Title
When Diverse Evidence is (and isn't) Inductively Privileged: The Influence of Evidence Presentation on Children's and Adults' Generalization

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 36(36)

ISSN
1069-7977

Author
Lawson, Chris

Publication Date
2014

Peer reviewed

When Diverse Evidence is (and isn't) Inductively Privileged: The Influence of Evidence Presentation on Children's and Adults' Generalization

Chris A. Lawson (lawson2@uwm.edu)
Department of Educational Psychology, UW-Milwaukee, 2400 E Hartford Ave
Milwaukee, WI 53211

Abstract

The ability to determine that diverse samples provide better evidence for generalization than non-diverse samples is an important inductive skill. Adults tend to utilize the diversity principle of induction (DP), but evidence regarding children's ability to do so is mixed. The two experiments reported here examined whether the method by which evidence is presented would have an influence on children's tendency to obey the DP. These experiments with undergraduates (N = 66, Mage = 21.12 years) and preschoolers (N = 62, Mage = 5.27 years) revealed that whether sample items were presented sequentially or simultaneously influenced diversity-based reasoning in children, and in some cases, adults. Specifically, sequential presentation facilitated diversity-based reasoning, and simultaneous presentation did not. Together these results indicate that processes elicited during the presentation of evidence have an important influence on how children and adults use evidence to make inductive generalizations.

Introduction

People utilize a small set of principles to make inductive decisions (Heit, 2000; Osherson, Smith, Wilkie, López, A., & Shafir, 1990). One such principle, the *diversity principle* (DP), dictates that evidence that covers a greater scope of phenomena provides better support for a conclusion than evidence that covers a narrow scope of phenomena. Consider the following example in which statements above the line are premises in an inductive argument and the statement below the line represents the conclusion:

(1) Hippos secrete uric acid crystals
Hampsters secrete uric acid crystals
____
All Mammals secrete uric acid crystals

(2) Hippos secrete uric acid crystals
Rhinos secrete uric acid crystals
____
All Mammals secrete uric acid crystals

Undergraduates find the argument presented in (1) to be stronger than the argument in (2) presumably because the premises in the former provide better coverage of the conclusion category (e.g., mammal), and thus better support for generalization, than the premises in the latter (e.g., Osherson et al., 1990).

Findings regarding the use of the DP in children are mixed. For example, studies that employed a category-based induction format (in which items took the form of the example outlined above) found that the DP does not emerge until sometime after 7 years of age (Gutheil & Gelman, 1997; Li, Cao, Li, & Deak, 2009; Lopez, Gelman, Gutheil, & Smith, 1992). However, research using other methodologies found that children younger than 5 years of age respect the diversity principle (Heit & Hahn, 2001; Lo, Sides, Rozelle, & Osherson, 2002; Shipley & Shepperson, 2006). For example, after learning about a child who chose to play with a diverse set of balls (e.g., basketball, cricket ball, and tennis ball) and another who chose to play with a non-diverse set of balls (e.g., 3 footballs), 5-year-olds predicted that the child who played with the diverse set of balls would pick a novel ball to play with (e.g., baseball; Heit & Hahn, 2001).

One explanation for these mixed findings is that they reflect differences in domain knowledge. For example children might respect the DP in the social domain because of their keen awareness of preferences as a guide for making social inferences (e.g., Woodward & Somerville, 2000). In contrast, failure to use the DP in category-based induction tasks, which typically include animal categories and novel biological properties, might be due to limitations in children's biological knowledge (e.g., Carey, 1985; cf. Gelman, 2003).

An alternative interpretation is that these mixed findings reflect children's difficulty incorporating certain features of evidence into their inductive decisions. Success on the category-based diversity task requires, at minimum, the ability to analyze each sample, determine the category shared by the evidence and conclusion animals, and evaluate the extent to which one sample provides better "coverage" of the category. Failure to use the DP in these tasks may be due to limitations in younger children's capacity to process the available input or to engage in, or integrate, all of these different processes (Lopez et al., 1992; see also Gutheil & Gelman, 1997). From this perspective, many have interpreted diversity-based reasoning as a mature inductive skill that undergoes significant developmental change between the ages of 5 to 8 years (Gutheil & Gelman, 1997; Lopez et al., 1992; Rhodes, Gelman, & Brickman, 2008).

A slightly different interpretation, and the focus of the current studies, is that young children are able to engage in diversity-based reasoning but that their ability to do so depends on specific task features (e.g., Heit & Hahn, 2001). Thus, rather than positing that the DP develops later in life, the idea is that the capacity to incorporate and effectively use diversity is a feature of early induction, but that it is masked under certain task demands. One reason to favor this interpretation is the substantial database of evidence indicating that well before their preschool years children have spent a large part of their cognitive lives detecting, and learning from, variability within the available evidence

(Quinn, Eimas, & Rosenkratz, 1993; Saffran, Aslin, & Newport, 1996; Younger & Cohen, 1986). Second, the claim that children's adherence to the DP is influenced by specific task features, is consistent with a host of other studies showing that children are flexible inductivists insofar as they are able to modify their generalizations to accommodate certain aspects of the evidence (Hayes & Thompson, 2007; Kalish & Gelman, 1992; Opfer & Bulloch, 2007).

The two studies reported here tested the prediction that the method of evidence presentation would influence how children use diversity information to make inductive generalizations. These studies draw from extensive work with adults in experimental and cognitive psychology on the distinct outcomes of simultaneous and sequential presentation (e.g., Krueger, 1983; Liu & Becker, 2013; Rescorla, 1980). In general, sequential presentation elicits identification of the differences between presented items (Lappin & Bell, 1972; Quinn & Bhatt, 2010), while simultaneous presentation supports identification of similarities between presented items (Gentner & Namy, 2006). Thus, sequential presentation can potentially facilitate two crucial components of diversity-based reasoning – identification of variability within samples and assessment of differences between the presented samples.

Several recent studies suggest these presentation formats lead to different outcomes in children's generalizations. For example, Lawson (2014) found that 3-year-olds obeyed the sample size principle of induction (i.e., they preferred to generalize from large, rather than small samples of evidence) when sample items were presented sequentially but not when they were presented simultaneously. Work by Spencer and colleagues (2011) revealed that presentation format influences the scope of children's label generalizations: When three instances from the same subordinate (e.g., 3 green peppers) were presented and labeled sequentially 3-year-olds generalized the label broadly (e.g., to other peppers), but when the same items were presented and labeled simultaneously children generalized the label narrowly (e.g., to green peppers). Similarly, in a property projection task, Lawson and Fisher (2011) found that when evidence about a property shared by 16 mammals was presented sequentially, 5-year-olds, but not adults, generalized the property broadly to a range of animals (e.g., vertebrate and invertebrate). At least for young children sequential presentation appears to facilitate broad generalization.

Drawing from these prior findings, the goal of the present studies was to test two predictions about how presentation format might influence diversity-based reasoning in children and adults. The first prediction was that sequential presentation, rather than the simultaneous presentation, would facilitate diversity-based reasoning. This pattern may be particularly true for young children due to the information processing demands of the task. In particular, because children may have difficulty coordinating all of the input in a way that would support diversity-based induction they may be more likely to benefit when the task structure facilitates processes necessary for diversity-based reasoning. Because adults spontaneously engage in diversity-based reasoning, the presentation format will likely have little or no effect on their inferences.

A second prediction is that presentation format might influence the scope of the diversity effect. Findings from Spencer et al. (2011) and Lawson and Fisher (2011) indicate that sequential presentation yields a broader scope of generalization. Thus one prediction is that, to the extent that sequential presentation facilitates the DP, it will support a broad pattern of generalizations. However, one normative prescription of diversity is that it serves an eliminative function (e.g., Heit, Hahn, & Feeney, 2005): Diverse evidence ought to eliminate alternatives and narrow hypotheses. Consider an example the following example. Suppose you learn that a diverse sample of items share a biological property (e.g., cat, wolf, whale all have plaxium blood) and a non-diverse sample share a different biological property (e.g., cat, tiger, zebra all have drotium blood). Does the diverse sample provide better evidence to support inferences to (only) mammals, or does it provide better evidence to generalize beyond the category of mammals? Experiment 2 was designed, in part, to explore whether one of the presentation formats was more likely to yield one of these generalization functions.

Experiment 1

Experiment 1 examined the potential influence of presentation format on diversity-based reasoning in young children. The main prediction was that sequential presentation of evidence, which elicits attention to variability within samples and encourages identification of differences between presented items, would support diversity-based reasoning in preschoolers while simultaneous presentation, because it supports identification of similarities between items, would not facilitate diversity-based reasoning. With the exception of the manipulation of presentation format the task was modeled after a version of category-based induction task that has shown later development of the DP (e.g., Gutheil & Gelman, 1997; Lopez et al., 1992). A sample of undergraduates was included to explore whether presentation format might affect the extent to which they engage in diversity-based reasoning and to serve as a developmental endpoint for comparison with the younger participants.

Method

Participants. Thirty-three undergraduates (*M*age = 21.45 years, *SD* = .98 years; 20 females, 13 males) and 32 five-year-olds (*M*age = 5.36 years, *SD* = .27 years; 18 females, 14 males) participated in Experiment 1. In both of the experiments reported here undergraduates were recruited from Psychology courses and received partial credit and children were recruited from local preschools, which were given small monetary donations. Participants were representative of the racial and socio-economic diversity of a medium-sized Midwestern US city.

Design and Procedure. Participants responded to 12 items each of which included a diverse sample, a non-diverse sample, and a target. All items were from the same basic-level (e.g., fish). For each item the diverse sample always included a range of animals from the same basic-level (e.g., trout, goldfish, surgeonfish) and the non-diverse sample always included a selection of animals from the same subordinate (e.g., 3 different goldfish). The target was a novel instance from the same basic-level as the individuals in both evidence samples (e.g., shark). The 12 items included animals familiar to young children. All items were presented in random order.

Method of evidence presentation (Simultaneous, Sequential) was manipulated between subjects and participants were randomly assigned to one of these two conditions. In the *Sequential* condition each evidence item was presented individually and attributed a novel biological property (e.g., "This animal has drotium blood"). The individuals from one sample were presented first and placed into a single pile to the left of the participant. Then, the individuals from the other sample were presented and attributed a different novel biological property (e.g., "This animal has plaxium blood"), and grouped into a different pile placed to the right of the participant. In the *Simultaneous* condition the items from both samples were presented at the same time and placed into two separate piles, one to the left of the participant and the other to the right. The experimenter gestured to one group and described, "These animals have drotium blood", and then to the other group and described, "These animals have plaxium blood". The location of the samples (left or right of the participant) was counterbalanced.

After presentation of the two samples participants were shown the target item to which they were asked to generalize the property attributed to the diverse sample or the property attributed to the non-diverse sample (e.g., "Do you think this (target) animal has drotium blood, like these animals, or plaxium blood, like these animals?"). The target was placed in front of the participant, and equidistant from the two samples until participants responded.

The task was identical for children and undergraduates. Undergraduates participated in a quiet laboratory room located on their campus. Children participated in a quiet location at their preschool. The entire experiment lasted approximately 10 minutes.

Results and Discussion

The analyses focused on the proportion of generalizations from the diverse sample. These responses were submitted to an ANOVA with Age (Adults, Children) and Condition (Simultaneous, Sequential) as between-subjects variables. Both effects were significant (Age, $F(1, 62) = 27.98$, $p < .001$, $\eta^2 = .31$, and Condition, $F(1, 62) = 7.19$, $p = .009$, $\eta^2 = .11$) as was the Age by Condition interaction, $F(1, 62) = 5.28$, $p = .03$, $\eta^2 = .09$. Overall adults were more likely to generalize from the diverse sample than children. However, as can be seen in Figure 1, there were no age differences in the Sequential condition, $F < 1$, *ns*, but there were differences in the Simultaneous condition, $F(1, 30) = 26.52$, $p < .001$, $\eta^2 = .47$, due to a higher rate generalizations from the diverse sample among adults than children. Supplemental analysis indicated the condition effect was significant for children $F(1, 31) = 14.67$, $p < .001$, $\eta^2 = .33$, but not adults, $F < 1$, *ns*. As expected, children exhibited a higher rate of diversity-based responses in the Sequential conditions than the Sequential conditions.

Follow-up analyses compared responses to chance ($M = .50$) to identify the cases for which participants exhibited a consistent preference to generalize from the diverse sample. In the Sequential condition both groups showed a consistent preference to generalize from the diverse sample, both *t*s > 4.45, *p*s < .001, *d*s > 2.28. In the Simultaneous condition adults exhibited a consistent preference to generalize from the diverse sample, $t(16) = 5.99$, $p < .001$, $d = 3.00$, while children's responses were no different from chance, $t < 1.3$, *ns*.

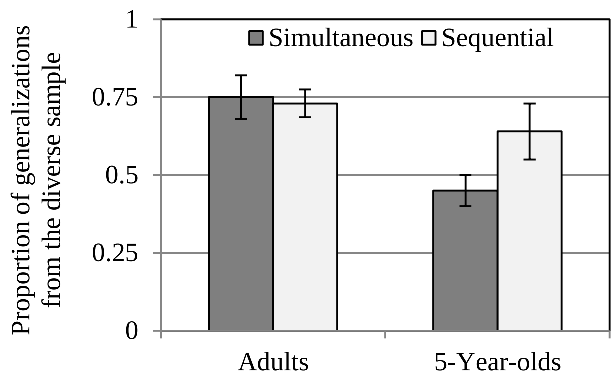

Figure 1. Proportion of generalizations from the diverse sample for Adults and 5-year-olds in the Sequential and Simultaneous conditions in Experiment 1. Bars represent one standard error from the mean.

These results are consistent with the prediction that sequential presentation of evidence would facilitate the DP. As predicted, this effect was pronounced for children: Adults consistently generalized from the diverse sample regardless of presentation format, while children only showed the diversity effect for sequential presentation. Overall these results are inconsistent with the idea that diversity-based reasoning is a mature inductive skill (e.g., Rhodes et al., 2008) and instead suggest that diversity is available early in development but only when evidence is presented in a way that engages processes that facilitate this type of reasoning.

Experiment 2

The results from Experiment 1 revealed that sequential presentation, but not simultaneous presentation, of evidence lead to diversity-based reasoning in children as young as 5 years of age. Adults likely did not show the effect because they spontaneously process input in a way that supports diversity-based reasoning, regardless of how that evidence is made available to them. It is possible that even though adults routinely utilize the DP, evidence presentation might influence the scope of generalization. For example, because sequential presentation supports detection of the variability within a sample primes a search of differences rather than similarities between items it could lead to a heightened awareness of the exemplars not "covered" by the sample. Thus, sequential presentation may support the narrowing function of the DP. Experiment 2 tested this prediction by asking participants to generalize properties to targets from a different basic-level (e.g., crustacean) than was represented by the individuals in the diverse and non-diverse samples (e.g., fish). Would participants prefer to generalize from a diverse sample even when doing so involved endorsing a broad generalization?

Method

Participants. 34 undergraduates (*M*age = 20.87 years, *SD* = .98 years; 20 females, 12 males) and 27 five-year-olds (*M*age = 5.31 years, *SD* = .54 years; 12 females, 15 males) participated in Experiment 2.

Design and Procedure. The experiment was identical to Experiment 1 with the exception that in Experiment 2 all of the targets were drawn for a different basic-level category (but still within the same superordinate) as the exemplars in the diverse and non-diverse samples. For example, for the fish item the diverse sample included a surgeonfish, trout, and goldfish, the non-diverse set included 3 goldfish, and the target was a crab. In all other respects the experiment was identical to Experiment 1.

Results and Discussion

The analytic approach was the same as in Experiment 1. The Age (Adults, Children) by Condition (Sequential, Simultaneous) ANOVA yielded both main effects; Age, $F(1, 57) = 6.99$, $p = .01$, $\eta^2 = .10$, and Condition, $F(1, 57) = 7.33$, $p = .009$, $\eta^2 = .11$, however the interaction was not significant, $F < 2$, *ns*. As in Experiment 1, adults showed a higher rate of generalizations from the diverse sample than children (see Figure 2). Also, there was an overall higher rate of generalizations from the diverse sample in the Sequential condition than in the Simultaneous condition. Although the Age by Condition interaction was not significant, separate ANOVAs were conducted to assess whether the Condition effect was consistent for both age groups. These analyses revealed the condition effect was significant for adults, $F(1, 32) = 4.41$, $p = .04$, but for children the effect failed to reach significance, $F(1, 24) = 2.74$, $p = .11$. Thus, these results are consistent with the idea that sequential presentation would lead to a broad, rather than narrow, pattern of generalizations, though the effect was pronounced for adults.

Further analyses involved comparisons to chance ($M = .50$). Adults exhibited a consistent preference to generalize properties from the diverse sample in both conditions, both *t*s > 3.1, *p*s < .006, *d*s > 1.54. Children did not show a consistent pattern of generalizations from the diverse sample in either condition, both *t*s > 1.50, *ns*.

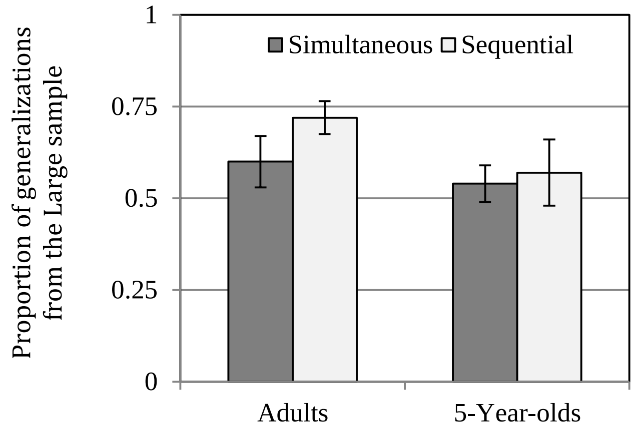

Figure 2. Proportion of generalizations from the diverse sample for Adults and 5-year-olds in the Sequential and Simultaneous conditions in Experiment 2. Bars represent one standard error from the mean.

These results confirm and extend the findings from Experiment 1. Sequential presentation facilitated diversity-based reasoning, and did so for targets from a different basic-level as was represented in the evidence samples. Finally, this effect was more consistent among adults than children.

General Discussion

The two studies described here examined the extent to which the method of evidence presentation would facilitate diversity-based reasoning in children. Results from Experiment 1 indicate that children as young as 5 years respect the diversity principle under specific conditions: Children preferred to generalize from diverse, rather than non-diverse samples, when evidence items were presented sequentially but not when the same items were presented simultaneously. Additionally, the results from Experiment 2 provided some evidence that sequential presentation elicited a broad pattern of generalizations: At least for adults, when the conclusion represented a category from a different basic-level as was represented in both evidence samples, they preferred to generalize from a diverse sample. Adults exhibited a clearer preference to generalize from diverse samples: They showed a broader pattern of generalizations in the sequential condition than the simultaneous condition, but unlike children they consistently showed the diversity effect regardless of evidence presentation. Moreover, in both experiments adults demonstrated an overall higher rate of generalizations from the diverse sample than did children.

Overall these results were consistent with the prediction that sequential presentation would facilitate diversity-based reasoning in young children. One explanation for why the sequential presentation facilitated diversity-based induction is that it simplified the task for children. In the context of category-based induction, adherence to the diversity principle poses considerable information processing demands. At the very least, a reasoner must assess the composition of each sample, judge how well each sample justifies the conclusion, and then compare the samples to determine which provides the best support for generalization. Sequential presentation supports detection of differences between items (Lappin & Bell, 1972; Quinn & Bhatt, 2010), and thus supports processes necessary for diversity-based reasoning. These results suggest that children, who otherwise are unable to engage and coordinate the processes necessary for diversity-based reasoning, benefitted from a presentation format that forces them to do.

That certain conditions can encourage diversity-based reasoning in preschoolers presents a challenge to the idea that this type of reasoning is the product of mature inductive ability (Gutheil & Gelman, 1997; Li et al., 2009; Lopez et al., 1992; Rhodes et al., 2008). Instead, the present findings are in-line with other work showing that children respect diversity when making inductive reasoning (Heit & Hahn, 2001; Lo et al., 2002; Shipley & Shepperson, 2006). Moreover, the current results indicate that mixed results in prior studies are not due to differences in knowledge about the content of the task (e.g., Carey, 1985), but instead were due to differences in task complexity. Thus, one explanation for why prior category-based induction studies undermined children's ability to engage in diversity-based reasoning is because the tasks involved simultaneous presentation of samples, which encourages identification of similarities, rather than differences, between samples (e.g., Gentner & Namy, 2006; Lawson, 2014). Indeed, children never showed the diversity effect in the Simultaneous condition in either experiment reported here.

Conclusions about the effect of diversity on the scope of generalization are less clear. A normative function of diversity is to limit, or narrow, the range of plausible conclusions (e.g., Heit et al., 2005). However, recent developmental studies indicate that for samples including diverse exemplars, sequential presentation elicits broad rather than narrow generalization (Lawson & Fisher, 2011; Spencer et al., 2011). One could argue that the narrowing effect of diversity develops later, and thus these developmental studies indicate children are not yet aware of this important function of diverse samples. The results from Experiment 2 do not support this conclusion. Instead, adults consistently preferred to generalize from the diverse sample to a broad category, and were *more* inclined to do so when items were presented sequentially than when they were presented simultaneously..

How do we reconcile the broad pattern of generalizations with the idea that diverse evidence ought to support narrow generalizations? One suggestion is that the standard category-induction task is poorly suited for addressing this question. For example, rather than showing unambiguously that participants use diverse samples to endorse broad conclusions, the results from Experiment 2 might suggest that diverse samples are always favored for induction regardless of the conclusion (though see Lawson & Fisher, 2011), or that homogenous evidence provides better support for narrow generalizations than diverse evidence.

It is also worth noting that the narrowing function of diverse evidence is often discussed in the context of optimal practices in science, such that one ought to establish diverse conditions to test a specific hypothesis. However, in most induction tasks participants are seldom asked to generate their own hypotheses (cf. Lopez, 1995), or establish their own methods for data collection. One could interpret performance on an induction task from the perspective of hypothesis-testing, though perceptual processes can also lead to success in these tasks (Jones & Smith, 1993; Sloutsky & Fisher, 2004). Thus, one goal for future research is to identify whether distinct processes influence the scope of generalizations.

Despite some uncertainty about how to interpret the findings from Experiment 2, the overall effect of presentation format was a replication of the results from Experiment 1. Across both experiments sequential presentation facilitated diversity-based reasoning, and this facilitative effect was greatest among 5-year-olds.

It is important to consider some alternative interpretations of the results. For example, in addition to changing the mode of presenting the items, there were other differences between the sequential and simultaneous conditions. Critically, in the sequential condition the evidence items were described as individuals (e.g., "this animal has plaxium blood….this animal...") and in the simultaneous condition they were described as a class (e.g., "these animals have plaxium blood"). It remains to be seen whether the observed findings were due to some, or all, of these particular features. It is possible children's reluctance to generalize in the simultaneous condition was because they viewed both samples as representing the same category. Moreover, the preference to generalize in the sequential condition may have been due to the redundant property label, which was highlighted three times, compared to the single presentation in the simultaneous condition. Both of these interpretations will be important to examine in future studies.

In sum, although an extensive body of literature reveals that by 5 years of age children are quite sophisticated inductive reasoners (e.g., Gelman, 2003), there is considerable debate about the mechanisms that guide induction and the developmental trajectory of some of these skills. The methodology outlined in these two studies provides a useful framework for addressing questions emerging from this debate. Indeed, using this approach the current work suggests that rather than representing a cognitive capacity that undergoes a substantial amount of change, the diversity principle of induction appears to be relatively stable across development, yet like other cognitive skills, is subject to task demands that influence performance.

Acknowledgments

Thanks to Susie Stanley, Chelsea Kieler, and Lily Cabrera for assistance with data collection. Special thanks to the generous support of the parents, children, and teachers from Grandma's House daycare and the UW-Milwaukee children's center.

References

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

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

Gentner, D., & Namy, L.L. (2006). Analogical processes in language learning. *Current Directions in Psychological Science*, *15*, 297-301.

Gutheil, G., & Gelman, S. A. (1997). Children's use of sample size and diversity information within basic-level categories. *Journal of Experimental Child Psychology, 64,* 159-174.

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

Heit, E. (2000). Properties of inductive reasoning. *Psychonomic Bulletin & Review, 7,* 569-592.

Heit, E., & Hahn, U. (2001). Diversity-based reasoning in children. *Cognitive Psychology, 47,* 243-273.

Heit, E., Hahn, U., & Feeney, A. (2005). Defending diversity. In W. Ahn, R. Goldstone, B. Love, A. Markman, & P. Wolff (Eds.), *Categorization inside and outside of the laboratory: Essays in honor of Douglas L. Medin*. APA.

Jones, S.S., & Smith, L.B. (1993). The place of perception in children's concepts. *Cognitive Development*, *8*, 113-140.

Kalish, C. W., & Gelman, S. A. (1992). On wooden pillows: Multiple classification and children's category-based inductions. *Child Development, 63,* 1536-1557.

Krueger, L.E. (1983). Probing Proctor's priming principle: The effect of simultaneous and sequential presentation on same-different judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*, 511-523.

Lawson, C.A. (2014). Three-year-olds obey the sample size principle of induction: The influence of evidence presentation and sample size disparity on young children's generalizations. *Journal of Experimental Child Psychology.*

Lawson, C. A., & Fisher, A. V. (2011). It's in the sample: The effects of sample size and sample diversity on the breadth of inductive generalization. *Journal of Experimental Child Psychology, 110,* 499-519.

Lappin, J.S., & Bell, H.H. (1972). Perceptual differentiation of sequential visual patterns. *Perception & Pscyhophysics, 12*, 129-134.

Li, F., Cao, B., Li, Y., Li, H., & Deák, G. (2009). The law of large numbers in children's diversity-based reasoning. *Thinking and Reasoning, 15,* 388–404.

Liu, T., & Becker, M.W. (2013). Serial consolidation of orientation information into visual short-term memory. *Psychological Science*, *24*, 1044-1050.

Lo, Y., Sides, A., Rozelle, J., & Osherson, D. (2002). Evidential diversity and premise probability in young children's inductive judgment. *Cognitive Science, 26*, 181-206.

Lopez, A. (1995). The diversity principle in the testing of arguments. *Memory & Cognition, 23*, 374-382 .

Lopez, A., Gelman, S. A., Gutheil, G., & Smith, E. E. (1992). The development of category-based induction. *Child Development, 63*, 1070-1090.

Opfer, J. E., & Bulloch, M. J. (2007). Causal relations drive young children's induction, naming, and categorization. *Cognition*, 105, 206 -217.

Osherson, D. N., Smith, E. E., Wilkie, O., López, A., & Shafir, E. (1990). Category-Based Induction. *Psychological Review, 97:2*, 185-200.

Quinn, P.C., & Bhatt, R.S. (2010). Learning perceptual organization in infancy: The effects of simultaneous versus sequential variability experience. *Perception, 39*, 795-806.

Quinn, P. C., Eimas, P. D., & Rosenkrantz, S. L. (1993). Evidence for representations of perceptually similar natural categories by 3-month-old and 4-month-old infants.*Perception, 22*, 463-475.

Rescorla, R.A. (1980). Simultaneous and successive associations in sensory preconditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 6,* 207-216.

Rhodes, M., Gelman, S.A., & Brickman, D. (2008). Developmental changes in the consideration of sample diversity in inductive reasoning. *Journal of Cognition and Development, 9*, 112-143.

Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science, 274,* 1926-1928.

Shipley, E. F., & Shepperson, B. (2006). Test sample selection by preschool children: Honoring diversity. *Memory & Cognition*, *34*, 1444-1451.

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.

Spencer, J.P., Perone, S., Smith, L.B., & Samuelson, L.K. (2011). Learning words in space and time: Probing the mechanisms behind the suspicious-coincidence effect. *Psychological Science, 22*, 1049-1057.

Woodward, A.L., & Sommerville, J.A. (2000). Twelve-month-old infants interpret action in context. *Psychological Science, 11*, 73-76.

Younger, B.A., & Cohen, L.B. (1986). Developmental change in infants' perception of correlations among attributes. *Child Development, 57*, 803-815.