Mechanisms of induction in children and adults: Evidence from the study of memory distortions

Anna V. Fisher (fisher.449@osu.edu)
Center for Cognitive Science & Department of Psychology
Ohio State University
1961 Tuttle Park Place, 210E
Columbus, OH 43210, USA

Vladimir M. Sloutsky (sloutsky.1@osu.edu)
Center for Cognitive Science
Ohio State University
1961 Tuttle Park Place, 208C
Columbus, OH 43210, USA

Abstract

This paper examines mechanisms underlying induction in young children and adults. Predictions of two alternative theories of induction were tested on 37 young children and 51 adults using the DRM paradigm. The findings indicate that young children are more likely to perform similarity-based induction, and adults are more likely to perform category-based induction. Also, the reported experiment demonstrates that under certain conditions young children exhibit better memory performance than adults.

Introduction

Our ability to perform inductive inferences, that is to form generalized conclusions from particular instances, is crucial for survival: "our knowledge that leopards can be dangerous leads us to keep a safe distance from jaguars" (Sloman, 1993, p. 321). Previous studies demonstrated that even infants are capable of performing simple inductions about objects' unobservable properties (Baldwin, Markman, & Melartin, 1993; Mandler and McDonough, 1996; Graham, Kilbreath, & Welder, 2001). However the precise mechanism that allows children to perform inductive generalizations remains unknown.

One possibility is that people, including young children, first identify presented stimuli as members of categories, and then perform inductive inferences on the basis of categorization (Gelman, & Markman, 1986; Gelman, 1988). For example, when presented with two rabbits and a dog and told that one of the rabbits has hollow bones inside its body, a child is more likely to generalize this property to another rabbit than to a dog because she (presumably) understands that both rabbits belong to the same category, and members of the same category share many properties. According to this view, when presented entities are labeled by count nouns, participants treat labels as markers of categories. Therefore, labels are especially helpful when entities are novel or ambiguous: labels point to common categories even if perceptual information does not. In short, under this view, induction is a function of categorization.

Another possibility is that children perform induction by computing similarity among presented entities. Because members of the same category often happen to be more similar to each other than they are to members of other categories, young children are more likely to induce properties to members of the same category. Under this view, labels are heavily weighted attributes contributing to the overall similarity (Sloutsky & Lo, 1999), and this overall similarity drives both induction and categorization.

In much of the previous research, the role of conceptual and perceptual information in induction was studied by directly pitting perceptual similarity and labels against each other, by employing the "triad" task (Gelman & Markman, 1986; Sloutsky, Lo, & Fisher, 2001). In this task, participants are presented with a target and two test stimuli. Then participants are familiarized with some facts about the test stimuli, and asked to make an inference to the target. The most important feature of the task is that the Target stimulus shares perceptual similarity with one Test stimulus, whereas it shares the same label with the other Test stimulus. The dependent measure in this paradigm is the proportion of label-based responses.

Overall, results indicate that children rely heavily on labeling information under a no-conflict condition (for instance, when the Target and both Test items look alike), whereas the proportion of label-based responses drops when one test item is similar to the Target, whereas the other test item has the same label (Gelman & Markman, 1987; Sloutsky, Lo, & Fisher, 2001).

These results, however, are inconclusive in that proponent of either position could claim these results as supporting evidence. Proponents of the former (i.e., "induction-via-categorization") position could argue that labels are "central" properties denoting categories, and therefore, results indicate that, under most conditions, induction is a function of categorization. Proponents of the latter (i.e., "similarity-based") position could argue that because participants exhibit flexible reliance on labels, it seems unlikely that labels are "central" properties.

In short, the "triad" task is somewhat limited in that both positions make quantitatively, but not qualitatively, different predictions regarding young children's induction performance. To further differentiate between these two theoretical positions, we deemed it necessary to create a task that would elicit conflicting predictions from the two theoretical positions. One possibility is to study memories created as a result of performing an induction task. The rationale for using such tasks is that processing of perceptual and conceptual information may manifest itself in different memory traces, in such a way that perceptual processing should lead to creation of memory traces that contain more perceptual information about the stimulus than conceptual processing, thus allowing us to differentiate between category-based and similarity-based inductive inferences.

One possible candidate is a memory task based on the DRM paradigm. The DRM paradigm was originally developed by Deese to study intrusions in free recall (1959), and recently used by Roediger and McDermott (1995) to demonstrate high levels of false recognition in the domain of word learning (hence obtaining its name: DRM for Deese, Roediger, and McDermott). The DRM task usually consists of two parts. In the first part participants are asked to listen to a list of words (i.e. "bed", "rest", "awake", etc.), with every list composed of associates of one word (i.e., "sleep") that is not presented. The first part of the experiment is immediately followed by a recall or a recognition task. The finding is that participants are very likely to recall or recognize with confidence the associates of study items that had never been presented. This phenomenon proved to hold true outside of the verbal domain, and was recently demonstrated with pictorial stimuli. Koutstaal and Schacter (1997) obtained high levels of false recognition using colored pictures representing different object categories (i.e., boats, cats, shoes, etc.).

It seems likely that if participants exhibit a high level of false alarms under regular memory conditions, they should exhibit even higher levels of false alarms if they are forced to categorize entities during the study phase. This is because categorization tasks may facilitate category-level representation instead of item-level representation of individual stimuli. Therefore this variant of the DRM paradigm may help distinguishing between the two theoretical approaches to induction. Suppose that during the study phase, the task is to perform induction. If participants perform induction on the basis of categorization, then their accuracy should decrease compared to the regularly memory or baseline condition: they should wrongly recognize new members of the target category as old items. At the same time, if participants perform induction on the basis of similarity, the introduction of induction in the study phase should not change their accuracy.

Therefore, the "induction-via-categorization" position would predict that both children (as well as adults) should exhibit a decrease in accuracy when asked to perform induction during the study phase of a DRM-type task. At the same time, the similarity-based position predicts a significantly larger drop off in accuracy in adults than in young children precisely because young children perform similarity-based induction, whereas adults perform category-based induction. Furthermore, under the induction condition young children may exhibit greater accuracy than adults.

To test this hypothesis we constructed a task that consists of a study phase and a recognition phase. The task was presented under three conditions: induction, blocked categorization, and memory, or baseline. These conditions had an identical recognition phase, but they differed in the study phase. Both the study and recognition phases used pictures of highly familiar animals, whose familiarity was established in a separate calibration study with 5-year-olds. In this calibration study, participants were presented with pictures of cats, bears, and birds, and were asked to name them. Over 86% of the pictures of bears, 94% of the pictures of birds, and 99% of the pictures of cats were correctly identified and labeled by young children. Because, these animals are so familiar and easily namable, young children should easily identify these animals as members of respective categories. Therefore, if induction in young children is category-based, and the typical role of labels is to denote common categories, then these pictures should prompt category-based induction even in the absence of labels.

In the study phase of the Induction condition, participants were presented with a picture of a Target animal (a cat), and were told that this animal has a certain biological property. Participants were then presented with pictures of animals from three different categories: cats, bears, and birds. When each picture was presented, participants were asked

whether or not the animal in question had the same property as the Target animal. They were then provided with feedback, such that only members of the same biological kind (i.e., cats), but not of other kinds (i.e., bears and birds) had this property.

In the study phase of the Blocked Categorization condition, participants were presented with a picture of a Target animal (a cat), and were told that this animal is young. They were then presented with pictures of animals from three different categories, cats, bears, and birds. When each picture was presented, participants were asked whether or not the animal in question is young. They were then provided with random feedback. In this condition, categorization based on the natural kind membership was blocked: participants were forced to look for cues distinguishing young and mature animals across categories. We expected that this manipulation should facilitate item-level processing and would block categorization.

Finally, in the study phase of the Baseline condition, participants were asked to remember the animals and they were told they their memory would be tested.

During the recognition phase, which was identical for all three conditions, participants were presented with pictures half of which they had seen in the study phase, whereas the other half consisted of new pictures. Participants were asked whether or not they had seen each picture during the study phase. If induction in the study phase is indeed based on prior categorization, the rate of false recognition of novel items from the Target category (i.e., cats) in the Induction condition should increase compared to the Blocked Categorization condition. This is because in the Induction condition participants perform category-level processing, and form category-level representations, whereas in the Blocked Categorization condition they should perform item-level processing, and form item-specific representations. If, however, induction is based on the computation of similarity, the proportion of false alarms should not increase in the Induction condition compared to the Blocked Categorization condition: in both conditions participants should be rather accurate in distinguishing old items from new items. This is because under both conditions, participants should perform item-level processing and from item-specific representations.

Because we expect children to perform similarity-based induction and adults to perform category-based induction (cf. Sloutsky, et al., 2001), we predict a marked increase in false alarm rates in the Induction condition compared to the Blocked Categorization or Baseline conditions for adults. At the same time, we predict no increase in false alarms

rate for children who should be quite accurate across these conditions. Furthermore, due to the increased levels of false alarms for adults, but not for children in the Induction condition, we expect children to exhibit greater accuracy than adults in the Induction condition. In short, we expect children to form item-specific representations and adults to form category-level representations in the Induction conditions, which, in turn, should affect their recognition accuracy.

This research has, therefore, two goals. Our first goal is to present evidence that that mechanisms underlying induction differ for young children and for adults. Our second goal is to demonstrate that this difference in induction manifests itself in different memory representations. These different representations may result in greater memory accuracy in children than in adults. Therefore, this research will help to differentiate between to alternative theories of induction in young children and contribute new knowledge to the studies of memory in children and adults.

Method

The goal of the reported experiment is to demonstrate that different mechanisms underlie inductive inference in children and adults. This general claim leads to two specific predictions with regards to the DRM procedure. First, the nature of a task (Induction vs. Blocked Categorization vs. Baseline) should alter adults' but not children's memory performance. Second, adults should exhibit greater decrease in accuracy in the Induction condition compared to the Block categorization and Memory conditions than children. And third, in the Induction condition, children would exhibit greater memory accuracy than adults.

Participants

Participants in the experiment were young children from the suburbs of Columbus (18 girls and 19 boys, $M = 5.3$ years, $SD = .25$ years), and introductory psychology students who received a partial credit for completing the task (33 women and 18 men, $M = 19.2$ years, $SD = 1.2$ years)

Materials and Design

Materials in this experiment were color photographs of animals presented against white background. A total of forty-five photographs were used in Experiment 1. The include photographs of cats, birds, bears, and squirrels. As mentioned above, to ensure that each item is familiar to young children, we also ran a calibration study, in which young

children were asked to label each of the presented items.

The task consisted of two phases (i.e., Study and Recognition), and there were three between-subject conditions (i.e., Baseline, Induction, and Blocked Categorization). In the Induction condition participants were first shown a picture of the Target animal (a cat in this case) and told that the animal has beta-cells inside its body. After that participants were presented with 30 pictures of animals from three different categories (cats, bears, and birds) and asked whether each depicted animal also had beta-cells inside its body. Feedback was provided after each trial. Participants were told that they were correct if they judged a cat, but not a bear or a bird to possess the Target biological property.

In the Blocked Categorization Condition participants were told that the Target animal is young. After that participants were presented with 30 pictures of animals from three different categories (cats, bears, and birds) and asked whether the animal in each picture was young or grown-up. Five animals from each category were randomly assigned to the "young" or "grown-up" status, and random feedback was provided to participants after each trial.

In the Baseline (or Memory) condition, participants were asked to accurately remember each picture for a future recognition test.

The recognition phase of the experiment immediately followed the study phase, and was identical for both conditions. Participants were presented with pictures, one at a time, some of which they saw in the first phase of the experiment (pictures of previously seen cats and bears; 7 pictures per category) and some of which were completely new pictures (pictures of novel cats and squirrels; 7 pictures per category). Participants were presented with a total of 28 pictures in the recognition phase of the experiment. They were asked to identify whether each presented picture was "old" (previously seen during the study phase) or "new" (not seen during the study phase). No feedback was provided in the recognition phase of the experiment.

Procedure

All stimuli were presented to participants on computer screens. Young children were interviewed by female experimenters in quiet rooms of several childcare centers. Introductory psychology students were tested individually, and had all instructions presented to them on a computer screen.

Results and Discussion

Both children and adults were very accurate in rejection of non-target distracters (i.e., squirrels,

none of which was presented in the study phase), 96% of children and 93% of adults accurately rejected these new items. Therefore, the majority of children and adults took the task seriously and paid attention to stimuli during the study phase. In addition, the majority of children and adults accurately performed induction in the Induction condition (85% of children and 89% of adults).

At the same time, participants exhibited different accuracy for the Target items and critical distracters (i.e., distracters coming from the same category as the Target items). Note that correct detection of an Old Target represents a hit, whereas an erroneous recognition of a new critical distracter as an Old Target represents a false alarm. Proportions of hits and false alarms by age and experimental conditions are presented in Table 1. Note that in the Induction condition, adults had 20% higher false alarm rate than children, whereas there were no differences in the other conditions.

Proportions presented in Table 1 were averaged across participants in each condition, and then used to compute A-prime values, which are non-parametric analogues of the d-prime statistic. Mean A-prime values by age and experimental conditions are presented in Figure 1.

Table 1. Mean proportions of Hits and False Alarms (FA) by age and experimental condition.

	Children		Adults	
	Hits	FA	Hits	FA
Induction	.72	.41	.83	.76
Blocked Categorization	.75	.54	.80	.50
Baseline (Memory)	.77	.50	.88	.47

A-prime values were submitted to the two-way (age level by experimental condition) ANOVA. The analysis indicated a significant age by condition interaction, $F(2, 137) = 4.8$, $MSE = .268$, $p < .01$. This result supports our qualitative prediction, indicating that the change in the task affected performance of adults but not children. The main effect of the experimental condition was significant, $F(1, 137) = 3.6$, $MSE = .199$, $p < .03$, whereas the main effect of age was not, $F(1, 137)$, $MSE = .121$, $p < .143$. Interaction is the most important: as predicted, A-prime in children did not differ across the conditions, whereas A-prime in adults was markedly lower in the Induction condition than in the other two conditions, $F(2,$

140) = 3.5, $MSE = .207$, p < .03. Furthermore, as predicted, in the Induction condition young children exhibited greater accuracy than adults, one-tailed $t = 1.8$, p < .06. Finally, there were more young children exhibiting near ceiling accuracy in rejecting critical distracters in the induction condition than there were adults: 41% of young children had A-prime values > .8, while only 10% of adults did. Also, only 8% of children had A-prime values < .25, whereas over 24 % of adults had these low A-primes.

Figure 1. Mean A-prime values by age and experimental condition in the recognition phase.
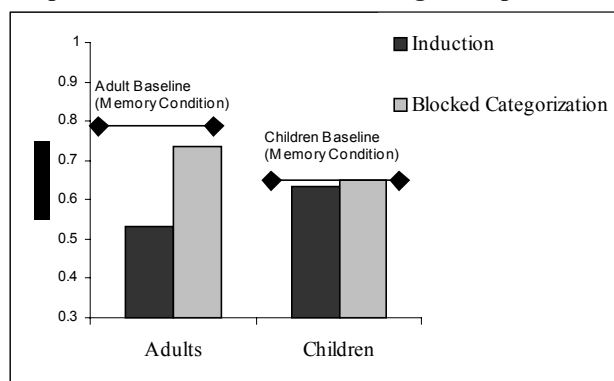
Overall, these results indicate that the study phase task does not affect memory performance of young children – across tasks they distinguished Old Targets from distracters equally well, whereas adults performed significantly better under the memory and blocked categorization conditions than they did under the induction condition. When the task was to induce biological properties from one member of a category to the other members of this category, adults were significantly more likely to false alarm on the members of this target category during recognition than in the other two conditions. Therefore, it seems likely that adults performed induction on the basis of categorization. At the same time, when the task was to induce biological properties, young children did not increase the rate of false alarms compared to the other conditions. Furthermore, in this condition, young children exhibited greater overall accuracy than adults, there was a greater proportion of young children than adults exhibiting near ceiling accuracy, and there was a smaller proportion of young children than adults exhibiting near floor accuracy.

These finding indicate that while adults construct a category–level representations, young children construct item-level representations when performing induction. Therefore, these findings do not support the theory that young children's induction is categorization-based. These findings are rather consistent with the similarity-based model of induction as they indicate that even at 5 years of age, there are many children who, unlike adults, perform induction by comparing each item in question with the Target.

References

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

Brainerd, C. J., Reyna, V. F., & Forrest, T. J. (2002). Are young children susceptible to the false-memory illusion? *Child Development, 73*, 1363-1377.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17-22.

Gelman S. A., & Markman E. M. (1986). Categories and induction in young children. *Cognition, 23*, 123-209.

Gelman S. A., & Markman E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development, 58*, 1532-1541.

Gelman, S. A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology, 20*, 65-95.

Graham S.A., Kilbreath C. S., & Welder A. N. (2001). Words and Shape Similarity Guide 13 month-olds' inferences about nonobvious object properties. *Proceedings of the XXIII Annual Conference of the Cognitive Science Society*, 352-357.

Koutstaal, W., & Schacter, D. L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory and Language, 37*, 555-583.

Mandler J. M., & McDonough L. (1996). Drinking and driving don't mix: Inductive generalization in infancy. *Cognition, 59*, 307-335.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not present in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 803-814.

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

Sloutsky, V. M., & Lo, Y-F. (1999). How much does a shared name make things similar? Part 1. Linguistic labels and the development of similarity judgment. *Developmental Psychology, 6*, 1478-1492.

Sloutsky, V. M., Lo, Y-F., & Fisher, A. V. (2001). How much does a shared name make things similar? Linguistic labels, similarity, and the development of inductive inference. *Child Development, 72*, 1695-1709.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*, 34-50.

Underwood, B.J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology, 70*, 122-129.