INDUCTIVE GENERALIZATION: UNDERLYING MECHANISMS AND DEVELOPMENTAL COURSE

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Anna V. Fisher, M.A.

The Ohio State University
2005

Dissertation Committee:
Professor Neal Johnson, Adviser
Professor Vladimir Sloutsky, Co-Adviser
Professor John Opfer
Professor James Todd

Approved by

Adviser
Graduate Program in Psychology
ABSTRACT

The ability to generalize from the known to the unknown is crucial for learning, and this ability is known to develop early in life. However, the mechanisms underlying early generalization remain contested. Some researchers argue that even early in development people rely on conceptual assumptions to perform induction. In particular, people, including very young children, assume that animals that share category labels belong to the same kind, and animals of the same kind share many properties. Therefore, according to this approach, people rely on common labels to first categorize presented entities, and then to perform category-based induction. However, other researchers suggest that early induction is similarity-based, and that labels are features of objects contributing to the overall similarity of entities. Furthermore, according to this approach, the ability to perform category-based induction is a product of learning. The current series of experiments was designed to differentiate between these theoretical approaches to early induction. Present work argues that if different processes underlie early and mature induction, young children and adults should form different memory traces when performing induction, and therefore, patterns of memory accuracy can be used to infer mechanisms underlying induction at different points in development. In particular, similarity-based induction should lead to encoding of item-specific information and therefore to accurate recognition on a subsequent memory test. At the same time, category-based induction should lead to encoding of predominantly
category-specific information, and therefore to poor recognition. Results of 10 experiments reported below suggest that 5- and 7-year-olds spontaneously perform similarity-based induction when presented with members of familiar categories, whereas adults perform category-based induction. Furthermore, category labels do not automatically promote category-based induction in young children. However, children can be trained to perform category-based induction by relying on shared labels, and retention of training is a function of age: 7-year-olds are more likely to retain training over short delays than 5-year-olds. Moreover, under certain conditions, associative training in label-based induction can be as effective as conceptual training. These results cast doubt on the assertion that young children rely on conceptual assumptions to perform induction, and support the similarity-based approach to early induction.
To my parents Galina and Valeriy
ACKNOWLEDGMENTS

I would like to express my deepest gratitude to Dr. Vladimir Sloutsky for nurturing me throughout my graduate career, for providing me with an opportunity to be a part of the Cognitive Development Lab, for intellectual stimulation, for generously sharing his research ideas, for guidance and inspiration, and for being a wonderful adviser, colleague, and a friend.

I would also like to thank Dr. Neal Johnson for an opportunity to work in the Behavioral Sciences Lab and be a part of the psycholinguistics seminar, for his kindness and endless support he provided over my years in the graduate school.

I thank Dr. John Opfer for our numerous theoretical discussions, for providing me with a challenge, for his input into my research, and for sharing valuable practical advice.

I thank Dr. James Todd for his advice and support, and for his helpful suggestions that he shared with me over the years.

I am very grateful to Tia Foster, Marie Elam, Loren Marulis, Kathy Ryan, Michelle Weiden, Meredith Williams, and David Valentine-Cooper for collecting the data that this thesis is based on.

I thank all of the members of the Cognitive Development Lab, past and present, for stimulating discussions, great advice, support, and for being wonderful and fun people to work with: thank you, Aaron Yarlas, Chris Robinson, Heidi Kloos, Brad Morris, Ya-Fen Lo, Jennifer Kaminski, Julie Hupp, Jakie von Spiegel, Amanda Napolitano, Erica
Howard, Renee Wilson, Cal Timbrook, Sravani Vinapamula, Aaron Rader, and Margie Spino. I am especially grateful to Aaron Yarlas for assistance with data collection for Experiment 6 at the Grand Valley State University.

I thank members of the Behavioral Sciences Lab for their input into my research, for support, and for being wonderful colleagues: thank you, Scott Jankowski, Leslee Martin, and Zach Schendel.

I would also like to thank the National Science Foundation for financial support of the research that this thesis is based on, through a grant To Vladimir Sloutsky.

And last but not least, I thank my wonderful family - my dear husband, Ivan Medvedev, for his love and support, my parents Galina and Valeriy, and my grandparents, Antonina and Alexander, for their love and faith in me.
VITA

May, 13, 1977……………… Born – Kharkov, Ukraine
1998 – 1999……………… Teaching Associate, Moscow Pedagogical State University
1999………………………. B.A. Education, Moscow Pedagogical State University
1999 – 2000…………….. Graduate Administrative and Technical Associate, The Ohio State University
2002……………………..M. A., The Ohio State University, Department of Education
2004………………….. Graduate Teaching Associate, The Ohio State University
2000 – 2005……………. Graduate Research Associate, The Ohio State University

PUBLICATIONS

Research Publications


FIELDS OF STUDY

Major Field: Cognitive/Experimental Psychology
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>VITA</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER 1. KNOWLEDGE-BASED AND SIMILARITY-BASED APPROACHES TO EARLY GENERALIZATION</td>
<td>3</td>
</tr>
<tr>
<td>CHAPTER 2. BASIC EXPERIMENTAL PARADIGM</td>
<td>21</td>
</tr>
<tr>
<td>CHAPTER 3. EFFECTS OF TO-BE-GENERALIZED PROPERTIES ON INDUCTION IN YOUNG CHILDREN</td>
<td>25</td>
</tr>
<tr>
<td>CHAPTER 4. LEARNING AND DEVELOPMENTAL COURSE OF CATEGORY-BASED INDUCTION</td>
<td>28</td>
</tr>
<tr>
<td>CHAPTER 5. EFFECT OF CATEGORY LABELS ON INDUCTION AND VISUAL RECOGNITION</td>
<td>50</td>
</tr>
<tr>
<td>CHAPTER 6. EFFECTS OF ASSOCIATIVE TRAINING ON INDUCTION AND RECOGNITION ACCURACY</td>
<td>71</td>
</tr>
<tr>
<td>CHAPTER 7. POSSIBLE MECHANISM OF CATEGORIZATION: EFFECTS OF DELAYED RECOGNITION ON MEMORY ACCURACY</td>
<td>84</td>
</tr>
<tr>
<td>CHAPTER 8. GENERAL DISCUSSION</td>
<td>91</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>103</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mean proportions of Hits and False Alarms (FA) in Experiment 1 by age group and experimental condition</td>
</tr>
<tr>
<td>2</td>
<td>Mean proportions of Hits, False Alarms on critical lures (FA), and A-prime scores by age and condition in Experiment 3</td>
</tr>
<tr>
<td>3</td>
<td>Mean proportions of Hits, False Alarms on critical lures (FA), and A-prime scores by age in Experiment 4</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Example of a morphing sequence</td>
</tr>
<tr>
<td>2</td>
<td>Examples of triads representing different similarity ratios: (A) Similarity ratio = 1; (B) Similarity ratio = 1.2; (C) Similarity ratio = 1.86; (D) Similarity ratio = 9</td>
</tr>
<tr>
<td>3</td>
<td>Aggregated findings of the research testing quantitative predictions of the SINC model</td>
</tr>
<tr>
<td>4</td>
<td>Aggregated findings of the research using Induction-then-Recognition paradigm. The dashed line represents the point of no sensitivity</td>
</tr>
<tr>
<td>5</td>
<td>Examples of stimuli used in the Basic Paradigm</td>
</tr>
<tr>
<td>6</td>
<td>Memory sensitivity scores (A-prime) across four age groups and three experimental conditions in Experiment 2. The dashed line represents the point of no sensitivity. Error bars represent the standard errors of the mean</td>
</tr>
<tr>
<td>7</td>
<td>Change in the A-prime scores in Induction compared to the Baseline across different age groups in Experiment 2. Error bars represent standard errors of the mean</td>
</tr>
<tr>
<td>8</td>
<td>Memory sensitivity scores (A-prime) in the Induction condition for 5- and 7-year-olds in Experiments 2-4. The dashed line represents the point of no sensitivity. Error bars represent the standard errors of the mean</td>
</tr>
<tr>
<td>9</td>
<td>Examples of stimuli used in Experiment 5</td>
</tr>
<tr>
<td>10</td>
<td>Memory Sensitivity A-prime scores for children (panel 10-A) and adults (panel 10-B) in Experiment 1. Error bars represent the standard errors of the mean. The dashed line represents the point of no sensitivity</td>
</tr>
<tr>
<td>11</td>
<td>Proportion of Hits and False Alarms (FA) in adult participants in the Baseline (panel 11-A) and Induction (panel 11-B) conditions Experiment 6. The error bars represent the standard errors of the mean. The dashed line represents chance level</td>
</tr>
</tbody>
</table>
12 Proportion of Hits and False Alarms (FA) in 5 year-olds in the Baseline (panel 12-A) and Induction (panel 12-B) conditions in Experiment 6. The error bars represent the standard errors of the mean. The dashed line represents chance level................................................................. 60

13 Memory Sensitivity A-prime scores for 5 year-olds in the Category Labels condition in Experiments 6 and 7. The error bars represent the standard errors of the mean. The dashed line represents the point of no sensitivity ............. 67

14 Proportion of Hits and False Alarms (FA) in 5 year-olds in the Baseline condition (panel A) and Induction condition (panel B) in Experiments 6 and 7. The error bars represent the standard errors of the mean. The dashed line represents chance level................................................................. 68

15 Examples of training stimuli used in Experiment 8. Notice, that all labeling and property information was presented auditorily ......................................................... 75

16 Memory sensitivity A-prime scores in the Induction and Baseline conditions in Experiments 7-8. Error bars represent the standard errors of the mean. Dashed line represents the point of no sensitivity ................................................. 77

17 Proportions of Hits and False Alarms in the Induction and Baseline conditions of Experiment 8. Error bars represent the standard errors of the mean. Dashed line represents chance level ................................................................. 78

18 Proportions of Hits and False Alarms in the Induction and Baseline conditions of Experiment 9. Error bars represent the standard errors of the mean. Dashed line represents chance level ................................................................. 82

19 Proportions of Hits and False Alarms in the Induction and Baseline conditions of Experiment 10. Error bars represent the standard errors of the mean. Dashed line represents chance level ................................................................. 88
INTRODUCTION

The ability to make inductive generalizations is crucial for acquiring new knowledge. For instance, upon learning that a particular cat uses serotonin for neural transmission, one can generalize this knowledge to other felines and possibly other mammals. The ability to perform induction appears very early in development (Gelman & Markman, 1986; Sloutsky & Fisher, 2004a, 2004b; Welder & Graham, 2001), however the underlying mechanisms remain unknown.

One possibility is that inductive generalization is a function of similarity between a source and a target. If similarity is construed as a function of overlap among elemental or relational features, similarity can be computed, and several researchers proposed models of such similarity computations (Estes, 1994; Goldstone, Medin, & Gentner, 1991; Nosofsky, 1984; Tversky, 1977). There is also empirical evidence that similarity among compared entities does play an important role in inductive generalizations of both young children (e.g., Jones & Smith, 2002; Smith, Jones, & Landau, 1996; Sloutsky et al., 2001; Sloutsky & Fisher, 2004a, 2004b) and adults (Osherson, et al., 1990; Sloman, 1993).

However, it has been counter argued that similarity cannot be used as an explanatory mechanism, because similarity is itself a construct in need of an explanation. For instance, Goodman (1992) argued that similarity is a “useless relation”, since two things can be considered similar if they have at least one property in common, and any two
entities always have at least one such property: For instance, an apple and a dog are both smaller than a cow. Therefore, it is unclear how similarity could direct induction when dimensions for similarity computation are unlimited and unspecified.

An alternative theoretical proposal was developed to address this problem. According to this proposal, generalization is constrained by previous knowledge, often in the form of naïve theories that people hold about the world (Murphy & Medin, 1985). Proponents of the knowledge-based approach claim that these theories are domain-specific, and that even very young children generalize on the basis of such theory-like beliefs when reasoning about natural kinds (Carey, 1985; Gelman & Coley, 1991; Gelman & Wellman, 1991; Keil, 1989). Therefore, people would not make an inference from an apple to a dog just because both have some features in common (i.e., both are smaller than a cow), since they believe that not all features are equally important, and that some features are more important than others. Instead, people generalize on the basis of conceptual assumptions, rather than surface similarities among objects (see Atran 1994; Keil, 1981; Keil, Smith, Simons, & Levin, 1998; Murphy, 2002, for reviews of these assumptions).

Both, the similarity-based and the knowledge-based approach have generated significant empirical support. In Chapter 1 I will review these theoretical positions and evidence supporting them.
CHAPTER 1

KNOWLEDGE-BASED AND SIMILARITY-BASED APPROACHES TO EARLY GENERALIZATION

As mentioned in the Introduction, there have been proposed two theoretical approaches to induction – a knowledge-based approach (often referred to as a naïve theory) and a similarity-based approach. In this chapter I review theoretical arguments and empirical evidence presented by proponents of both approaches.

KNOWLEDGE-BASED APPROACH TO EARLY GENERALIZATION

Many researchers believe that organization of children’s knowledge is theory-like in nature. It has been proposed that young children possess naïve theories of mind (Wellman, 1990), physics (Baillargeon, 1987; Spelke, 1990), and biology (Keil, 1994; Simons & Keil, 1995; Hatano & Inagaki, 1999). The latter theory guides people’s inferences about living things: It has been proposed that people, including young children, rely on several conceptual assumptions (Keil, Smith, Simons, & Levin, 1998; also see Murphy, 2002 for a review of these assumptions) when making inductive generalizations. In particular, they rely on the (1) Category Assumption, the assumption that each individual entity is a member of a category (i.e., each individual bear is a
member of the category “BEAR”) and (2) the Linguistic Assumption, the assumption that 
linguistic labels are category markers providing direct access to conceptual knowledge. 
According to the naïve theory proposal, linguistic labels convey essential properties of 
objects, and are therefore central for induction, whereas perceptual features are 
peripheral. As a result, when the labeling information is provided, children first identify 
presented entities as members of categories (i.e., assume that entities sharing a common 
label belong to the same category), and then induce properties on the basis of this 
categorization.

Note, that conceptual assumptions discussed above are argued to be a prerequisite for 
acquiring new knowledge rather than a product of learning and development. Therefore, 
children do not learn that labels are category markers; rather the label-as-a-marker 
assumption is necessary for the learning to occur (Gelman, 2003).

Evidence supporting the naïve theory position comes from several lines of research. 
First, there is evidence that children, as well as adults, are more willing to generalize 
properties that are biological rather than social or psychological in nature (Springer & 
Keil, 1989; Springer, 1992). For instance, when asked which features a baby would 
inherit from a parent, both children and adults made inferences about such biological 
properties as “have tiny bones inside them”, but not about non-biological properties such 
as “very dirty from playing in the mud” or “knows a really good place to find food” 

Second, it has been argued that adults and young children generalize on the basis of 
category membership information, even when appearance information conflicts with 
category information (Gelman and Coley, 1990; Gelman & Markman, 1987). Thus, when
presented with a bluebird and told that this bird lives in a nest, a child is more likely to generalize this property to a dissimilarly looking dodo bird than to a similarly looking pterodactyl, because the child presumably understands that since both, a dodo and a bluebird are referred to with the same name (i.e., “bird”) they belong to the same category, and therefore share many important properties (Gelman and Coley, 1990).

Third, young children exhibit different patterns of responses when asked about transformations of natural kind objects and artifacts over time. For example, when presented with a baby animal, a plant sprout, or a small artifact object and asked about the change that these items are likely to undergo over time, 4- to 5-year-old children claim that artifacts should not undergo any size changes, whereas animals and plants should increase in size over time (for review see Hatano & Inagaki, 1999).

Further support for the knowledge-based approach came from the studies on the maintenance of identity over transformations (Gelman & Wellman, 1991; Keil, 1989). Participants in these experiments were told a story, for instance, about a skunk that was surgically altered to look like a raccoon. These studies demonstrated that children judged the category membership of natural kind objects on the basis of their origins (i.e., being born from a skunk) and not perceptual features (Keil, 1989).

SIMILARITY-BASED APPROACH TO EARLY GENERALIZATION

The alternative approach to inductive generalization requires little to no reliance on conceptual knowledge (Sloman, 1993; Sloutsky & Fisher, 2004a). Instead, it is suggested that people, including young children, generalize on the basis of multiple commonalities.
among entities. According to this position, it is unnecessary to postulate multiple assumptions about innate or early developing conceptual knowledge to explain early generalizations: advocates of this position argue that powerful learning mechanisms coupled with simple attentional processes are responsible for many “smart behaviors” exhibited by young children (Sloutsky & Fisher, 2004b; Smith, Jones, Landau, 1996).

One of the claims of the naive theory approach is that young children often ignore perceptual information in favor of conceptual information. However, several recent studies found little evidence to support this claim. For instance, when supposedly conceptual information (i.e., linguistic labels) conflicts with perceptual information (i.e., patterns of motion) children were found to rely on perceptual information to make inductive generalizations (Mak & Vera, 1999). Moreover, it has been demonstrated that results of the earlier studies claiming that children disregard conflicting appearances in favor of conceptual cues were driven by high within-set similarity of the stimuli: when all compared entities are highly similar, appearances become poor predictors, and children rely on other cues (i.e., labels) to perform induction. However, when predictiveness of appearances increases, children are as likely to rely on matching appearances as on matching labels (Sloutsky, 2003; Sloutsky & Fisher, 2004a).

Furthermore, there is a general agreement that category-based generalization is not the only kind of generalization early in development. There is ample evidence that infants are capable of using similarity as a basis for making generalizations (Quinn & Eimas, 1997). It has also been demonstrated that young children often generalize in a similarity-based manner. First, young children perform similarity-based generalizations when presented entities are not labeled or when all presented entities are accompanied by
identical labels (Fisher & Sloutsky, 2004; Sloutsky & Fisher, 2004a). For instance, when Sloutsky & Fisher (2004b) presented 5 year-olds with triads of animal pictures, none of which was labeled, and asked participants to generalize a property from one member of the triad to the rest of the members, young children’s generalizations were guided by perceptual similarity of presented entities.

Second, when the task involves artifacts rather than natural kind objects, young children often rely on perceptual similarity information over information about objects’ function when generalizing a label or category membership (Gentner, 1978; Keil, 1989). For example, Gentner (1978) presented participants with two kinds of artifact objects, which had different functions and appearances: one of the objects, called a jiggy, consisted of a square box that would light-up when a lever was pressed, and the other object, a modified gumball machine called zimbo, would drop jellybeans when a lever was pressed. Participants were then presented with hybrid objects, in which functions and appearances were crossed, and asked whether the hybrid object was a jiggy or a zimbo. Unlike 5- to 15-year-olds who often extended labels based on the object’s function, 2.5-year-olds generalized labels based on perceptual similarity of presented objects.

Third, even proponents of the knowledge-based approach presented evidence that children make similarity-based generalizations. For example, Gelman (1988) presented 4- and 7-year-olds with a task in which they had to generalize novel properties, such as “used to make fumet” (a property generalizable for artifacts) or “has pectin inside” (a property generalizable for natural kinds). It was found that for both kinds of properties children made similarity-based generalizations. Specifically, children were very likely to generalize the property of having pectin from a red apple to another red apple, less likely...
to generalize this property to a yellow apple, even less likely to generalize this property to a banana, and unlikely to generalize this property to a stereo.

Therefore, the important role of perceptual similarity in early generalization has been well documented. Furthermore, there have been numerous demonstrations that perceptual and attentional mechanisms alone can potentially account for much of the learning in infancy and early childhood in a variety of domains, such as categorization (French, Quinn, & Mareschal, 2001), word learning (Smith, Jones, & Landau, 1996), and numerosity (Mix, Huttenlocher, & Levine, 2002). Therefore, the main point of controversy is the following: is it necessary to postulate multiple generalization strategies, when most of the empirical evidence can be accounted for by perceptual and attentional mechanisms alone?

The advantages and disadvantages of both theoretical approaches are well known. Theories relying on similarity computations run into the problem pointed out by Nelson Goodman (1992): they need to specify what attributes the similarity is computed over, and give a reason why observers would take into account some but not other dimensions. Theories relying on prior knowledge avoid this problem altogether, however they postulate other assumptions (such as multiple inductive generalization strategies) that need justification. An alternative model of early generalizations has been recently proposed by Sloutsky and his colleagues (Sloutsky & Lo, 1999; Sloutsky et. al., 2001; Sloutsky & Fisher, 2004a). In the next section I provide a brief review of this model and empirical evidence supporting it.
THE SIMILARITY–INDUCTION–NAMING–CATEGORIZATION (SINC) MODEL

According to SINC, induction and categorization are variants of the same process based on the overall similarity computation. Due to space limitations, only a brief sketch of the model is presented in what follows, while a more detailed account is presented elsewhere (Sloutsky & Fisher 2004a).

Unlike other similarity-based models, SINC treats linguistic labels as object features (rather than conceptual markers); therefore similarity of compared entities is computed over both, visual and auditory attributes:

\[
Sim(i, j) = W_{Label}^{1-L} S_{vis.attr}^{N-k} \\
\left\{ \begin{array}{l l}
L = 1, & \text{if } L_i = L_j \\
L = 0, & \text{otherwise}
\end{array} \right.
\]

where, \( N \) denotes the total number of visual attributes, \( k \) denotes the number of matches, \( S_{vis.attr.} \) denotes attentional weights of a mismatch on a visual attribute, \( W_{Label} \) denotes values of label mismatches, and \( L \) denotes a label match. When there is a label match, \( L = 1 \), and \( W_{Label} = 1 \); when there is a label mismatch, \( L = 0 \), and \( W_{Label} < 1 \). Note that \( S \) and \( W \) (\( 0 \leq S, W \leq 1 \)) denote attentional weights of mismatches and the contribution of \( S \) and \( W \) is large when these parameters are close to 0 and is small they are close to 1.

In the case when entities are not labeled, Equation 1 reduces to the product-rule model of similarity (Estes, 1994). However, when labels are introduced, the model predicts that common labels (\( L = 1 \)) will increase the similarity of compared entities, whereas distinct labels (\( L = 0 \)) will lead to the decrease in the similarity of compared
objects. This prediction of the model has been tested with young children and adult participants. It has been demonstrated that linguistic labels contribute to the similarity judgments of young children but not adults (Sloutsky & Fisher, 2004a); therefore, further predictions of the model have been restricted to young children.

When a child is presented with a stimulus set consisting of three non-labeled entities, a Target stimulus (T) and two test stimuli (A and B), and asked to estimate which of the test stimuli is more similar to the Target, the model claims that the child’s choices can be predicted using a variant of the Luce’s choice rule (Luce, 1959) presented in Equation 2:

\[
P(B) = \frac{S^x}{S^x + S^y} = \frac{S^x}{S^x (1 + S^{y-x})} = \frac{1}{1 + \frac{S^y}{S^x}}
\]

where \(S^x\) and \(S^y\) denote the similarity of test A and B to the Target respectively. When features of objects can be individuated (i.e., when schematic line drawings are used) it is easy to estimate featural overlap, and to calculate the similarity among compared objects using the product-rule model of similarity (Estes, 1994) when entities are not labeled, or Equation 1 when entities are labeled. However, when naturalistic perceptually rich entities are used (i.e., color photographs of real objects) similarity estimates need to be obtained empirically, since individuation of features is not possible. The methodology for obtaining similarity estimates \(S^x\) and \(S^y\) for naturalistic perceptually rich stimuli was developed by Sloutsky & Fisher (2004a).

In the situation when entities are labeled, children’s choices on the similarity judgment task can be predicted using Equation 3:
Notice, that the derivations remain essentially the same, except for the $W_{Label}$ parameter. This parameter equals to 1, if there is a label match, otherwise it varies from 0 to 1, and the smaller the value of $W$, the greater the contribution of label mismatch.

If it is the case that categorization and inductive generalization are similarity-based processes, then Equations 2 and 3 should predict performance of young children on categorization and induction tasks as well as on the similarity judgment tasks, when entities are not labeled or labeled respectively.

Quantitative predictions of the SINC model were tested in a series of experiments with young children (Sloutsky & Fisher, 2004a). These experiments used a triad task in which participants are presented with a Target and two Test stimuli, and asked to make generalizations from one of the Test items to the Target. As follows from Equations 2 and 3, quantitative predictions of the model are a function of the similarity ratios of each of the Test stimulus to the Target (i.e., $S^x / S^y$). Estimates of various similarity ratios were obtained in a separate calibration experiment in which children were presented with triads of animal pictures and asked whether the Target looked more like Test A or Test B. Triads of pictures consisted of images obtained by morphing the images of two different animals into each other (see Figure 1).

As a result of calibration, 16 triads representing four different similarity ratios were selected for experimentation: similarity ratio $= 1$ (both test items are equally similar to

\[
P(B) = \frac{S^x}{S^x + WS^y} = \frac{S^x}{S^x(1 + WS^{y-x})} = \frac{1}{1 + \frac{WS^y}{S^x}}
\]
the Target), similarity ratio = 1.22 (Test A is slightly more similar to the Target than Test B), similarity ratio = 1.86 (Test A is noticeably more similar to the Target), and similarity ratio = 9 (Test A is almost identical to the Target and Test B is very different). Examples of triads representing different similarity ratios are presented in Figure 2.

![Step 1 Step 5 Step 10 Step 15 Step 20](image)

Figure 1. Example of a morphing sequence.

To estimate the unique contribution of appearances and labels to induction and categorization in young children, these two sources of information were pitted against each other: Test B, which was equally or less similar to the Target than Test A, always shared the label with the Target, whereas Test A had a different label. If children do indeed rely on matching labels and are able to ignore conflicting appearances to make generalizations (as the proponents of the naïve theory claim), children should consistently induce from Test B to the Target regardless of the similarity ratios. If however, SINC is correct, children’s reliance on matching labels (i.e., proportion of Test-B choices in this case) should decrease as the similarity ratio increases.
Figure 2. Examples of triads representing different similarity ratios: (A) Similarity ratio = 1; (B) Similarity ratio = 1.2; (C) Similarity ratio = 1.86; (D) Similarity ratio = 9.
Figure 2 continued

C.

Figure 2. Examples of triads representing different similarity ratios: (A) Similarity ratio = 1; (B) Similarity ratio = 1.2; (C) Similarity ratio = 1.86; (D) Similarity ratio = 9.
Aggregated findings of this research are presented in Figure 3. Data in the figure indicate that as predicted by SINC, children’s reliance on matching labels was a function of the similarity ratios of the triads. When the similarity ratio was low (i.e., when overall similarity of entities in the set was high, and similarity was therefore non-predictive) children tended to rely on labels to perform induction and categorization tasks; however, when similarity ratio was high (i.e., similarity was predictive) children were as likely to rely on matching labels as on matching appearances.

Figure 3. Aggregated findings of the research testing quantitative predictions of the SINC model.
However, proponents of the naïve theory approach can counter argue that while young children’s performance on induction tasks is influenced by the similarity of presented entities to some degree, even early in development people’s generalizations are guided mainly by their conceptual assumptions. Therefore, traditional approaches to the study of induction may not be an optimal way of examining mechanisms underlying performance on induction tasks, because these approaches cannot provide evidence distinguishing between the similarity-based and knowledge-based approaches. An alternative framework has been recently suggested (Fisher & Sloutsky, 2005; Sloutsky & Fisher, 2004a, 2004b). In this framework, mechanisms underlying induction performance are studied by examining memory traces formed in the course of induction. From the work of Koutstaal & Schacter (1997) we know that spontaneous categorization often leads to memory intrusions in the form of false recognition of “critical lures” - unstudied items that belong to the same category as the studied items. Therefore, if generalizations are made on the basis of category membership, then performance on a memory test preceded by an induction task should reveal low discrimination of studied items from non-studied items (compared to a baseline no-induction condition). However, if participants perform similarity-based rather than category-based induction, we should expect their memory to be as accurate in the induction as in the baseline conditions.

Furthermore, it is possible to predict not only the overall recognition accuracy of participants, but also specific patterns of hits (i.e., responses, reflecting true recognition of the previously studied items) and false alarms (i.e., responses, reflecting false recognition of novel items). It has been demonstrated that deeper conceptual processing (as opposed to more shallow perceptual processing) leads to the increase in the rates of
both, hits and false alarms (Rhodes & Anastasi, 2000; Thapar & McDermott, 2001). Therefore, if participants perform category-based induction, we can predict not only that their overall recognition memory accuracy will be low, but also that the rate of hits and false alarms will be high. On the other hand, if participants perform similarity-based induction, we can predict that their high recognition memory accuracy will be generated by the high level of hits and low level of false alarms.

The above considerations made possible a nontrivial prediction (Fisher & Sloutsky, 2005; Sloutsky & Fisher, 2004a, 2004b), that in a recognition memory test following an induction task about members of familiar animal categories, young children (who were expected to perform similarity-based induction) should be more accurate that adults (who were expected to perform category-based induction). This prediction is surprising, because developmental studies show a general improvement of memory performance with age (see Schneider & Pressley, 1997, for a review).

The predictions outlined above were confirmed in a series of experiments (Fisher & Sloutsky, 2005; Sloutsky & Fisher, 2004a, 2004b). Specifically, adults were highly accurate on a recognition memory test in the no-induction baseline condition, however after performing an induction task their discrimination of studied items from critical lures dropped to chance. In contrast, young children’s discrimination remained high on the memory test in both, the induction and no-induction conditions. In other words, after performing induction adults could not differentiate studied items from critical lures, because in the course of induction they encoded the category-level information, but not the item-specific information. For instance, they might remember that they saw pictures of cats, but fail to remember whether they had seen a particular picture of a cat. Children,
on the other hand, encoded item-specific information, and therefore were more accurate than adults on the memory test. However, when young children were trained to perform category-based induction before the experiment proper, their recognition accuracy decreased dramatically to the level of adults. The training, however, had no adverse affect on the memory accuracy of participants in the no-induction baseline condition, therefore it is unlikely that accuracy decrease in the induction condition was due to fatigue or other extraneous factors. Aggregated findings of these studies are presented in Figure 4. The Figure represents memory sensitivity A-prime scores. A-prime is a non-parametric analogue of the signal-detection d-prime statistic (Snodgrass & Corwin, 1988). An A-prime score of .5 indicates that participants do not discriminate studied items from critical lures, and as discrimination accuracy increases, A-prime scores approach 1.

Furthermore, adult participants in these studies demonstrated above-chance levels of false alarms (FA) only in the Induction condition (i.e., FA = .74), but not in the no-induction Baseline (i.e., FA = .40), thus suggesting involvement in deep versus shallow encoding respectively. Children, on the other hand, exhibited chance-level rates of false alarms in both testing condition (i.e., FA = .41 and .59 in the Induction and Baseline conditions respectively). However, after children were trained to perform category-based induction, the level of false alarms dramatically increased in the Induction condition, but not in the Baseline condition (i.e., FA = .70 and .40 after training in the Induction and Baseline conditions respectively). These findings demonstrate that unlike adults, young children spontaneously performed similarity-based generalizations.
However, these studies left a number of important questions unanswered. First, the developmental course of category-based induction is unclear. It is possible that 5 year-olds are a “transitional” group: since it was relatively easy to train young children to perform category-based induction, it is conceivable that by 5 years of age children are on the verge of spontaneously generalizing in a category-based manner. Second, the role of category labels in early induction requires further investigation. If category labels promote category-based induction, then providing labels in the Induction-then-Recognition paradigm should lead to an increased level of false recognition in young children. On the other hand, this should not happen if labels for young children are features of objects rather than category markers. Third, it remains unknown whether
associative training (as opposed to conceptual training used in Sloutsky & Fisher 2004a, 2004b) can influence inductive generalizations in young children. Finally, it is unclear whether the nature of to-be-generalized properties affects the nature of induction in young children: specifically, it is possible that properties that are perceived to be more “biological” promote category-based induction in young children to a larger extent than blank predicates. The series of experiments presented in Chapters 3-7 was designed to address these important questions.
CHAPTER 2

BASIC EXPERIMENTAL PARADIGM

This Chapter presents the Basic Paradigm and materials that were used in all of the reported experiments. Any deviations from the Basic Paradigm are described in detail in the appropriate chapters. Materials in Experiments 1-10 were 44 color photographs of familiar animals used in the previous research (Fisher & Sloutsky, 2005; Sloutsky & Fisher, 2004a, 2004b). Experiments consisted of two phases – a Study Phase and a Recognition Phase. During the Study Phase, participants were presented with 30 pictures representing three different categories (10 cats, 10 bears, and 10 birds), one picture at a time. During the Recognition Phase, participants were presented with 28 pictures, half of which were previously presented during the Study Phase, and the other half were new pictures. The recognition pictures consisted of (a) previously presented pictures (7 cats and 7 bears), (b) novel pictures from the studied category (7 cats, which served as critical lures), and (c) novel pictures from a non-studied category (7 squirrels, which served as catch items). Examples of stimuli are presented in Figure 5.

There were three between-subjects experimental conditions – Baseline, Induction, and Blocked Categorization. In the Study Phase of the Baseline condition participants were presented with 30 pictures of animals, and their task was to remember these pictures for a subsequent recognition test.
Figure 5. Examples of stimuli used in the Basic Paradigm.
In the Study Phase of the Induction condition participants were first presented with a picture of a cat, and informed that it had “beta-cells inside its body”. Participants were then presented with 30 pictures of animals (identical to those presented in the Baseline condition), and asked whether each of the animals also had beta-cells inside. After responding, participants were provided with “yes/no” feedback, indicating that only cats, but not bears or birds, had beta-cells. The recognition test was not mentioned in the Study Phase of the Induction condition.

The goal of the Blocked Categorization condition was to control for the difference in task demands between the Baseline condition (which had neither induction task nor feedback) and Induction condition (which had both induction task and feedback). Participants were first presented with a picture of a cat, and informed that the animal was young. Participants were then presented (in a random order) with 30 study pictures of animals (identical to those in the Baseline and Induction conditions), and asked whether each of the animals was young or mature. Participants were provided with random “yes/no” feedback, intended to block inferences based on the animal kind information and to force all participants to concentrate on perceptual features of individual items. Similar to the Induction condition, the recognition test was not mentioned during the study phase.

The Recognition Phase immediately followed the Study Phase, and was identical across the Baseline and Induction conditions. Participants were presented with 28 recognition pictures, and were asked to determine whether each picture was “old” (i.e., exactly the one presented during the Study Phase) or “new.” No feedback was provided during the Recognition Phase.
In all of the experiments reported in Chapters 3-7, children were tested individually in their schools or day care centers by female hypothesis-blind experimenters. Undergraduate students were tested individually in a laboratory on campus. For all participants stimuli were presented on a computer screen in a self-paced manner, and stimuli presentation was controlled by Super Lab Pro 2 software (Cedrus Corporation, 1999). Adult participants responded by pressing a keyboard button, whereas children responded verbally, with their responses recorded by experimenters.
CHAPTER 3

EFFECTS OF TO-BE-GENERALIZED PROPERTIES ON INDUCTION IN YOUNG CHILDREN

Most researchers of induction avoid using familiar attributes (i.e. has red blood) as to-be-generalized properties in order to prevent participants from deploying their knowledge rather than performing an inductive generalization or a reasoning task. Rips (1975) was the first to start using blank predicates to examine induction. Blank predicates communicate clearly the nature of properties in question without being specific in respect to the concepts they describe (i.e., disease X, vitamin N, etc.). Previous research in the Induction-then-Recognition paradigm described in Chapter 2 (e.g., Sloutsky & Fisher, 2004a, 2004b) used a blank predicate “has beta cells in its body”. This blank predicate was intended to communicate to participants that the property was biological in nature. However, it can be argued that the property of having beta cells was not perceived as “biological” by young children, because they are not familiar with the term “cells”. According to this argument, a blank predicate that is more likely to communicate its biological nature to young children should be more likely to elicit category-based induction. Experiment 1 was designed to examine this possibility.
EXPERIMENT 1

Method

Participants

Participants were 13 preschool children (6 girls and 7 boys; $M = 5.4$ years; $SD = .18$ years). In this and all other experiments reported in Chapters 3-7, children were recruited from several daycare centers located in middle class suburbs of Columbus, Ohio.

Design, Materials, and Procedure

Design, materials and procedure were identical to the Induction condition described in the Basic Paradigm (Chapter 1) with one important difference: the blank predicate participants were asked to generalize consisted of terms familiar to 5 year-olds. Specifically, the property in question was “has thick blood inside its body” (according to the MRC Online Database, 1987, the age of acquisition of words blood and body is 2.5 years).

Results and Discussion

Catch Trails Accuracy

During the Recognition Phase participants were highly accurate in rejecting catch items, averaging 100% of correct responses. Therefore, no participants were eliminated from the sample.
Induction Accuracy

During the Study Phase, similar to the previous reports (Fisher & Sloutsky, 2005; Sloutsky & Fisher, 2004a, 2004b) participants performed accurate inductive generalizations, averaging 80% of correct inductions, above chance, one-sample $t (12) > 4.6, p < .01$. It is worth noting that the rate of correct inductions in Experiment 1 was statistically equivalent to the previously observed rate of correct inductions (average induction accuracy across published reports is 79%), $t (73) < 1$, even when children were asked to generalize a property that they were likely to perceive as “biological”.

Also similar to the previous reports, 5 year-olds demonstrated high recognition accuracy, average memory sensitivity A-prime score = .63, above chance, one-sample $t (12) > 2.3, p < .05$. This level of recognition accuracy is also statistically equivalent to the previously reported recognition accuracy levels (average A-prime score across published reports = .66), $t (73) < 1$.

SUMMARY

In Experiment 1, 5 year-olds were asked to generalize a property that they were likely to perceive as “biological”. Children were accurate in both, making correct generalizations and recognition judgments, which suggests that they encoded perceptual details of presented pictures in the course of induction. Furthermore, the levels of both induction and recognition accuracy did not exceed those reported in the previously published research (Sloutsky & Fisher, 2004a, 2004b). Therefore, it can be concluded that young children perform similarity-based induction regardless of the properties of to-be-generalized predicates.

27
As mentioned in Chapter 1, young children spontaneously perform similarity-based induction, but they can be trained to perform category-based induction. Training procedures used in Sloutsky and Fisher (2004a, 2004b) were short and it was relatively easy to train children in category-based induction. Therefore, it can be argued that the ease with which young children performed category-based induction signifies that children did not learn to perform induction in a category-based manner during training; rather training simply activated their pre-existing conceptual knowledge. If this is the case, then it follows that 5-year-olds should be able to retain effects of training: it should not be hard to retain something that is supposed to be already known. At the same time, finding that even after short delays young children revert to similarity-based induction would indicate that the learned knowledge of the category and linguistic assumptions is very fragile, thus undermining the “priming of preexisting knowledge” possibility. One of the goals of experiments presented in Chapter 3 was to examine this issue.

Another goal was to examine the developmental course of category-based induction. In particular, finding that recognition accuracy following an induction task decreases with age would strongly indicate that category-based induction develops gradually, which
would further indicate that it is not a default, but is rather a product of learning and development. At the same time, finding that recognition accuracy decreases with age would be novel and counterintuitive, because memory (in general) is known to increase and not decrease from early childhood to young adulthood (see Schneider & Bjorklund, 1998, for a review).

The four reported experiments were designed to achieve these goals. To establish the developmental pattern of effects of performing induction on recognition memory accuracy, Experiment 2 presented participants from four different age groups (5-, 7-, and 11-year olds, and adults) with the same Induction-then-Recognition memory paradigm that was used by Sloutsky and Fisher (2004a, 2004b). Based on the previous research (e.g., Sloutsky & Fisher, 2004a, 2004b; Sloutsky, et al., 2001), we expected that category-based induction would increase with age, and thus that memory accuracy after an induction task would decrease with age. Another goal was to provide more direct evidence that similarity-based induction results in accurate recognition memory: findings that when forced to perform similarity-based induction, adults exhibit high recognition accuracy would constitute such evidence.

In Experiments 3 and 4, younger children (5- and 7-year-olds) were trained to perform category-based induction, and effects of training on memory accuracy were examined either immediately after training (Experiment 3) or after a delay (Experiment 4). Finally, in Experiment 5, adults were presented with members of novel artificial categories within the Induction-then-Recognition paradigm. It was expected that when presented with members of novel categories adults will be forced to perform similarity-
based processing, and therefore will exhibit no decrease in recognition accuracy in the
Induction condition compared to the Baseline condition.

EXPERIMENT 2

Method

Participants

Participants were 69 5-year-olds (29 girls and 40 boys, $M$ age = 5.11 years, $SD = 0.38$), 49 7-year-olds (26 girls and 23 boys, $M$ age = 7.87 years, $SD = 0.59$ years), 66 11-year-olds (31 girls and 35 boys, $M$ age = 11.87 years, $SD = 0.59$ years), and 60 introductory psychology students (25 women and 35 men, $M$ age = 19.77 years, $SD = 1.14$). In this and all other experiments reported in Chapters 3-7 (except Experiment 6, Chapter 5) adults were undergraduate students at The Ohio State University participating in the experiment in partial fulfillment of a course requirement.

Materials, Design, and Procedure

Materials, design, and procedure were identical to those described in the Basic Paradigm (Chapter 1). There were three between-subject conditions: Baseline, Induction, and Blocked Categorization. Participants in each of the four age groups used in Experiment 2 (5-, 7-, and 11-year olds, and adults) were randomly assigned to one of the testing conditions.

Results and Discussion

Catch Trials Accuracy
Seven 5-year-olds (one in the Induction, two in the Baseline, and three in the Blocked Categorization condition), one 7-year-old (in the Baseline condition), and twelve adults (all in the Blocked Categorization condition) did not reliably reject catch items, and their data were eliminated from further analysis. The rest of participants were highly accurate in rejecting catch items, averaging 96%, 97%, 99%, and 97% of correct responses across conditions for 5-, 7-, 11-year-olds and adults, respectively.

**Induction Accuracy**

In the Induction condition participants promptly realized that the target property should be extended to cats but not to other animals. The rate of correct inductive inferences in 5-, 7-, 11-year-olds and adults was 75%, 85%, 94%, and 91% respectively, all above chance, one-sample $t > 5.7$, $p < .0001$.

**Recognition Accuracy**

Across conditions participants were also very accurate in correctly recognizing old items; however their recognition accuracy on the critical lures (i.e., novel pictures of cats) differed across conditions and age groups. Proportions of hits (i.e., correct recognitions) and false alarms (FA) on critical lures by age group and condition are presented in Table 1. Measures of memory sensitivity (i.e., A-prime scores) by age group and condition are presented in Figure 6. Data in the Table and in the Figure indicate that 5- and 7-year-olds experienced no decrease in memory accuracy in the Induction condition compared to Baseline, whereas memory accuracy of 11 year-olds and adults in the Induction condition decreased dramatically compared to Baseline.
To examine the significance of these tendencies, A-prime scores were subjected to a two-way (Age by Condition) between-subjects ANOVA. The analysis indicated that there was no significant main effect of age, $F (3, 215) < 2.6, p > .05$, whereas the main effect of condition was significant, $F (2, 215) = 9.2, p < .0001$, with accuracy in the Induction condition being significantly lower than accuracy in the Baseline and Blocked Categorization conditions, post-hoc Tukey test, all $p$s < .05. This main effect, however, was driven by a significant Age by Condition interaction, $F (6, 215) = 4.1, p < .001$.

![Figure 6](image_url)

Figure 6. Memory sensitivity scores (A-prime) across four age groups and three experimental conditions in Experiment 2. The dashed line represents the point of no sensitivity. Error bars represent the standard errors of the mean.
<table>
<thead>
<tr>
<th>Condition</th>
<th>5 year-olds</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hits</td>
<td>FA</td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>.82</td>
<td>.59</td>
<td></td>
</tr>
<tr>
<td>Blocked Categorization</td>
<td>.62</td>
<td>.38</td>
<td></td>
</tr>
<tr>
<td>Induction</td>
<td>.70</td>
<td>.41</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Condition</th>
<th>7 year-olds</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hits</td>
<td>FA</td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>.79</td>
<td>.40</td>
<td></td>
</tr>
<tr>
<td>Blocked Categorization</td>
<td>.77</td>
<td>.25</td>
<td></td>
</tr>
<tr>
<td>Induction</td>
<td>.77</td>
<td>.45</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Condition</th>
<th>11 year-olds</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hits</td>
<td>FA</td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>.79</td>
<td>.39</td>
<td></td>
</tr>
<tr>
<td>Blocked Categorization</td>
<td>.77</td>
<td>.31</td>
<td></td>
</tr>
<tr>
<td>Induction</td>
<td>.79</td>
<td>.59</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Condition</th>
<th>Adults</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hits</td>
<td>FA</td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>.88</td>
<td>.40</td>
<td></td>
</tr>
<tr>
<td>Blocked Categorization</td>
<td>.86</td>
<td>.49</td>
<td></td>
</tr>
<tr>
<td>Induction</td>
<td>.81</td>
<td>.74</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Mean proportions of Hits and False Alarms (FA) in Experiment 1 by age group and experimental condition
Follow-up analyses indicated that 5- and 7-year-olds discriminated old items from critical lures well in all three conditions, and there were no significant differences in their levels of discrimination across the task conditions, as indicated by the one-way ANOVAs performed on the A-prime scores, both $F$s < 2.3, $p$s > .1. At the same time, adults and 11-year-olds well discriminated old items from critical lures in the Baseline and Blocked Categorization conditions, whereas their discrimination in the Induction condition was dramatically lower than in the other two conditions, both $F$s > 8.8, $p$s < .001, post-hoc Tukey test $p$s < .05 for all significant differences.

To further examine this developmental trend, differential effects of induction on recognition accuracy were analyzed as a function of age. Figure 7 showcases average differences in A-prime scores in the Induction condition compared to the Baseline condition across different age groups. Data presented in Figure 7 suggest that whereas Induction in 5- and 7-year-olds did not result in any appreciable change in memory accuracy, it resulted in marked changes in accuracy of 11-year-olds and adults. This contention was supported by statistical analyses: one-way ANOVA with age as a factor revealed the significant main effect of age, $F (3, 73) = 15.49, p < .0001$. Post-hoc Tukey tests indicated that after performing induction, 11 year-olds and adults exhibited greater drop in accuracy (compared to the Baseline) than younger participants, all $p$s < .05.

Taken together, results of Experiment 1 replicate and further extend findings reported by Sloutsky and Fisher (2004a, 2004b): 11 year-olds and adults spontaneously performed category-based induction, whereas 5- and 7-year-olds spontaneously performed similarity-based induction. These findings suggest that the ability to perform category-
based induction develops gradually between 7 years of age and adulthood rather than being a default present early in development.

Figure 7. Change in the A-prime scores in Induction compared to the Baseline across different age groups in Experiment 2. Error bars represent standard errors of the mean.

The experiment, however, left several questions unanswered. First, it could be argued that even if younger participants do perform category-based induction, their greater memory accuracy could be driven by their elevated interest in pictures. This argument was undermined in previous research with 5 year-olds (Sloutsky & Fisher, 2004b), and the goal of Experiment 3 was to replicate these findings and extend them to 7 year-old
children. Second, although results suggest that younger and older participants perform
induction differently, cross-sectional data alone do not provide an insight into how and
why the transition from “child-like” to “adult-like” induction takes place. To address
these issues, we conducted Experiment 3.

The goals of Experiment 3 were (1) to provide a learning account of category-based
induction found in older participants in Experiment 2, and (2) to further examine the
possibility that results of Experiment 2 stemmed from younger participants having
greater interest in pictures than older participants. In Experiment 3, 5- and 7-year-olds
were trained to perform induction in a category-based manner. If, as a result of training,
children start performing category-based induction, their memory accuracy in the
Induction condition (but not in the Baseline condition) should decrease.

**EXPERIMENT 3**

**Method**

**Participants**

Participants were 46 5-year-olds (27 girls and 19 boys, $M_{\text{age}} = 5.30$ years, $SD = 0.15$
years) and 30 7-year-olds (22 girls and 8 boys, $M_{\text{age}} = 7.39$ years, $SD = 0.59$ years).

**Materials, Design and Procedure**

Materials were identical to those in Experiment 2. The experiment had a 2
(Condition: Induction and Baseline) by 2 (Age Group: 5-year-olds and 7-year-olds)
between-subjects design, with participants being randomly assigned to either the
Induction or the Baseline condition. The procedure of Experiment 2 was similar to that
of Experiment 2, with one important exception: prior to the experiment proper, 5- and 7-year-olds were presented with training, in which they were taught to perform category-based induction.

Training was focused on the following three pieces of information: (1) animals that have the same name belong to the same kind; (2) animals that belong to the same kind have similar insides; and (3) animals that have the same name have similar insides. Training materials included three boxes, with each box identified by a black outline of a lion, a rabbit, or a dog. The materials also included pictures of lions, rabbits, and dogs (none of these categories was used in the experiment proper, and a separate naming study revealed that these animal categories were familiar to 5 year-olds).

Children were first told that “animals that have the same name are the same kind of animal,” and these animals should be placed in the same box. Children were then presented with six categorization trials, in which their task was placing pictures in appropriate boxes face down. After each trial children were provided with corrective “yes/no” feedback.

This categorization training was followed by induction training. First, participants were reminded that animals that have the same name are the same kind of animal. They were then told that animals of the same kind have “the same stuff inside”, and given 6 induction trials, each accompanied by “yes/no” feedback. On each trial, children were shown a picture, and told that the animal had a particular biological property (i.e. “this dog has thick blood inside its body”). Children were then asked to generalize this property to other animals. After responding participants were provided with feedback
indicating that generalizations should be made to the animals of the same kind, and an explanation that animals of the same kind have the same name and same stuff inside.

At the conclusion of training participants were reminded that “animals that have the same name are the same kind of animal, and have the same stuff inside.” All participants completed training successfully (i.e., gave either 5 out of 6, or 4 correct answers in a row in the categorization and induction training tasks), and were then presented with either Induction or Baseline condition of the experiment proper (identical to those in Experiment 2).

Results and Discussion

Catch Trials Accuracy

Data from five 5 year-olds (two in the Induction and three in the Baseline condition) who did not reliably reject catch items were eliminated from further analysis. As in Experiment 1, the rest of participants were very accurate in rejecting catch items, averaging 96%, and 100% of correct responses across conditions for 5- and 7-year-olds respectively.

Induction Accuracy

Similar to previous findings, participants were highly accurate in the Induction task. The average rate of correct generalizations was 91% and 94% for 5- and 7-year-olds respectively, above chance, both one-sample ts > 13.9, ps < .0001.
Recognition Accuracy

Proportions of hits, false alarms on critical lures, and A-prime scores by age and condition are presented in Table 2. Data in the Table indicate that the pattern of recognition accuracy changed dramatically compared to that observed in Experiment 2. Unlike in Experiment 2, in the Induction condition, memory sensitivity of participants in both age groups, as indicated by the A-primes scores, dropped to chance, both one-sample \( t < 1.7, p_s > .1 \). Furthermore, after training, memory accuracy of 5- and 7-year-olds (A-prime scores of .58 and .57, respectively) dropped to the level of adults in Experiment 1 (A-prime score = .54).

<table>
<thead>
<tr>
<th></th>
<th>5 year-olds</th>
<th>7 year-olds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hits</td>
<td>FA</td>
</tr>
<tr>
<td>Training + Baseline</td>
<td>.70</td>
<td>.39</td>
</tr>
<tr>
<td>Training + Induction</td>
<td>.82</td>
<td>.66</td>
</tr>
</tbody>
</table>

Note: * -- above chance recognition accuracy, \( p < .005 \)

Table 2. Mean proportions of Hits, False Alarms on critical lures (FA), and A-prime scores by age and condition in Experiment 3.
At the same time, participants exhibited above chance memory accuracy in the Baseline condition, both one-sample $t$s > 3.4, $ps < .005$. In fact, there was no drop in memory accuracy for either age group as a result of training in the Baseline condition (A-prime scores in the Training + Baseline in Experiment 3 were .69 and .78, for 5- and 7-year-olds respectively, whereas A-prime scores in the Baseline condition of Experiment 2 were .66 and .72, for 5- and 7-year-olds respectively). Therefore, it is unlikely that memory decrease in the Training + Induction condition stemmed from extraneous factors.

Results of Experiment 3 indicate that training to perform category-based induction significantly attenuated memory accuracy of both 5- and 7-year-olds in the Induction condition, bringing their recognition performance to chance and making it comparable to the performance of older participants in Experiment 2. Results also demonstrate that training per se did not have a general adverse effect on participants’ accuracy, because training did not result in reduced memory accuracy in the Baseline condition. Thus, Experiment 3 replicated earlier findings demonstrating that 5 year-olds can learn to perform category-based induction, and further extended these findings to 7-year-olds.

However, it could be argued that children did not learn to perform induction in a category-based manner during training; rather training merely activated their preexisting ability to perform category-based induction. If this is the case, then both 5- and 7-year-olds should be able to retain effects of training. At the same time, finding that even after short delays children revert to similarity-based induction would undermine this “preexisting knowledge” possibility. Experiment 4 was designed to address this issue by investigating the ability of 5- and 7-year-olds to retain the effect of training over a relatively short delay.
EXPERIMENT 4

Method

Participants

Participants were 29 5 year-olds (17 girls and 12 boys, $M$ age = 5.05 years, $SD$ = 0.35 years) and 20 7 year-olds (5 girls and 15 boys, $M$ age = 7.62 years, $SD$ = 0.45 years).

Materials, Design, and Procedure

Materials and procedure were identical to those of Experiment 3 with one important difference: there was a delay between training to perform category-based induction and the Induction-then-Recognition experiment. The delay was on average 14.6 days ($SD$ = 1.5 days, range 14 – 18 days). Participants were presented with the Induction condition only, with both the training procedure and the Induction condition being identical to those in Experiment 3.

Results and Discussion

Catch Trials Accuracy

Eight 5-year-olds and one 7-year-old did not reliably reject catch items, and their data were eliminated from further analysis. As in the previous experiments, the majority of participants were highly accurate in rejecting catch items, averaging over 97% of correct rejections.
**Induction Accuracy**

During the Study Phase participants were also highly accurate in making inductive generalizations. The average rate of correct inductions was 84% and 86% for 5- and 7-year-olds respectively, above chance, both one-sample $t > 7.6, ps < .0001$.

**Recognition Accuracy**

Proportions of hits, false alarms on critical lures, and A-prime scores across the two age groups are presented in Table 3. As can be seen in the table, in contrast to Experiment 3, memory accuracy differed across the age groups: whereas 7-year-olds retained high level of false alarms and hence low accuracy (A-prime scores not different from chance, one-sample $t (18) = 1.9, p > .07$), 5-yeas-olds recovered their high pre-training accuracy (A-prime scores reliably above chance, one-sample $t (16) = 4.9, p < .0001$).

<table>
<thead>
<tr>
<th>Training - Delay - Induction Condition</th>
<th>Age Group</th>
<th>Hits</th>
<th>FA</th>
<th>A-prime Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5 year-olds</td>
<td>.83</td>
<td>.61</td>
<td>.68*</td>
</tr>
<tr>
<td></td>
<td>7 year-olds</td>
<td>.91</td>
<td>.76</td>
<td>.58</td>
</tr>
</tbody>
</table>

Note: * -- above chance recognition accuracy, $p < .05$

Table 3. Mean proportions of Hits, False Alarms on critical lures (FA), and A-prime scores by age in Experiment 4.
These results indicate that the retention of the learned ability to perform category-based induction was a function of age: while 5-year-olds were unable to retain the trained ability to perform category-based induction even over a short two-week delay, 7-year-olds were able to do so. Therefore, it seems unlikely that training merely activated the pre-existing ability of 5-year-olds to perform category-based induction: given how fragile this ability is, it seems more likely that young children learned de novo how to perform category-based induction.

A-prime scores of 5- and 7-year-olds in the Induction condition across Experiments 2-4 are presented in Figure 8. Data in the Figure point to an interesting developmental pattern: (1) while neither 5- nor 7-year-olds perform category-based induction spontaneously participants in both age groups can be successfully trained to perform category-based induction, and (2) only 7-year-olds, but not 5-year-olds, are able to retain the effects of training over short delays.

Experiments 2-4 demonstrated that young children spontaneously perform similarity-based induction and that category-based induction develops gradually between 7-years of age and adulthood. However, in these experiments similarity-based induction was inferred from patterns of correct recognition memory. The goal of Experiment 5 was to present converging evidence that similarity-based induction results in accurate recognition memory. To achieve this goal, we forced adults to perform similarity-based induction by presented them with artificial (and thus completely novel) categories of animal-like creatures.
EXPERIMENT 5

Method

Participants

Participants were 48 introductory psychology students (21 men and 27 women, $M$ age $= 22.51$ years, $SD = 6.70$ years).

Material, Design, and Procedure

Design and procedure were identical to those of the Induction and Baseline conditions of Experiments 2-4, however materials consisted of a set of artificial animal-like stimuli created by combining seven features: body, upper wings, lower wings, buttons, antennas, tails, and spirals. Each feature varied on several dimensions (i.e., shape, size, color,
and/or number). A total of four categories were created: Category 1, Category 2, Category 3, and Catch Items. Stimuli were created such that Category 1 creatures all had round bodies, long upper wings, and lower wings, Category 2 creatures had round bodies, short upper wings, and lower wings, Category 3 creatures had long bodies, long upper wings, and no lower wings, and the rest of the features varied. Catch Items were created to be perceptually different from the rest of the creatures: they were blue colored and had spirals attached to the upper wings (neither of these features was used for the other animal categories). Examples of stimuli used in Experiment 5 are presented in Figure 9. Note that there was greater within-category similarity than similarity across categories, which was confirmed empirically in a separate experiment with 11 undergraduate students (none of whom participated in the experiment proper). These participants were not informed about the structure of these artificial categories, and yet they accurately grouped members of each category together (the overall accuracy was 94%).

In the experiment proper, participants were tested individually in a laboratory on campus, and all instructions were presented to them on a computer screen. During the Induction phase, corrective “Yes/No” feedback was provided to indicate that only Category 1 creatures, but not members of other categories had beta-cells inside.

Results and Discussion

Catch Trials Accuracy

Eight participants did not reliably reject catch items during the recognition test, and their data were excluded from further analyses. The rest of participants were accurate in both rejecting catch items (95% correct)
Figure 9. Examples of stimuli used in Experiment 5.
**Induction Accuracy**

During the Study Phase participants were successful in restricting their generalizations only to the members of Category 1. On the average, participants provided 74% of correct responses, above chance, one-sample $t(16) > 5.2$, $p < .0001$.

**Recognition Accuracy**

As predicted, participants exhibited accurate recognition memory in both the Baseline (Hits = .74, FA = .50, Accuracy = .24, and A-prime = .64) and the Induction conditions (Hits = .74, FA = .57, Accuracy = .17, and A-prime = .64), with both A-prime measures being above chance, both one sample $t$s > 2.99, $p$s < .01. Furthermore, unlike Experiment 2, accuracy scores were not statistically different in the Induction and Baseline conditions, $t(38) = .84$, $p > .40$. The fact that participants exhibited above-chance accuracy when performing induction with completely new categories is remarkable: participants were more accurate with these novel stimuli than they were with pictures of highly familiar categories used in Experiment 2, in which A-prime scores did not differ from chance.

Recall that in this condition adult participants were forced to perform similarity-based induction, and as a result they exhibited the same pattern of high recognition accuracy as young children in Experiments 2 and 4. These findings present direct evidence that similarity-based induction results in higher memory accuracy.
SUMMARY

Results of the four reported experiments point to several important regularities. First, in Experiment 2 recognition accuracy decreased with age in the Induction condition, but not in the Baseline or Blocked Categorization conditions. In Experiment 3, 5- and 7-year-olds were successfully trained to perform category-based induction, and training attenuated their recognition memory accuracy in the Induction but not in the Baseline condition. These findings replicated and further extended earlier findings of Sloutsky and Fisher (2004a, 2004b) suggesting that (a) induction attenuated memory accuracy in older but not in younger participants, and (b) when younger participants were trained to perform category-based induction, their memory accuracy dropped to the level of adults. In addition, Experiment 2 yielded a novel finding, revealing a clear developmental trend: after performing induction, 5- and 7-year-olds were more likely to exhibit accurate recognition than 11-year-olds and adults. These findings are remarkable because under regular conditions memory is known to increase rather than decrease from early childhood to young adulthood (see Schneider & Bjorklund, 1998, for a review). These results (a) support the hypothesis that early induction is similarity-based and (b) point to a gradual development of category-based induction from early childhood to adulthood.

In Experiment 4, when a time delay was introduced between induction training and the experiment proper, recognition memory accuracy of 5-year-olds increased to the pretraining level, whereas accuracy of 7-year-olds remained low. Finally, Experiment 5 demonstrated that when adults are forced to perform similarity-based induction, they exhibit patterns of accurate recognition in both the Baseline and Induction conditions.
Taken together, reported findings suggest that category-based induction is not a default, but is rather a product of development and learning. These results elucidate the development of induction, and they have theoretical implications for the study of categorization and induction, as well as broader implications for the study of memory and its development. These implications will be discussed in Chapter 7.

However, in all experiments reported in Chapter 4 presented pictures were never labeled with category labels, and it can be argued that even the youngest children would have performed induction in a category-based manner, had the pictures been labeled. In this case, findings reported in Chapter 4 reflect properties of the testing procedure used rather than the mechanism underlying inductive generalization early in development. The issue of the effects of category labels on induction in young children will be addresses in Chapter 5.
CHAPTER 5

EFFECT OF CATEGORY LABELS ON INDUCTION AND RECOGNITION

Linguistic labels have been demonstrated to promote inductive generalizations even early in development, however, the mechanism by which labels contribute to induction remains unknown. According to the proponents of the knowledge-based approach, when presented with entities that share a name (i.e., both are called *Cats*), people, including young children, first infer (by the linguistic assumption) that the entities belong to the same category. Then (by the category assumption) they infer that things that belong to the same category share important properties, and this induction has been referred to as “category-based induction” (e.g., Gelman, 2003). In other words, even early in development people realize that labels denote categories; therefore, it is the conceptual meaning behind labels that influences induction.

Unlike the knowledge-based approach, SINC assumes that for young children labels are features of objects contributing to the overall similarity of compared entities. Furthermore, there are reasons to believe that early in development effects of linguistic labels might stem from auditory input overshadowing, (i.e., attenuating processing of) the corresponding visual input (Napolitano & Sloutsky, 2004; Robinson & Sloutsky, 2004a; Sloutsky & Napolitano, 2003). In particular, when visual input is accompanied by
auditory input, young children are less likely to encode visual input compared to encoding of visual input presented in isolation (a single modality baseline). Therefore, when two differently looking entities are referred to by the same label, participants are less likely to encode differences between these entities than when these entities are presented in isolation (Sloutsky, Robinson, & Timbrook, 2005). In contrast, the knowledge-based approach argues that even early in development labels affect induction because induction is category-based with labels denoting categories.

Therefore, overall predictions of each position are quite straightforward. According to the knowledge-based position, providing category labels should promote category-based induction in children. At the same time, induction that is based on categorization is a variant of deep semantic processing, and thus should lead to an increase in the level of false alarms, and therefore to an overall decreased recognition accuracy.

On the other hand, SINC predicts that early induction is similarity-based (with labels contributing to the overall similarity), and thus requires perceptual processing. Therefore, providing category labels during induction should not promote category-based induction, and therefore should not lead to an elevated level of false alarms. At the same time, labeling information may lead to auditory overshadowing, and thus may disrupt children’s encoding of visual information. Therefore, SINC predicts that labels may decrease recognition accuracy, however the pattern of responses should differ dramatically from that exhibited by participants performing category-based induction. Recall that category-based induction is a variant of deep semantic processing, and thus results in the elevated levels of both hits and false alarms. If labels facilitate deep semantic processing, then participants performing category-based induction by relying on
category labels should exhibit a pattern of high hits and high false alarms. However, according to SINC, labels should not facilitate deep semantic processing in young children, and therefore providing labels should not lead to an elevated level of false alarms. At the same time labels may disrupt encoding of perceptual information due to auditory overshadowing, and this disrupted encoding could result in a decreased levels of hits. However, it is also possible that labels will not have any appreciable effect on recognition memory. This possibility, while consistent with SINC, is inconsistent with the knowledge-based approach. However, after training to perform induction by relying on category labels, young children should exhibit a pattern of elevated hits and false alarms – a pattern indicative of deep semantic processing. Thus, according to both positions labels may have negative effects on young children’s recognition accuracy. However, the negative effects exerted by interrupted perceptual processing should differ from the negative effects exerted by semantic processing: while the latter should result in an elevated level of false alarms, the former should result in a decreased level of hits.

The goal of the experiments presented below is to test these hypotheses, thus elucidating mechanisms underlying effects of labels on induction, and distinguishing between the two theoretical alternatives.

EXPERIMENT 6

In Experiment 6, 5 year-old children and adults were tested using the stimuli and the same paradigm used in our previous research (Fisher & Sloutsky, 2005; Sloutsky & Fisher, 2004a, 2004b) and in Experiments 1-5 (Chapters 3-4). There was, however, one critical difference in the procedure: in contrast to the previous research, in the Study
Phase of Experiment 1 items were accompanied by appropriate basic-level category labels (i.e., all cats were referred with a label “cat”). If category labels promote spontaneous category-based induction in young children, then this manipulation should result in an increased level of category-based induction, thus increasing the level of false alarms during the recognition test. This, however, should not happen if labels are features of objects contributing to the overall similarity. Also note that performance of adult participants was expected to be unaffected by labels since adults have been shown to be performing category-based induction with familiar entities even in the absence of labels (Fisher & Sloutsky, 2005; Sloutsky & Fisher, 2004a, 2004b).

Method

Participants

Participants were 83 preschool children recruited from several daycare centers located in middle class suburbs of Columbus, Ohio (44 girls and 39 boys; \( M = 5.17 \) years; \( SD = .35 \)), and 100 college undergraduates (70 women and 30 men; \( M = 19.18 \) years, \( SD = 1.36 \) years) from the Grand Valley State University and the Ohio State University, participating in the experiment in partial fulfillment of a course requirement.

Experimental Materials, Methods, and Procedures

Experiment 6 used the Induction-then-Recognition and Baseline-then Recognition paradigm described in Chapter 1. In addition to pictorial materials described in the Basic Paradigm (Chapter 1), Materials also included a set of category labels – *Cat, Bear,* and *Bird.* There were two between-subjects Labeling conditions (Category Label and No
Label) nested within the Study Phase task conditions (Baseline and Induction). In the No Label condition, Study Phase tasks were identical to the Baseline and Induction tasks described in the Basic Paradigm (Chapter 1). In Study Phase of the Category Label condition each picture was accompanied by a basic level category label (i.e., *Cat*, *Bird*, or *Bear*). The Recognition phase was the same for all labeling and task conditions, and identical to the one described in Chapter 1.

Results and Discussion

*Catch Trials Accuracy*

Data from participants who failed to correctly reject at least 5 out of 7 catch items in the Recognition Phase were excluded from further analysis. There were 9 children and 10 adults excluded: five 5 year-olds in the Baseline condition (3 in the No Label and 2 in the Category Label condition), four 5 year-olds in the Induction condition (1 in the No Label, and 3 in the Category Label condition), eight adults in the Baseline (7 in the No Label and 1 in the Category Label condition), and two adults in the Induction (both in the No Label condition). The rest of the participants were highly accurate in rejecting catch items, averaging across all experimental conditions over 96% of correct responses for 5 year-olds and over 97% of correct responses for adults.

*Induction Accuracy*

Similar to the previous reports (Fisher & Sloutsky, 2005; Sloutsky & Fisher, 2004a, 2004b), when no labels were presented, all participants performed accurate inductive generalizations averaging 75% and 88% of correct responses for 5 year-olds and adults.
respectively, both above chance, one-sample $t > 2.3, p < .0001$. Introduction of category labels during the Study Phase had no effect on the induction accuracy of adults (93% of correct generalizations in the Category Label condition), independent-samples $t(44) < 1.5, p > .6$.

However, in 5 year-olds the rate of correct generalizations was somewhat lower in the Category Labels condition compared to the No Label condition (67% compared 75% of correct inductions respectively). Although the difference did not reach significance, it was clearly in the direction opposite from that predicted by the knowledge-based approach, which argues that category labels should have promoted category-based induction. However, most important finding stem from the recognition accuracy data.

**Recognition Accuracy**

Memory sensitivity A-prime scores, computed for each participant and averaged across participants in each between-subject condition, are presented in Figure 10. Data in the Figure indicate that introduction of category labels had no effect on the patterns of recognition accuracy of adult participants: similar to the previous reports there was a marked drop in recognition accuracy in the Induction condition compared to the Baseline, regardless of the labeling condition: in the Baseline condition average A-prime scores were .81 and .76 in the No Label and Category Label conditions respectively, whereas in the Induction condition average A-prime scores were .61 and .59 in the No Labels and Category Labels conditions respectively. This decrease in recognition accuracy in the Induction condition compared to the Baseline was significant, both independent-samples $t > 3.1, p < .01$).
Figure 10. Memory Sensitivity A-prime scores for adults (panel 10-A) and children (panel 10-B) in Experiment 1. Error bars represent the standard errors of the mean. The dashed line represents the point of no sensitivity.
At the same time, introduction of category labels affected recognition accuracy of 5-year-olds: when no labels were presented during the Study Phase, participants demonstrated accurate recognition memory in both, Induction and Baseline conditions (average A-prime scores were .63 and .66 respectively, both above chance, $t_s > 2.5$, $p_s < .05$). However, when category labels were introduced, children’s memory accuracy in the Induction condition decreased to chance (average A-prime = 56, $t (22) < 1.1$, $p > .3$), whereas participants remained accurate in the Baseline condition (average A-prime = .67, one-sample $t (15) > 4$, $p < .001$).

Recall, both SINC and the knowledge-based approach predict a possible decrease in recognition accuracy in the Category Label condition, however these theoretical positions predict that attenuated recognition should stem from different patterns of hits and false alarms. Specifically, if labels facilitate category-based induction (as predicted by the naïve theory), the decrease in recognition accuracy should occur due to an increased level of false alarms, stemming from semantic processing. However, if labels overshadow visual information, thus disrupting encoding of visual information (as predicted by SINC), attenuated recognition accuracy should stem from a decreased level of hits. At the same time, labels might have no effect on children’s performance, and this outcome would be consistent with SINC, but not with the naïve theory approach. To distinguish between these possibilities, analyses of hits and false alarms were conducted. Proportions of hits and false alarms were submitted to a mixed ANOVA with response type (hits vs. false alarms) as a repeated measure and labeling condition (Category Label vs. No Label), age (children vs. adults), and task (Induction vs. Baseline) as factors. The analysis revealed a significant response type by task by age interaction, $F (1, 157) > 4$, $p < .05$. 

57
The patterns of hits and false alarms in Experiment 6 are presented in Figures 11-12. Data in Figure 11-A indicate that high accuracy of adults in the Baseline condition was generated by a pattern of high hits and low false alarms in both Category Label and No Label conditions (in each labeling condition the average level of hits was .82 and .85 respectively, and the level of false alarms was .43 and .36 respectively). Furthermore, in both labeling conditions the level of hits was above chance (one-sample \( t \) > 12.3, \( p < .0001 \)), whereas the level of false alarms was either at or below chance (\( t \) (24) < 1.2, \( p > .26 \) in the Category Label condition; \( t \) (24) > 2.5, \( p < .05 \) in the No Label Condition).

Data presented in Figure 11-B suggest that low recognition accuracy of adult participants in the Induction condition was generated by a pattern of high hits and high false alarms – a pattern indicative of conceptual processing (in both labeling conditions the average level of hits was .86 and the level of false alarms was .70). In contrast to the Baseline, in the Induction condition both, the level of hits and the level of false alarms were above chance, regardless of the labeling condition (one-sample \( t \) > 2.9, \( p < .01 \)).

Data presented in Figure 12-A demonstrate that in 5 year-olds, similar to adult participants, high recognition accuracy in the Baseline condition was generated by a pattern of high hits (.74 and .79 in the No Label and Category Label conditions respectively, both above chance, one-sample \( t \) > 5.7, \( p < .0001 \)) and relatively low level of false alarms (.58 and .46 in the No Label and Category Label conditions respectively, both not different from chance, one-sample \( t \) < 1, \( p > .3 \)). Furthermore, the difference in the level of hits and false alarms in both labeling conditions was statistically significant, both paired-samples \( t \) > 3.6, \( p < .005 \).
Figure 11. Proportion of Hits and False Alarms (FA) in adult participants in the Baseline (panel 11-A) and Induction (panel 11-B) conditions Experiment 6. The error bars represent the standard errors of the mean. The dashed line represents chance level.
Figure 12. Proportion of Hits and False Alarms (FA) in 5 year-olds in the Baseline (panel 12-A) and Induction (panel 12-B) conditions in Experiment 6. The error bars represent the standard errors of the mean. The dashed line represents chance level.
However, patterns of hits and false alarms in the Induction condition were markedly different in 5 year-olds and adults. Data in Figure 12-B indicate that in the No Label condition, the level of hits was significantly above chance (average level of hits = .68, one-sample $t(20) > 3.8, p < .001$) and statistically different from the level of false alarms (average level of false alarms = .46), paired-samples $t(20) > 3.6, p < .005$); however in the Category Label condition the rates of hits (.61) and false alarms (.53) were indistinguishable (paired-samples $t(22) = 1, p > .3$) and not different from chance (both $ps > .06$).

Thus, results of Experiment 6 indicate that neither induction nor recognition accuracy of adult participants was affected by category labels: in both labeling conditions adults demonstrated equally high induction accuracy and a pattern of hits and false alarms indicative of conceptual processing. At the same time, providing category labels did not promote category-based induction in 5 year-old children: the rate of correct inductions in the Category Labels condition did not increase compared to the No Label condition, and analysis of the hits and false alarms patterns pointed to disrupted perceptual rather than to conceptual processing.

Results of Experiment 6 do not support predictions of the knowledge-based approach that category labels promote category-based induction in young children. At the same time, these results provide support to the hypothesis that labels attenuate processing of the corresponding visual input. This hypothesis was tested further in Experiment 6-A, in which 5 year-olds were provided with labeling information which could not possibly facilitate category-based induction: if under these conditions children’s performance is comparable to that in Experiment 6, then this would provide further support to the
hypothesis that early in development labels are features of objects and may attenuate processing of the corresponding visual input.

EXPERIMENT 6A

Method

Participants

Participants were 20 preschool children recruited from several daycare centers located in the suburbs of Columbus, Ohio (12 girls and 8 boys; $M = 5.07$ years; $SD = .24$).

Materials, Design, and Procedure

Design and procedure of Experiment 6A were similar to the Induction condition of Experiment 6 with one important difference: during the Study Phase each picture was labeled with an individual nonsense label (instead of a basic level category label), so that there were a total of 30 nonsense labels used in Experiment 6A (i.e., a fika, a guma, a dabi, etc.). Because each item was referred to with a different label, these labels could not possibly promote category-based induction. This manipulation was based on the following reasoning: if even early in development labels are category markers, then effect of familiar category labels on induction and recognition memory should be different from that of individual nonsense labels. However, if early in development labels exert general auditory effects on visual processing, then performance with category labels should be similar to performance with individual nonsense labels.
Results and Discussion

Catch Trials Accuracy

Data from participants who failed to correctly reject at least 5 out of 7 catch items in the Recognition Phase were excluded from further analysis. Data from four participants were excluded based on this criterion. The rest of participants were very accurate in rejecting catch items, averaging over 98% of correct rejections.

Induction Accuracy

In Experiment 6A Participants averaged 66% of correct generalizations when pictures were accompanied by individual nonsense labels, which is equivalent to children’s performance with category labels in Experiment 6 (67% of correct generalizations), $t(37) < 1$. This finding indicates that in 5 year-old children, category labels and individual nonsense labels exert similar influence on induction.

Recognition Accuracy

In Experiment 6A average A-prime scores were equal to .62, above chance, one-sample $t(15) > 2.7, p < .05$. Nonetheless, the patterns of hits and false alarms of 5-year-olds in Experiment 6-A, were similar to those observed in Experiment 6. Specifically, participants demonstrated a pattern of low hits (.56, not different from chance, one-sample $t(15) < 1, p > .3$) and low false alarms (.41, not different from chance, one-sample $t(15) < 1.3, p > .2$).

Taken together, results of Experiments 6 and 6A indicate that early in development category labels do not make a contribution to induction that is above and beyond the
contribution of individual nonsense labels. This finding supports predictions of SINC, and further undermines the argument that early in development effects of labels on induction stem from their conceptual rather than general auditory properties. However, it could be argued that the pattern of low hits and low false alarms observed in the Induction-with-Labels conditions of Experiment 6 and 6A stems from extraneous factors, such as elevated task demands. In this case, it could be argued that due to the elevated task demands, participants performed at their functional ceiling with respect to their level of hits. The goal of Experiment 7 was to eliminate this possibility. In Experiment 7 young children were trained to perform category-based induction relying on common category labels. Previous findings (Sloutsky & Fisher, 2004a, 2004b) indicate that such training results in a pattern of recognition accuracy indicative of conceptual encoding (i.e., high hits and high false alarms). Therefore, if results of Experiments 6 and 6A stemmed from interrupted similarity-based induction, training to perform category-based induction should result in low accuracy stemming from the pattern of high hits and high false alarms.

EXPERIMENT 7

In Experiment 7 participants were first trained to perform category-based induction relying on common category labels. After training participants were presented with the experiment proper, which was identical to the Induction and Baseline tasks of the Category Labels condition of Experiment 6. It was expected that after training participants should perform category-based induction relying on category labels; no effects of training on the performance in the Baseline condition were expected.
Therefore, (1) induction accuracy was predicted to increase (compared to that in Experiments 6 and 6-A), (2) recognition accuracy in the Induction condition was expected to be low due to an increase in the level of false alarms (rather than to a decrease in the level of hits as in Experiments 6), and (3) recognition accuracy in the Baseline condition was expected to be high with the pattern of high hits and relatively low false alarms.

**Method**

*Participants*

Participants were 41 preschool children recruited from several daycare centers located in middle class suburbs of Columbus, Ohio (19 girls and 22 boys; \( M = 5.07 \) years; \( SD = .36 \)).

*Materials, Design, and Procedure*

Materials, design, and procedure of Experiment 7 were similar to those of the Category Labels condition of Experiment 6, with one important difference: prior to the experiment proper, participants were given training in category-based induction. The training procedure was identical to the one used in Experiment 3. At the completion of training participants were randomly assigned to an Induction or a Baseline condition, in which presentation of animals during the Study Phase was accompanied by appropriate category labels.
Results and Discussion

*Catch Trials Accuracy*

Data from participants who failed to correctly reject at least 5 out of 7 catch items in the Recognition Phase were excluded from further analysis. Based on this criterion, data from four participants in the Induction condition and one participant in the Baseline condition were excluded. The rest of the participants were highly accurate in rejecting catch items, averaging over 96% of correct rejections.

*Induction Accuracy*

As predicted, after training participants demonstrated high induction accuracy: the rate of correct generalizations was 88%, above chance, one-sample $t (16) > 10.6$, $p < .0001$. Furthermore, participants’ induction accuracy in Experiment 7 (i.e., 88%) exceeded that of Experiment 6 (i.e., 67%), independent-samples $t (38) > 3.3$, $p < .001$.

*Recognition Accuracy*

The patterns of recognition accuracy in Experiment 7 are summarized in Figures 13-14. Data in Figure 13 indicate that similar to Experiment 6, recognition accuracy in the Induction condition of Experiment 7 was quite low, with an average A-prime score of .52 (not different from chance, $p > .7$). However, the pattern of hits and false alarms changed dramatically compared to Experiment 6: as predicted, after training in category-based induction, participants exhibited a pattern of high hits (.77) and high false alarms (.71), both above chance, one-sample $ts > 2.4$, $p < .05$ (Figure 6b). Furthermore, the rate of hits was statistically equivalent to the level of false alarms, paired-samples $t (16) < 1$, $p > .4$.  

66
Overall, Experiment 7 demonstrated that after training to perform category-based induction by relying on common labels, children exhibited evidence of category-based induction, which resulted in a marked increase in the level of hits and false alarms. Note that the pattern of results observed in the Induction condition of Experiment 7 (high hits/high false alarms) is indicative of conceptual processing, and is different from (1) the pattern observed in the Baseline conditions of Experiment 6 (high hits/low false alarms) – a pattern indicative of perceptual processing, and (2) from the pattern observed in the Induction condition of Experiment 6 (low hits/low false alarms) – a pattern that suggests interrupted perceptual processing.

![Graph showing memory sensitivity A-prime scores](image)

**Figure 13.** Memory Sensitivity A-prime scores for 5 year-olds in the Category Labels condition in Experiments 6 and 7. The error bars represent the standard errors of the mean. The dashed line represents the point of no sensitivity.
Figure 14. Proportion of Hits and False Alarms (FA) in 5 year-olds in the Baseline condition (panel A) and Induction condition (panel B) in Experiments 6 and 7. The error bars represent the standard errors of the mean. The dashed line represents chance level.
SUMMARY

Several important findings stem from the experiments reported in Chapter 5. First, as demonstrated in Experiment 6, introduction of category labels does not automatically promote category-based induction: category labels seem to disrupt perceptual processing, rather than promote conceptual processing in the Induction tasks – as evidenced by slightly decreased induction accuracy and dramatically decreased level of correct recognition. At the same time, no evidence for auditory overshadowing was found in the Baseline tasks. One potential explanation for this finding is a possibility that overshadowing effects are, at least in part, moderated by the task complexity that interacts with stimulus familiarity and age of participants (both of the latter variables have been implicated in overshadowing effects, see Napolitano & Sloutsky, 2004 and Robinson & Sloutsky, 2004a for a discussion). However, this possibility requires further investigation.

Experiment 6A demonstrated that contribution of category labels to induction early in development does not go above and beyond the contribution of individual nonsense labels. This finding provides further support to the hypothesis that auditory information tends to attenuate processing of the visual information in young children, and that early in development effect of labels on induction stem from their general auditory rather than conceptual properties. Finally, Experiment 7 demonstrated that training young children to rely on common category labels when performing induction, increases their induction accuracy, and leads to a pattern of recognition accuracy (high hits/high false alarms) indicative of conceptual processing. At the same time, recognition accuracy in the Baseline condition was unaffected by training.
Taken together, results of Experiments 6-7 suggest that even when provided with category labels, young children perform similarity-based induction, and contribution of category labels to early induction is similar to the contribution of individual nonsense labels. At the same time, young children can be trained to perform category-based induction by relying on common labels. These results support predictions of the SINC model, while challenging the idea of the knowledge-based approach that early induction is driven by the category and linguistic assumptions. These results also have implications for the study of effects of language on cognition, and these implications will be discussed in Chapter 7.

Previous research (Fisher & Sloutsky, 2005; Sloutsky & Fisher, 2004a, 2004b) and results of Experiment 3 (Chapter 3) and 7 (Chapter 5) indicate that young children can be trained to perform induction by relying on common category labels. However, in all these experiments, training included an explanation given to a participant. It remains unclear whether this component is necessary for successful training. One possibility is that purely associative training would be sufficient to promote label-based generalizations in young children. This possibility is predicted by SINC. Alternatively, it is possible that explanation-based training used in previous research is required to obtain this pattern of results. This possibility would be predicted by the knowledge-based approach). This issue was explored in Chapter 6.
CHAPTER 6

EFFECTS OF ASSOCIATIVE TRAINING ON INDUCTION AND RECOGNITION ACCURACY

In Chapters 3-4 (Experiments 3, 4, and 7) it has been demonstrated that young children can be trained to perform category-based induction, and to rely on shared labels to perform induction. The training procedure used in Experiments 3, 4, and 7, and in other previous research (Sloutsky & Fisher, 2004a, 2004b) consisted of two components – a conceptual component (i.e., providing an explanation about why labels are important for categorization and why labels can be relied upon in induction) and an attentional component (i.e., attracting participants’ attention to the labeling information). However, it is unclear what proportion of success of the training procedure used in previous research is attributable to conceptual and attentional components.

The goal of the experiments presented in Chapter 5 was to investigate the possibility that increasing attentional weights of labels without providing conceptual explanations is sufficient to influence induction in young children. In Experiment 8 participants were given training that was aimed at increasing attentional weights of labels in induction and was devoid of any conceptual components. After training, participants were presented with the Induction and Baseline tasks identical to those in the Category Label condition of Experiment 6 (Chapter 4), in which pictures of animals were accompanied by category
labels during the Study phase. In Experiment 9 participants were presented with the same attentional-only training as in Experiment 8, however following training children were presented with the Induction and Baseline tasks identical to those of the No Label condition of Experiment 6, in which no labels accompanied presented pictures. If it was the attentional component of the original training procedure that was primarily responsible for the success of training of category-based induction, then purely attentional training should be successful in influencing induction in young children: as a result of attentional training participants should exhibit patterns of recognition accuracy observed in Experiments 3 and 7. Recall, that in Experiments 3 and 7, following training, participants exhibited high hits and high false alarms in the Induction but not in the Baseline condition. Alternatively, if it was the conceptual component of training that was primarily responsible for the effectiveness of the original training procedure, then purely attentional training should prove ineffective in both Experiment 8 and 9. At the same time, if conceptual components of training have any added benefits compared to associative components, then purely attentional training might differentially affect participants’ performance in Experiment 8 (in which training conditions closely match testing conditions: i.e., labeling information is provided during both training and testing) and Experiment 9 (in which training conditions do not closely match testing conditions: i.e., labeling information is provided during training only).
EXPERIMENT 8

Method

Participants

Participants were 51 preschoolers (22 girls and 29 boys; $M = 5.18$ years; $SD = .23$).

Materials

Experiment 8 consisted of two parts – Associative Training and an experiment proper (which used the Induction-then-Recognition paradigm). Materials for the experiment proper were identical to those used in Experiment 6 (Chapter 4). These materials consisted of 44 pictures of members of animal categories familiar to 5 year-olds (as established by a separate calibration study, see Sloutsky & Fisher, 2004b) – cats, bears, birds, and squirrels, and a set of basic level category labels (i.e., Cat, Bear, and Bird).

Materials for Attentional Training consisted of a set of photographs of 48 guinea pigs and hedgehogs (28 pictures per category), a set of 48 artificial two-syllable labels, and a set of 24 pseudo-biological properties. A separate calibration study with 27 five-year-olds established that animal categories used in training were unfamiliar to young children (less than 4% of participants could correctly identify presented animal pictures). Animal pictures used in training were organized into triads, consisting of a target and two test items, such that one of the test items was identical to the target, while the other test item looked different and was drawn from the opposite animal category (see Figure 14 for examples of training materials).
Procedure

At the beginning of training participants were told a story about a witch who had captured some animals and kept them caged in her castle. Children were told that by playing the game with the experimenter, they could help to free the animals. Participants were then presented with 24 triads of animal pictures (in a pseudo-randomized order), labeled with artificial labels, such that appearance similarity and labeling information were pitted against each other (i.e., the animal that was different from the Target had the same name as the Target, and the animal that was identical to the Target had a different name, see Figure 15 for examples). Participants were told that names of animals are very important in this game, and they should pay attention to animals’ names. Participants were familiarized with a pseudo-biological property of the Target, for instance they could be told that the Target animal “has thick blood inside”. Participants were then asked to generalize the property to one of the test items. Upon responding, participants were provided with feedback, indicating that the target property should be generalized only to the animals with matching names, but not appearances. If participants provided a label-based response, they received feedback in a form of a short movie, in which one animal from the witch’s castle was freed. If participants provided an appearance-based response, they received feedback in the form of a picture of caged animals. Notice, that participants were never given an explanation as to why labels are important for induction – participants were merely encouraged to increase their attention to matching labels, through positive feedback reinforcing label-based generalizations. At the completion of training participants were randomly assigned to one of the two between-subjects testing conditions: Induction or Baseline. In both testing conditions, during the Study Phase
pictures were accompanied by the basic level category labels. The Recognition task, administered immediately after the Study tasks, was identical to the one described in the Basic Paradigm (Chapter 1).

Figure 15. Examples of training stimuli used in Experiment 8. Notice, that all labeling and property information was presented auditorily.

Results and Discussion

Data from participants who provided less than 17 correct responses out of 24 trials (i.e., less than 70% of correct responses) during the Training Phase were excluded from further analysis. Based on this criterion data from eight participants were excluded. The
rest of participants were very accurate and averaged over 93% of label-based responses during the Training Phase.

Additionally, data from participants who failed to correctly reject at least 5 out of 7 catch items in the Recognition task of the Testing Phase were also excluded from further analysis. Based on this criterion data from nine more participants were excluded from the analysis – five in the Baseline and four in the Induction condition. The rest of participants were very accurate and averaged over 96% of correct rejections of catch items across the Baseline and Induction conditions.

**Induction Accuracy**

As predicted, in the Induction condition participants were also very accurate in making correct generalizations, averaging 79% of correct inductions, above chance, one-sample \( t (19) > 6.06, p < .0001 \). However, the most important findings stem from the recognition accuracy data.

**Recognition Accuracy**

The pattern of recognition accuracy in Experiment 8, in which children were presented with purely attentional training, was similar to the pattern of recognition accuracy in Experiment 7, in which training included a conceptual explanation of why labels are important for induction. Specifically, recognition accuracy of participants in the Induction condition of Experiment 8 decreased to chance, average A-prime score = .51, not different from chance, one-sample \( t (19) < 1, p > .8 \). At the same time, similar to Experiment 7, in the Baseline condition participants remained accurate even after
training, average A-prime score = .62, above chance, one-sample $t(21) > 2.7, p < .05$ (see Figure 16).

![Graph showing memory sensitivity A-prime scores in the Induction and Baseline conditions in Experiments 7-8. Error bars represent the standard errors of the mean. Dashed line represents the point of no sensitivity.]

Figure 16. Memory sensitivity A-prime scores in the Induction and Baseline conditions in Experiments 7-8. Error bars represent the standard errors of the mean. Dashed line represents the point of no sensitivity.

In addition to the aggregated measures of recognition memory being similar in Experiments 7-8, the patterns of hits and false alarms in these experiments were also similar. Specifically, in the Baseline condition of Experiment 8 participants demonstrated a pattern of high hits and low false alarms (hits = .71, above chance, one-sample $t(21) > 3.7, p < .01$; false alarms = .56, not different from chance, one-sample $t(21) < 1, p > .3$). Furthermore, the level of hits in the Baseline condition of Experiment 8 was greater than
the level of false alarms, paired-samples $t\ (21) > 2.6, p < .05$. At the same time, in the Induction condition of Experiment 8 participants demonstrated a pattern of high hits and high false alarms (hits = .69, false alarms = .68, both above chance, both $ts\ (19) > 2.9, ps < .01$), with the level of hits being statistically equivalent to the level of false alarms, paired-samples $t\ (19) < 1, p > .8$. The patterns of hits and false alarms in the Induction and Baseline conditions of Experiment 8 are presented in Figure 17.

![Figure 17. Proportions of Hits and False Alarms in the Induction and Baseline conditions of Experiment 8. Error bars represent the standard errors of the mean. Dashed line represents chance level.](image)

Notice, that the pattern of hits in false alarms in Experiment 8 is different from the pattern observed in Experiment 6, in which participants received no training, and similar to the pattern in Experiments 3 and 7, in which participants received a conceptual
explanation about the importance of labels. Therefore, results of Experiment 8 indicate that purely associative training is sufficient to influence the pattern of induction in 5 year-olds. However, in both the Training and the Testing phases of Experiment 8 participants were provided with labeling information, unlike Experiment 3 (Chapter 3), in which participants were provided with labeling information only during training. Removing labeling information from the experiment proper, should inform us whether effects of associative training influence induction only when testing conditions closely match training conditions (i.e., when labeling information is provided in both cases) or whether effects of associative training transfer to the testing conditions that are different from the training conditions (i.e., when labeling information is provided only during training but not testing). These possibilities were investigated in Experiment 9.

EXPERIMENT 9

Method

Participants

Participants were 50 preschoolers (24 girls and 26 boys; \( M = 5.19 \) years; \( SD = .21 \)).

Materials, Design, and Procedure

Materials, designs, and procedure of Experiment 9 were similar to those of Experiment 8 with one important difference: no labeling information was presented to participants during the Study Phase of the Induction and Baseline conditions that immediately followed associative training.
Results and Discussion

Data from participants who provided less than 17 correct responses out of 24 trials (i.e., less than 70% of correct responses) during the Training Phase were excluded from further analysis. Based on this criterion data from 10 participants were excluded. The rest of participants were very accurate and averaged over 92% of label-based responses during the Training Phase.

Additionally, data from participants who failed to correctly reject at least 5 out of 7 catch items in the Recognition task of the Testing Phase were also excluded from further analysis. Based on this criterion data from one participant in the Induction condition and four participants in the Baseline condition were excluded from further analysis. The rest of participants were very accurate and averaged over 97% of correct rejections across the Baseline and Induction conditions.

Induction Accuracy

Similar to Experiment 8, in the Induction condition of Experiment 9 participants were very accurate in making correct generalizations. The average rate of correct inductions was over 80%, above chance, one-sample $t (15) > 5.5, p < .0001$.

Recognition Accuracy

The patterns of recognition accuracy were different in Experiments 8 and 9: unlike Experiment 8, in which participants demonstrated greater recognition accuracy in the Baseline than in the Induction condition, in Experiment 9 participants exhibited equally high recognition accuracy in both testing conditions (average A-prime score in the
Induction condition = .69, average A-prime score in the Baseline condition = .66, both above chance, one-sample $t > 3$, $p < .01$, independent-samples $t (33) < 1$, $p > .4$).

A follow-up analysis of the hits and false alarms patterns indicated that in both Baseline and Induction conditions participants’ accuracy was generated by a pattern of high hits and low false alarms (see Figure 18). Specifically, in both testing conditions the level of hits was greater than chance (hits = .69 in the Induction condition, hits = .72 in the Baseline condition, both above chance, one-sample $t > 3$, $p < .01$), whereas the level of false alarms was not different from chance (false alarms = .46 in the Induction condition, false alarms = .51 in the Baseline condition, both not different from chance, $t < 1$, $p > .4$). Furthermore, in both Induction and Baseline conditions the level of hits was greater than the level of false alarms, both paired-samples $t > 2.7$, $p < .05$.

Overall, results of Experiments 9 indicate that brief associative training is not sufficient to influence weights of labels in 5 year-olds when testing conditions do not closely match training conditions (i.e., when labeling information is provided only during training but not testing). Taken together, results of Experiments 8 and 9 provide suggestive evidence that conceptual training may have added benefits in terms of broader generalization of training effects, compared to purely associative training. However, this issue requires further investigation.
SUMMARY

Experiments 8 and 9 investigated effects of purely associative training of the weights of labels in early induction. Results of these experiments indicate that brief associative training is sufficient to increase attentional weights of labels in 5 year-olds when training conditions closely match testing conditions (Experiment 8): patterns of recognition accuracy and patterns of hits and false alarms of trained children changed dramatically compared to those of untrained children (i.e., Experiment 2, Chapter 3). Specifically, following associative training children’s recognition accuracy decreased to the level of chance, and follow-up analyses indicated that this decrease in accuracy was generated by a marked increase in the level of false alarms. However, as Experiment 9 demonstrated,
when conditions of the experiment proper do not closely match training conditions, brief associative training was not sufficient to influence attentional weights of labels, as evidenced by children’s pattern of recognition accuracy after training. Recall that in Experiment 9, children exhibited high recognition accuracy, a pattern that was similar to that of untrained children.

These results contrast with the previous findings on the effects of training that includes a conceptual component, presented in published reports (Fisher & Sloutsky, 2005; Sloutsky & Fisher, 2004, 2004b) and in Experiment 3 (Chapter 3): even when training conditions did not closely match testing conditions, training that included a conceptual component was effective in increasing attentional weights of labels in 5 year-olds. Therefore, it is possible that there is an added benefit of training that includes conceptual components compared to purely associative training. However, a more direct investigation of this possibility is necessary, as well as investigation of other factors that can elucidate potential benefits of one kind of training over the other (i.e., sustainability of training over delays and minimal amount of training necessary to influence attentional weights of features). Directions for future research will be further discussed in the General Discussion section (Chapter 7).
CHAPTER 7

POSSIBLE MECHANISM OF CATEGORIZATION: EFFECTS OF DELAYED RECOGNITION ON MEMORY ACCURACY

Results of experiments presented in Chapters 2-5 demonstrated that unlike adults young children spontaneously engage in item-specific (or similarity-based) processing when making generalizations about members of familiar categories. However, item-specific information often decays rapidly, and as a result of this rapid forgetting only information that is common to all experienced exemplars is retained. In other words, rapid forgetting of perceptual information may lead to formation of generalized (or category-like) representations.

If true, then forgetting of item-specific information could potentially be one of the mechanisms of categorization. The possibility that forgetting of perceptual information underlies the mechanism of categorization and category learning has been suggested by findings of Nosofsky and Zaki (1998). Nosofsky and Zaki (1998) presented adult participants with study pictures, consisting of dot patterns generated from a prototype, and then tested participants in a classification or a recognition task, which was administered either immediately after the study phase or after a delay. Nosofsky and Zaki (1998) demonstrated that delay lead to forgetting of item-specific information, and
therefore to a decrease in participants’ recognition performance, compared to the no-delay condition. At the same time, forgetting of item-specific information had no appreciable effect on participants’ classification performance, since item-specific information was irrelevant to a categorization task. These findings provide support to the possibility that forgetting of item-specific information plays an important role in categorization and category learning.

Furthermore, it has been demonstrated that in 5- and 6-month-olds forgetting of perceptual information leaves infants with a generalized representation of previous experience. In particular, Borovksy and Rovee-Collier (1990) trained 6-month-old infants to move a mobile by kicking it, and later tested infants’ memory with the same or different mobile after delays ranging from 1 to 3 days. Results indicated that performance with a mobile different from the one used in training, is severely impaired after short delays, when context-specific memories are presumably intact, and is unaffected at longer delays, when surface characteristics of the mobile used in training are presumably forgotten.

More recently Dueker, Modi, and Needham (2003) demonstrated similar effects in 5-month-old infants using a preferential looking paradigm in an object segregation task. They pre-familiarized 5-month-olds with objects and 24 hours later tested infants in an object segregation task using these pre-familiarized objects. Dueker et. al. (2003) established that infants are able to use prior experience to aid their performance on the objects segregation task, and that test objects do not have to be identical to the pre-familiarized objects – presumably, infants forget perceptual details of the pre-familiarized
objects over the course of the delay, and succeed on the task even when it involves novel objects that are sufficiently similar to the pre-familiarization objects.

These findings suggest that generalized representations can be an emergent property of item-based processing. The goal of the experiment presented in Chapter 6 was to examine this possibility with 5 year-old children using Induction-then-Recognition paradigm.

In Experiment 10 participants were presented with the Induction and Baseline tasks described in the Basic Paradigm chapter, however there was a delay introduced between the Study and Recognition phases of the task. If forgetting of perceptual details leads to formation of a generalized representation, then introduction of a delay between the Study and Recognition phases of the task should result in accuracy patterns similar to those generated by category-based processing (i.e., low recognition accuracy generated by a pattern of high hits and high false alarms). Finding support for this prediction would point to a possibility that low-level memory mechanisms, rather than conceptual assumptions, may underlie category learning in young children.

EXPERIMENT 10

Method

Participants

Participants were 40 preschoolers (21 girls and 19 boys; \( M = 5.14 \) years; \( SD = .18 \)).

Materials and Procedures
Materials and procedure in Experiment 10 were identical to those described in the Basic Paradigm section (Chapter 1), with one important difference: the Recognition Phase did not immediately follow the Study Phase – it was administered after a 2-day delay. Participants were first randomly assigned to be tested in either Induction or Baseline condition, and the recognition test, identical for both conditions, was administered 48 to 54 hours later.

Results and Discussion

*Catch Trials Accuracy*

Data from participants who failed to correctly reject at least 5 out of 7 catch items in the Recognition Phase were excluded from further analysis. Based on this criterion data from two participants (one in the Induction and one in the Baseline condition) were excluded. The rest of the participants were highly accurate in rejecting catch items, averaging over 97% of correct responses across both experimental conditions.

*Induction Accuracy*

Similar to the previous reports (Sloutsky & Fisher, 2004a, 2004b; also see Experiments 1-2 of the present manuscript), in the Induction condition participants demonstrated high accuracy in the Study Phase of the Experiment, averaging over 85% of correct inductions. However, unlike previous reports, in this Experiment the recognition test was delayed by two days (instead of being administered immediately after the Study Phase). Under these conditions, participants’ recognition accuracy did not exceed chance level, A-prime = .49, \( t (17) < 1, p > .8 \). A follow-up analysis indicated that low
recognition accuracy was driven by a pattern of high hits (.77) and high false alarms (.71), both above chance, $t_s > 2.8, ps < .05$.

It is possible that low recognition accuracy in the Induction condition and a pattern of high hits and high false alarms were generated by rapid forgetting of item-specific information. However, it is also possible that patterns of recognition accuracy observed in Experiment 10 were driven by participants performing category-based induction during the Study Phase. Results of the Baseline condition help to distinguish between these possibilities: if low recognition accuracy in the Induction condition is due to participants being engaged in category-based processing during the Study phase, rather than to rapid forgetting of item-specific information, then in the Baseline condition we should observe high recognition accuracy with a pattern of high hits and relatively low false alarms. However, the pattern of results in the Baseline condition was similar to that in the Induction condition (see Figure 19): participants demonstrated overall low recognition accuracy ($A$-prime = .58, not different from chance, one-sample $t(19) = 1.36, p > .18$), generated by a pattern of high hits (.84) and high false alarms (.71), both above chance, one-sample $t_s > 3.1, ps < .05$. This indicates that results in the Induction condition were not driven by participants performing category-based induction during the Study Phase, but rather can be attributed to rapid forgetting of item-specific information.
SUMMARY

In Experiment 10 participants were presented with Induction and Baseline tasks of the basic paradigm, in which recognition test was delayed by two days. Results indicated that in both experimental conditions participants’ recognition accuracy decreased to chance (compared to non-delayed recognition in Experiments 1-2) due to a marked increase in the level of false alarms. This finding provides further evidence in support of the possibility that forgetting may be the mechanism underlying category learning and categorization. Specifically, these results suggest that generalized representations can be an emergent property of item-based processing.
Based on the hypothesis that forgetting of item-specific information plays a crucial in
category formation, it is possible to make several important predictions. In particular, if
forgetting is the mechanism underlying categorization and category learning, then
participants’ performance on tasks similar to the task used in Experiment 10, should be
sensitive to such factors as the amount of items to be held in memory and duration of the
delay, in such a way that performance should deteriorate with increase in both the
number of items and the length of the delay. Additionally, testing participants with novel
categories (which preclude the possibility of generalized representations existing prior to
testing) will provide a crucial test to the hypothesis that memory mechanisms underlie
category learning. Testing these predictions will further elucidate mechanisms underlying
category learning and categorization.
CHAPTER 8

GENERAL DISCUSSION

Findings presented in Chapters 3-7 point to several important regularities. First, when presented with members of familiar categories young children, unlike adults, spontaneously perform similarity-based induction, and category-based induction is a product of gradual development (Experiment 2, Chapter 4). Second, when adults are presented with members of novel categories, they generalize in a similarity-based manner (Experiment 5, Chapter 4). Third, providing category labels does not automatically promote category-based induction in young children (Experiment 6, Chapter 5). Fourth, children can be trained to perform induction by relying on category labels, with retention of training being a function of age: older children are more likely to retain effects of training over time than younger children (Experiment 7, Chapter 5). Fifth, when training conditions closely match testing conditions, purely associative training is as effective in increasing attentional weights of labels for induction in young children, as training including conceptual components (Experiments 7-8, Chapters 5-6). Finally, results of Experiment 10 (Chapter 7) elucidate the role of rapid forgetting of item-specific information in categorization and category learning. These findings have several important theoretical implications, which are discussed below.
Development of Induction

Theoretical approaches assuming reliance on conceptual knowledge in the course of induction have been very influential in the past decade; however many recent findings pose a challenge to this approach (e.g., Smith, Jones, & Landau, 1996; Jones & Smith, 2002; Sloutsky, 2003; Sloutsky & Fisher, 2004a, 2004b; Sloutsky & Spino, 2004; Colunga & Smith, 2004). Results presented here further challenge the idea that early in development induction is category-based: the reported findings are inconsistent with the idea that young children first spontaneously identify presented entities as members of the known categories, and then generalize properties to other members of the same categories. These findings suggest that it is unnecessary to posit conceptual assumptions to account for inductive generalizations in young children, thus supporting the recently proposed SINC model (Sloutsky & Fisher, 2004a), which argues that for young children, both induction and categorization are similarity-based processes.

Note that the claim that there is a gradual transition from similarity-based to category-based induction does not imply that adults perform induction solely in the category-based manner. When categories are novel (as in Experiment 5, Chapter 3), adults perform similarity-based induction, whereas when categories are familiar (as in Experiment 2, Chapter 3), adults perform category-based induction. Therefore, while mature induction can be either similarity-based or category-based (depending on conditions), early induction is likely to be similarity-based. Thus, it seems reasonable to conclude that similarity-based induction is a default, whereas category-based induction is a product of learning and development.
The reported results also present developmental and learning accounts of category-based induction. First, category-based induction gradually emerges in the course of development: there is little evidence that 5- or 7-year-olds spontaneously perform category-based induction, however 11-year-olds are more likely to perform it than younger children, and adults are more likely to perform it than 11-year-olds. Second, conceptual assumptions (such as a category assumption and a linguistic assumption) are not necessarily prerequisites for early generalization: Experiments 3-4 demonstrated that young children can be trained to perform induction by relying on shared category labels, and retention of this training was shown to be a function of age, with older children being more likely to retain effects of training than younger children.

Recall that in Experiment 3, participants were taught that (a) animals that have the same name belong to the same kind, (b) animals that belong to the same kind share important properties, (c) animals that have the same name share important properties. It is possible that (a) and (b) are acquired through formal schooling, with (c) being a direct consequence of (a) and (b). These considerations suggest that category-based induction could be a product of learning.

Furthermore, Experiment 9 demonstrated that increasing attentional weights of labels without introducing conceptual explanations, is sufficient to promote label-based induction in young children, at least when training conditions closely match testing conditions. Therefore, results of Experiments 3, 4 and 8 can explain the transition from similarity-based to label-based induction, suggesting that mature induction could be a product of feedback-based learning. It is, however, still unclear whether this learning has
to be conceptual in nature or purely associative learning can account for this transition. This question is of great importance and needs to be addressed in future research.

Generalization and Category Labels

It has long been known that labels play an important role in cognition, facilitating such processes as categorization and induction (Gelman & Markman, 1986; Balaban & Waxman, 1997; Welder & Graham, 2001; Waxman & Booth, 2003; Sloutsky & Fisher, 2004). It has also been proposed that these facilitative effects of labels on induction and categorization stem from language-specific factors: according to the language-specific proposal, linguistic, but not other auditory input, direct attention to commonalities among presented entities, thus facilitating categorization (Gelman & Markman, 1986; Balaban & Waxman, 1997; Booth & Waxman, 2002; Waxman & Booth, 2003; Gelman, 2003). Furthermore, infants and young children assume that labels refer to categories, and this linguistic assumption drives early categorization and induction.

However, recent evidence suggests that early in development linguistic and non-linguistic auditory input have similar effects on processing (Campbell & Namy, 2003; Robinson & Sloutsky, 2004b). Therefore, effects of labels on induction and categorization could also stem from general auditory rather than language-specific contribution of labels, and learning is likely to be a major factor in the development of specific effects of language on cognition (Samuelson & Smith, 1998; Campbell & Namy, 2003; Sloutsky & Napolitano, 2003; Napolitano & Sloutsky, 2004; Robinson & Sloutsky, 2004a).
Findings reported in Chapter 5 contribute to the growing body of evidence, suggesting that effects of language on cognition early in development are likely to stem from general auditory rather than from language specific factors. It has been previously demonstrated that (1) in infants auditory input often overshadows visual input when visual and auditory input co-occur, such that only auditory but not visual information is encoded (Robinson & Sloutsky, 2004a), (2) overshadowing effects in 16-month-olds are moderated by familiarity with the auditory stimulus, such that increasing familiarity with auditory input, increases infants’ ability to encode input from both, visual and auditory modality (Robinson & Sloutsky, 2004a, 2004b), (3) similarly, in 4 year-olds preference for encoding auditory or visual input is moderated by the stimulus familiarity, such that more familiar stimuli are more likely to be encoded regardless of the presentation modality, when bimodal stimuli are presented for brief periods of time (Napolitano & Sloutsky, 2004), and (4) adults exhibit strong visual preference, though they are fully capable of encoding all components of brief co-occurring bimodal stimuli (Sloutsky & Napolitano, 2003). These findings suggest that early in development, auditory and visual information might compete for attention, and that this tendency decreases with age. If auditory input competes for attention with visual input and tends to overshadow visual input early in development, then one might expect decreased categorization performance when visual stimuli are accompanied by auditory stimuli, and there is recent empirical evidence to support this prediction: it has been demonstrated that non-speech sounds, as well as linguistic labels decrease categorization in 8- and 12-month-old infants, compared to the condition in which to-be-categorized entities are presented in isolation (Tobin, Howard, Robinson, & Sloutsky, 2004).
It has been widely accepted that linguistic labels facilitate categorization in infants, however studies that demonstrated these facilitative effects of labels, compared labels to other auditory input, but not to the no-auditory baseline (Balaban & Waxman, 1997). When no-auditory controls were introduced, it was found that (1) infants accumulate more looking time when visual input is accompanied by auditory input, but at the same time (2) both labels and non-speech sounds hinder categorization compared to the no-auditory baseline. Findings presented in this manuscript contribute to this line of research, suggesting that when making inductive generalizations about members of familiar categories, 5 year-old children are more likely to (1) process visual information and (2) to make correct generalizations, when visual input is not accompanied by linguistic labels (unless children are trained to rely on matching labels to perform induction).

Mechanisms of Learning Early in Development

Associative learning mechanisms have been widely studied in both animals and humans (for review see Hall, 1991), however it is still unclear “to what extent … the processes of human learning emerge from complex configurations and elaborations on the elementary learning processes observed in animals” (Gluck & Bower, 1988). Some researchers suggest that most of early human learning is driven by such elementary associative processes (Smith, Jones, & Landau, 1996; Smith, 2000; French, Quinn, & Mareschal, 2001; McClelland & Rogers, 2003; Sloutsky, 2003; Colunga & Smith, 2004), whereas others argue that children rely on conceptual knowledge to assist their learning
even early in development (Gelman & Coley, 1991; Soja, Carey, & Spelke, 1991; Akhtar & Tomasello, 2000; Booth & Waxman, 2002; Gelman, 2003).

For example, Booth and Waxman (2002) argued that young children rely on their knowledge of ontology when presented with a task of word learning. They presented 3 year-olds with novel objects in either an animate or an artifact context, and manipulated context through short stories, in which a novel object was either loved, kissed, and hugged by mommy and daddy (i.e., animate context), or was used by astronauts and replaced if broken (i.e., artifact context). After hearing the stories children were presented with a label generalization task. It was found that children used information from the stories to guide their generalizations: they generalized novel labels by both shape and texture in the animate condition, and solely by shape in the artifact condition. Booth and Waxman (2002) concluded that children used conceptual information presented in the stories to access their knowledge of ontological kind, and their generalizations were influenced by this knowledge (i.e., that shape is central for artifacts, and that texture and shape are both important for living things). However, Colunga and Smith (2004) argued that short stories used by Booth and Waxman (2002) influenced children’s label generalizations not because of their conceptual but rather because of their associative properties: they argued that words like mommy, daddy, and kiss automatically activated a cluster of correlated features, such as eyes, legs, and breathe, and it was this automatic activation that directed children’s label generalizations. To support this argument, Colunga and Smith (2004) demonstrated that priming children with a list of relevant features, presented outside of a coherent story context, is sufficient to influence children’s performance on a label extension task. In particular, participants primed with artifact
context words (i.e., take, use, break, etc.) generalized labels from the target object to the items that were identical to the target, matched the target by shape only, shape and texture, and shape and color. In contrast, participants primed with animate context words (i.e., love, sleep, happy, etc.) generalized labels only to the objects that were either identical to the target object, or matched the target in both shape and texture.

Findings presented Chapters 5 and 7 provide further support to the associative accounts of early learning, while challenging the notion that learning is conceptually driven even early in development. In particular, in Chapter 5 it has been demonstrated that purely associative training can be as effective in promoting label-based generalizations, as training including conceptual components. At the same time, results presented in Chapter 7 suggest that forgetting of item-specific information could be one of the mechanisms underlying category learning: individual details that differentiate among stored exemplars tend to be rapidly forgotten, thus leaving one with a summary-like representation of encountered items. Other attentional mechanisms that have been implicated in category learning and categorization are overshadowing and tuning. Specifically, it has been demonstrated that in infants and young children novel labels as well as vowel strings partially overshadow visual information, and as a result of this auditory overshadowing, visual discrimination gets attenuated compared to the no auditory input condition (Sloutsky, Robinson, Timbrook, 2005; Sloutsky & Robinson, under review). This evidence suggests that auditory information promotes grouping or categorizing together entities that are discriminable under the no auditory input conditions, by overshadowing differences among compared items. However, when infants are pre-familiarized to labels or sounds, they are more likely to discriminate visual
stimuli compared to the no auditory input condition. In other words, instead of overshadowing visual input, familiar labels and sounds tune attention to the corresponding visual stimuli (Sloutsky & Robinson, 2005). Taken together, these findings provide evidence that low-level memory and attentional mechanisms may underlie early category learning, and cast on the arguments that early learning is conceptually driven.

However, previous studies left unanswered many important questions that need to be addressed in further research. Specifically, it is unclear whether there is an added benefit of conceptually driven learning. To answer this question one needs to compare purely associative training with training that includes both conceptual and associative components. These training procedures need to be compared directly to one another within the same experimental task (notice, that Experiments 7-8, which provided preliminary evidence for an added benefit of conceptual training over purely associative training, used different training procedures). Another important set of questions concerns the minimum required amount and retention rates of associative training compared to conceptual training. One possibility is that for young children purely conceptual training is ineffective. Another possibility is that purely conceptual training is effective for young children and requires less practice to be successful, but is retained for shorter periods of time than purely associative training. Finally, it is important to examine changes in the effectiveness of conceptual and associative training at different points in development: it is possible that while effectiveness of purely associative training remains constant, effectiveness of purely conceptual training increases with age.
Implications for the Study of Memory

Results reported in Chapters 2-6 have broader implications for the study of memory and its development. First, present research further indicates that deeper levels of processing do in fact reduce recognition accuracy: across all experiments, category-based induction (which is associated with deeper semantic processing) resulted in less accurate recognition than similarity-based induction (which is associated with shallow perceptual processing).

Second, current results indicate that young children are less prone to the levels-of-processing manipulations: without training, even with familiar categories young children exhibit evidence of perceptual rather than semantic processing. These results suggest that the ability to encode the semantic level, or category-information, is a product of development. As a result, while adults can form both category-level (or “gist”) representations and item-specific representations, young children tend to form mostly item-specific representations. These considerations are consistent with the fuzzy trace theory of memory (Brainerd, et al. 2002).

Third, the reported findings point to interesting interrelations between categorization and recognition. While recognition has been typically thought of as a process closely correlated with categorization, current results suggest that the relationships could be inverse: greater categorization may result in lower recognition accuracy for individual items, and high recognition accuracy for individual items may indicate lower categorization of these items.

Finally, findings reported in Chapter 7 suggest that forgetting of item-specific information could be one of the low-level memory mechanisms underlying categorization
and category learning. These findings have important implications for the debate between the proponents of prototype and exemplar-based models of classification and memory. It has been suggested that single-system exemplar-based models are unable to account for the prototype effects in classification and recognition tasks (i.e., high false recognition rates for category prototypes never presented during study; for discussion see Murphy, 2002; Nosofsky, 1991). There have been several demonstrations that exemplar-based models can in fact account for the prototype effects (Nosofsky, 1988; Nosofsky, 1991); findings presented in Chapter 7 further support this notion and propose a mechanism by which a high-level process of prototype abstraction can be simulated with low-level exemplar-based processing.

Conclusions

Results of experiments reported in Chapters 2-6 indicate that (1) when presented with members of familiar animal categories, young children spontaneously perform similarity-based induction and adults spontaneously perform category-based induction; (2) there is a gradual transition from similarity-based to category-based induction with familiar categories, with category-based induction likely being a product of learning; (3) category labels do not automatically promote category-based induction in young children, however children can be trained to perform label-based induction; moreover, both associative and conceptual training procedures are effective in training children to rely on shared labels to perform induction; and (4) generalized representations are an emergent property of item-based processing, and forgetting of item-specific information might be one of the mechanisms underlying categorization and category learning. These results support the
similarity-based account of young children’s induction, and present a challenge to the naïve theory approach, which assumes that young children’s induction is a function of preexisting conceptual knowledge.


