}} = 0.96$, $SD = 0.04$; $M_{4\text{-year-olds}} = 0.94$, $SD = 0.06$; $M_{3\text{-year-olds}} = 0.86$, $SD = 0.10$; $M_{2\text{-year-olds}} = 0.71$, $SD = 0.13$). This result suggests that participants were familiar with the stimuli utilized in the matching tasks.

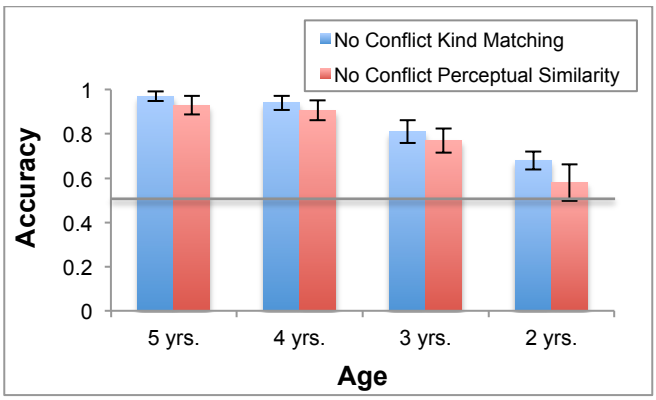

Figure 3. Summary of children's performance on the no conflict matching tasks (Kind and Perceptual Similarity matching) across age groups. Error bars represent the standard error of the means. Line indicates chance performance (0.50)

Discussion

Taken together, the results from this study suggest that neither of the current theoretical accounts of inductive generalization in young children can fully explain the reported findings. As stated above, the naïve theory account is unable to account for: (1) preschoolers' failure to make consistent category-based inferences in the presence of perceptual conflict, and (2) developmental increases in category-based responding. The similarity-based approach is unable to explain the lack of perceptual similarity-based responses, even in the youngest participants tested in this study. Therefore, the reported findings call for a revision of the current theoretical perspectives. Below, we briefly outline a revised version of the similarity-based account, which can capture the findings reported in this paper.

The basic premise of our revised similarity account is that one can distinguish two forms of featural similarity: perceptual and representational similarity. Perceptual similarity refers to features that can be compared on-line and in-the-moment. An attentional weighting parameter can be used to specify why some features should make a larger contribution to the overall similarity (i.e., based on differential saliency; Sloutsky & Fisher, 2004; 2012). Nonetheless, perceptual similarity refers to the features of a directly observed entity.

Representational similarity, as the name implies, refers to the featural overlap in mental representations. All of the features that have been encoded and stored in memory contribute to representational similarity. Representational similarity includes properties that have been perceived directly as well as properties that have come from conversations, books, or other indirect sources.

Representational similarity in our view is synonymous with *semantic knowledge* (or semantic memory). At the same time, we see this type of knowledge as distinct from what is often referred in the literature as *conceptual knowledge*. Although this term can be used in the lean sense as simply knowledge about concepts (e.g., Sloutsky, 2010), proponents of the naïve theory approach often use this term

to refer to knowledge that children may have "independent of experience" (Gelman & Markman, 1986, p. 207). In contrast, semantic knowledge is rooted in experience. Therefore, conceptual knowledge is broader than semantic knowledge: for instance, conceptual knowledge of what a 'dog' is may include semantic features (Clark, 1973), such as 'four legs', 'furry', and 'barks', but it may also include an essentialist belief that "there is some unobservable property … the essence" that makes something a 'dog' (Gelman, 2003, p. 7). The latter belief would constitute conceptual but not semantic knowledge.

We propose that inductive generalization early in development is similarity-based, with both perceptual and representational similarity contributing to the overall similarity of presented entities. Developmental changes in performance on induction tasks with familiar categories are hypothesized to stem from: (1) developmental changes in representational similarity (Godwin, Matlen, & Fisher, 2013) or from (2) changes in the relative contribution of perceptual and representational similarity. These issues remain to be addressed in future research.

Acknowledgements

We thank Malika Sinha, Manon Sohn, Amy Barrett, Laura Pacilio, Anna Loiterstein, Rachel Walsh, and Jennifer Shin for their help collecting data. We thank the children, parents, teachers, and Phipps Conservatory 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).

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.

Baldwin, D. A., Markman, E. M., & Melartin, R. L. (1993). Infant's ability to draw inferences about nonobvious object properties: Evidence from exploratory play. *Child Development,* 64(3), 711-728.

Booth, A. E., & Waxman, S. R. (2002). Word learning is 'smart': Evidence that conceptual information affects preschoolers' extension of novel words. *Cognition*, 84, B11–B22.

Clark, E. V. (1973). What's in a word? On the child's acquisition of semantics in his first language. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 65-110). New York: Academic Press.

Fisher, A. V. (2007). Are developmental theories of learning paying attention to attention? *Cognition, Brain, and Behavior, 11, 635-646.*

Fisher, A. (2011). Processing of perceptual information is more robust than processing of conceptual information in preschool-age children: Evidence from costs of switching. *Cognition*, *119, 253-264.*

Fisher, A., Matlen, B., & Godwin, K. (2011). Semantic similarity of labels and inductive generalization: Taking a second look. *Cognition*, 118(3), 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, 583-597.

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

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.

Godwin, K., Matlen, B., & Fisher, A. (2013). 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,* 115(1), 74-90.

Graham, S. A., Kilbreath, C. S., & Welder, A. N. (2004). Thirteen-Month-Olds Rely on Shared Labels and Shape Similarity for Inductive Inferences. *Child Development*, 75(2), 409-427.

Horst, J. S. (2009). Novel Object and Unusual Name (NOUN) Database. Retrieved from: http://www.sussex.ac.uk/wordlab/noun

Jaswal, V. K. & Markman, E. M. (2007). Looks aren't everything: 24-month-olds' willingness to accept unexpected labels. *Journal of Cognition and Development*, 8(1) 93-111.

Jones, S. S. & Smith, L. B. (1998). How children name objects with shoes. *Cognitive Development,* 13(3)*,* 323-334.

Kalish, C. W. (2006). Integrating normative and psychological knowledge: What should we be thinking about? *Journal of Cognition and Culture,* 6, 161-178.

Keil, F. (1989). *Concepts, kinds, and cognitive development.* Cambridge, MA: MIT press.

Markman, E., (1990). Constraints children place on word meanings. *Cognitive Science,* 14, 57-77.

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 (2), 166 – 188*.

Sloutsky, V. M., & Fisher, A. V. (2012) Linguistic labels: Conceptual markers or object features? *Journal of Experimental Child Psychology, 111* 65-86.

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.

Smith, L. B. & Heise, D. (1992). Perceptual similarity and conceptual structure. Percepts, Concepts and Categories. B. Burns (Ed). Elsevier Science Publishers B. V., 233-272.

Smith, L.B., Jones, S. S., & Landau, L. (1996), Naming in young children: A dumb attentional mechanism? *Cognition, 60,* 143 – 171.

Smith, L. B. & Samuelson, L. (2006) An attentional learning account of the shape bias: Reply to Cimpian & Markman (2005) and Booth, Waxman & Huang (2005). *Developmental Psychology,* 42(6), 1339-1343.