FUNCTIONS, FORMS AND FEPS:
HOW INFANTS USE LANGUAGE TO LEARN ABOUT ARTIFACTS

by

ANJULA JOSHI

B.A., Simon Fraser University, 1995

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

(Psychology)

THE UNIVERSITY OF BRITISH COLUMBIA

August 2005

© Anjula Joshi, 2005
Abstract

Two experiments investigated the influence of object labels and perceptual similarity in 18-month-olds' inductive inferences within the domain of artifact kinds. In Experiment 1, infants learned structure-dependent functions of training objects, and were presented with test objects that varied in perceptual similarity. When objects were labeled with an adjective or were not labeled, infants generalized functions to perceptually similar test objects, but when objects were labeled with a count noun, infants generalized functions to all test objects, regardless of perceptual similarity. In Experiment 2, we found a different pattern emerge with structure-independent functions. Infants generalized these functions to other objects, but perceptual similarity, count noun labels, and adjective labels did not affect performance. These results suggest that 18-month-olds differentiate count nouns from adjectives, and only count nouns refer to kinds, override perceptual similarity and guide inductive inferences about structure-dependent functions.
# Table of Contents

Abstract .......................................................................................................................... ii

Table of Contents ........................................................................................................... iii

List of Tables .................................................................................................................... iv

List of Figures ................................................................................................................... v

Acknowledgements .......................................................................................................... vi

INTRODUCTION ........................................................................................................... 1
   Studies of Inductive Inference – Evidence from Preschoolers .................................. 3
   Studies of Inductive Inference – Evidence from Infants ............................................. 6

EXPERIMENT 1 ............................................................................................................. 13
   Method ......................................................................................................................... 14
      Participants .............................................................................................................. 14
      Stimuli ...................................................................................................................... 15
      Design ...................................................................................................................... 16
      Procedure ............................................................................................................... 17
   Results ......................................................................................................................... 18
   Discussion .................................................................................................................... 21

EXPERIMENT 2 ............................................................................................................. 23
   Method ......................................................................................................................... 24
      Participants .............................................................................................................. 24
      Stimuli ...................................................................................................................... 24
      Procedure ............................................................................................................... 24
   Results ......................................................................................................................... 25
   Discussion .................................................................................................................... 27

GENERAL DISCUSSION .............................................................................................. 28
   Summary ...................................................................................................................... 28
   Perceptual Similarity vs. Count Noun Labels .............................................................. 30
   Count Noun Labels vs. Adjective Labels .................................................................... 32
   Structure-Function Relationships ............................................................................. 33
   Conclusion ................................................................................................................... 36

References ...................................................................................................................... 38

Footnotes ........................................................................................................................ 42
List of Tables

Table 1. Mean and standard deviations of target actions performed on test objects .................43
List of Figures

Figure 1. Stimulus object sets used in Experiments 1 and 2 ..........................................................44

Figure 2. Mean number of target actions performed on the test objects in the Surprised Trials in
Experiment 1 (Structure-dependent functions) ..................................................................................45

Figure 3. Mean number of target actions performed on the test objects in the Surprised Trials in
Experiment 2 (Structure-independent functions) ..............................................................................46
Acknowledgements

Many thanks to my thesis supervisor, Fei Xu, for her invaluable guidance and support, and my committee members, Geoff Hall, and Janet Werker, for their very helpful suggestions and insights. I would also like to acknowledge the members of the Baby Cognition Lab for all of their assistance with these studies, their ideas and discussions, and Sean, for his suggestions, interesting discussions, patience and humour.
Introduction

What is the relationship between language and conceptual development? Although most psychologists reject extreme versions of the Whorfian hypothesis – that language determines thought and behavior (Whorf, 1956) – in recent years, some researchers have explored the possibility that the language we hear may have a strong impact on the way we understand the world, and have suggested that the process of language acquisition may help structure developing concepts (Bowerman & Levinson, 2001; Gentner & Goldin-Meadow, 2003). One of the ways in which researchers have begun to look at the impact of language on conceptual development has been to study how infants and young children categorize and make inferences about objects.

Inductive inferences involve reasoning beyond available evidence and inferring that a property of one member of a category holds true for other members of the same category. For example, if we know a cat that meows, purrs, and eats tuna, we might infer that other cats also meow, purr and eat tuna. As such, categories provide a way to organize our knowledge and guide our expectations about new instances of a category. In this way, inductive reasoning can be a powerful learning mechanism for infants and young children. There is evidence that young children and even infants are able to infer that members of a category share properties that are not perceptually obvious, and are able to make inferences about these properties to new exemplars of a category (Baldwin, Markman, & Melartin, 1993; Gelman & Coley, 1990; Welder & Graham, 2001)

A continuing controversy in this field of study is the relative importance of perceptual and conceptual information when forming categories. Some researchers suggest that infants and children are “perceptually-bound”, or inclined to attend to perceptual features, such as shape and parts, when categorizing objects (Imai, Gentner, & Uchida, 1994; 1993; Landau, Smith, & Jones, 1988; Quinn & Eimas, 2000; Wellman & Gelman, 1988). That is, infants and children are
thought to categorize objects on the basis of common perceptual appearances, rather than underlying conceptual qualities. Other researchers have suggested that while perceptual appearance, such as shape, is often the best cue to an object's kind, it is not the sole factor in determining kind (Bloom, 2001). For example, for objects such as spoons, shape is a very reliable cue to their category identity. However, object kind is generally related to deeper conceptual properties that objects share, and objects of the same kind do not necessarily share common perceptual appearances. For example, the conceptual category of fruit may include apples and bananas, which do not share many perceptual characteristics, yet are easily understood to belong to the same conceptual category.

Some studies have shown that young children are able to form categories on the basis of conceptual cues in the form of a shared label (Davidson & Gelman, 1990; Gelman & Coley, 1990; Welder & Graham, 2001). Count noun labels are generally interpreted as identifying kind membership (Markman, 1989), and objects that we encounter in the world that share a label generally share perceptual similarities as well. Thus, in order to determine which cue is weighted more heavily, young children are presented with situations in which perceptual cues conflict with shared labels. These studies have found that young children demonstrate an understanding that objects that have been labeled with the same word belong to the same category, and infer that they may share nonobvious, internal properties, despite differences in perceptual appearance.

The present research focuses on the domain of artifact kinds and is aimed at investigating how infants' patterns of inference change as a function of perceptual similarity and conceptual information in the form of shared labels. If infants interpret count noun labels as a strong cue to an object's kind, we may observe infants to rely more on count noun labels than perceptual similarity when judging category membership. Following this logic, we may also observe that infants will make inductive inferences between objects that share count noun labels, and do not necessarily share perceptual appearances.
Many of the early investigations of categorization and inductive inference in preschoolers have demonstrated that children are capable of using categories as the basis for novel inferences about objects. Gelman and Markman (1986) developed a triad task to examine how four-year-old children understand natural kind categories and whether they place more weight on category membership or outward appearance when drawing inferences between objects. In this task, category membership was put into conflict with perceptual similarity. The experimenter showed the child three pictures, such as a flamingo, a bat, and a blackbird. As the pictures were presented, they were labeled (“Here is a bird, here is a bat, and here is a bird”). The blackbird was similar in appearance to the bat and dissimilar to the flamingo. The labels, however, identify the blackbird and flamingo as belonging to the same category: bird. The child was then told a new fact about the flamingo and the bat, and asked which they thought was true of the blackbird. If children categorize according to perceptual similarity, they would be expected to generalize the property of the bat to the blackbird. In contrast, children categorized the blackbird according to its label, and generalized the property of the flamingo to the blackbird, despite perceptual differences. This study demonstrates that young children are able to draw inferences from one member of a category to another based on the presence of a common label to identify them as the same kind, even when their appearances are markedly different.

An alternative account might suggest that children’s inductive inferences are guided by their familiarity with and knowledge about specific categories of objects, such as birds and bats. Davidson and Gelman (1990) tested this hypothesis by using a similar paradigm to examine four- and five-year-old children’s ability to draw inductive inferences with pictures of novel objects that have been labeled with novel words. They reasoned that if children are able to make inferences about unfamiliar objects, their knowledge of and experience with the categories could not explain their performance. Indeed, preschoolers did make inductive inferences on the basis of
novel labels for unfamiliar objects, even when conceptual cues were not entirely consistent with perceptual appearances. Four- and five-year-old children appear to be able to rapidly form categories with objects they have never seen before, and generalize properties to novel exemplars of those categories. However, children did not make these inferences indiscriminately. When objects sharing a label were radically different in appearance children did not draw inferences across all exemplars. They tended to generalize properties to perceptually similar objects. That is, when children were presented with test objects that, in relation to the target object, had the same appearance but a different label, or a different appearance and same label, they tended to rely on perceptual appearance. Children took category labels into account when the conflict between labeling and perceptual similarity was reduced, perhaps enabling children to understand how or why the category labels mapped onto the objects. These findings suggest that by the time children are about four years old, their real-world experiences with objects lead them to realize that it is highly unusual to encounter objects that are dramatically different in appearance, yet share the same name. Preschool-age children appear to be sensitive to the reliable relationship between appearance and labels, and are able to use this expectation to guide their inductive inferences.

Gelman & Coley (1990) established that children as young as two years are able to employ linguistic information to make inferences about categories of objects. Demonstrating this ability with younger children, however, required some scaffolding and additional support. In this study, children were shown pictures of familiar, prototypical objects and reminded of a familiar property of that object. For example, the experimenter showed children a picture of a bluebird, labeled it as a bird, and reminded them that it lives in a nest. Children were then tested on whether they generalized that property to a bat that was similar in appearance to the bluebird, or a flamingo that was different in appearance from the bluebird, but was also labeled as a bird. In the absence of a label, children generalized the property to perceptually similar bat. When
objects were labeled, two-year-olds relied on category labels and inferred that objects that shared
the same label would also share the same property. Thus, by two years of age, language can
provide a reliable source of information and enable children to look beyond surface appearances
and form categories based on linguistic information and underlying properties.

The experiments described thus far have examined children’s ability to make inferences
about nonobvious properties in natural kinds – animals, plants, nonliving natural kinds such as
gold (Gelman & Coley, 1990; Gelman & Markman, 1986) and novel animals (Davidson &
Gelman, 1990). Gopnik and Sobel (2000) demonstrated that 2-, 3-, and 4-year old children are
able to draw inferences about the causal properties of objects. In this study, children were shown
four wooden blocks. In some trials, two blocks were labeled with a novel word (e.g. “blickets”),
and the other two blocks were described as “not blickets”. Labeling was put into conflict with
perceptual similarity, whereby the labeled pair differed from each other in shape and/or color,
and each one was perceptually identical to one of the blocks in the unlabeled pair. The
experimenter demonstrated that placing one of the labeled blocks on a machine causes the
machine to activate. Children were then asked to choose another object that would activate the
machine. Three- and four-year old children accurately inferred that the other blicket would
activate the machine, even when faced with conflicting perceptual information. In contrast, two-
year-old children did not rely on object labels, but instead, selected the perceptually similar
object from the unlabeled pair. These researchers suggest that very young children are more
likely to assume a correlation between perceptual appearance and causal property than older
children, and have not yet developed the ability to overlook perceptual information and rely on a
category label. Three- and four-year olds however, appear to interpret object labels as more
important than perceptual similarity when making inferences about object categories and their
causal properties. These findings are somewhat at odds with the results from Gelman and Coley,
who found that two-year-old children did rely on category labels. However, the tasks and
inferred properties are very different in each study. While Gopnik and Sobel examine whether young children are able to infer a causal relation through observation, the object properties in Gelman and Coley's study are conveyed verbally. These differences may explain the apparent discrepancy in two-year-olds' performance in these two studies.

This body of research demonstrates that 2- to 5-year-olds categorize perceptually similar objects, and expect members of the same category to possess the same underlying properties. Preschoolers are also sensitive to category labels, although not indiscriminately. They are not as likely to draw inferences between objects that are radically different in appearance, even if they share a label. This is likely due to their experiences with objects in the world, where they observe reliable associations between perceptual appearances and object labels.

*Studies of Inductive Inference – Evidence from Infants*

In recent years, researchers have begun to examine inductive reasoning abilities in infants. These studies have shown that the ability to draw inferences between objects that are perceived to be members of the same category has its roots in infancy. Because verbal responses that are typically used in induction tasks with preschoolers are not possible with infants, generalized imitation techniques involving manipulations of stimulus objects have been developed (Baldwin et al., 1993; Mandler & McDonough, 1998). In this paradigm, the experimenter demonstrates an action on a target object and then presents the infant with a test object. If the infant perceives the test object as a member of the same category as the target object, the infant is expected to perform the target action on the test object, demonstrating that they have made an inference.

Studies by Mandler and McDonough (1996; 1998) have focused on infants’ existing knowledge of conceptual categories to guide their inductive generalizations. Using the generalized imitation technique, the experimenter modeled actions on toy animals and vehicles and observed whether infants would imitate the actions on different exemplars from the same or
different domains. For instance, the experimenter modeled giving a dog a drink from a cup, and then presented infants with an object from the same domain (e.g. cat), and an object from a different domain (e.g. car). Infants generalized the modeled actions (e.g. drinking) to other members of the same category (e.g. other animals), but did not cross category boundaries and generalize to objects of a different category (e.g. vehicles). Mandler and McDonough suggest that infants rely on their conceptual knowledge about animal and vehicles, and their inductive generalizations reflect these conceptual categories.

Using novel objects, Baldwin et al. (1993) investigated how readily 9- to 16-month-old infants make inferences about nonobvious properties of objects. The presence of these nonobvious properties could not be detected directly, but instead, had to be inferred by the infant. The experimenter demonstrated a nonobvious property of a target object by performing a specific target action. For example, the experimenter presented the infant with a can that wails when it is tilted or shaken. In the violated expectation condition, the target objects possessed the nonobvious property, but the test objects did not. This condition enabled the experimenters to assess whether infants expected the test object to possess the nonobvious property. If infants held the expectation that the test object possessed the property, they should persist in performing the target action in an attempt to elicit the sound property. The interest control condition, in which both target and test objects were altered so that they did not possess the property (e.g. the can was silent), provided a baseline measure of infants’ exploratory manipulations of the objects without receiving any demonstration or sound reinforcement. The results indicated that infants performed more target actions in the violated expectation condition than the interest control condition, confirming that infants formed an expectation based on the target object. In Experiment 2 infants were presented with test objects that were either perceptually similar or different from the first object. Infants expected perceptually similar objects to share similar underlying properties, and were less likely to make this inference with objects that differed in
appearance. Although this finding appeared to hold for infants across the 9 to 18 month age range, the authors acknowledged that it was not as apparent until infants were approximately 11 months old. Thus, by about 11 months, infants are able to quickly learn a nonobvious property of an object after a very brief experience, and expect perceptually similar objects to share this property. The results of these two studies provide clear evidence of inductive abilities in infancy, and support the notion that infants are sensitive to the underlying properties of objects, and not just surface features.

Using a similar paradigm, Welder and Graham (2001) demonstrated that infants can rely on both perceptual and conceptual information that is provided during the experiment. Although Baldwin et al.'s (1993) findings point to infants' use of perceptual appearance to guide their inferences, their use of stimulus objects that were either very similar or radically different did not provide any information regarding the degree of similarity required to support inductive inferences. To address this issue, Welder and Graham developed stimulus objects that systematically varied in shape along a continuum to find the degree of shape similarity necessary for infants to perceive objects of the same category. In addition to examining the influence of shape similarity, they also introduced object labels to test the effect of language on infants' inductive inferences.

In this study, 16- to 21-month-old infants were presented with novel target objects, and the experimenter demonstrated a target action that elicits a nonobvious (sound) property. For example, the experimenter demonstrated to the infant that a cloth-covered shape produces a squeaking sound when it is squeezed. Infants were then presented with a test item that was a high, medium, or low similarity match, or a dissimilar object. The high, medium, and low similarity objects shared similar textures, but varied in shape and color. The dissimilar objects differed from the target object in shape, color and texture. The number of target actions that infants performed on the test objects indicated whether they generalized the nonobvious property
from the target object to the test objects. Replicating findings from Baldwin et al., (1993), infants inferred that objects that were perceptually similar to the target object would also have the same nonobvious property.

In addition to perceptual similarity, infants were also sensitive to the object labels. Infants performed more target actions at each level of shape similarity when the objects were labeled with a novel noun than when they were not labeled. However, they did not find an interaction between labeling and similarity, and infants continued to rely on perceptual similarity. Infants did not overlook shape information, but were sensitive to both labeling and perceptual cues. The role of count noun labels was unclear, and it is possible that count nouns only increased infants’ attention to the objects and actions, but did not convey conceptual information.

In a follow-up study, Graham, Kilbreath and Welder (2004) tested 13-month-old infants, and found similar results. In this study, infants were not presented with high, medium and low similarity objects, but high and low similarity and dissimilar objects. When objects were not labeled, 13-month-old infants generalized the nonobvious properties to the high similarity objects. When objects were labeled with a novel noun, infants appeared to perceive high and low similarity objects, but not dissimilar objects, as members of the same category and generalized the nonobvious property to these objects. However, the authors did not address why 13-month-olds’ performance reflects an ability to rely solely on object label information for low similarity objects, while in their previous study, 16- to 21-month-olds relied on both perceptual similarity and shared labels. In addition, the overall number of target actions performed by 13-month-old infants was quite low, in contrast to the 16- to 21-month-olds. These two issues suggest that the findings from 13-month-old infants may not be robust.

The current studies are aimed at addressing several gaps in the literature that have not been adequately addressed thus far. The first specific question is whether count noun labels enable 18-month-old infants to overlook perceptual appearances when determining category
membership. The second question asks if count noun labels are privileged over words from other form classes, such as adjectives, in guiding infants’ inductive inferences. The final question is with regard to the nature of infants’ conceptions of artifact kinds and functions. Specifically, we ask, what are infants’ patterns of inference when the demonstrated property is a function that is structure-dependent (e.g., a castanet is made to have a particular structure to produce a sound) versus structure-independent (e.g., the same castanet can be used to ring a bell, as could many other kinds of objects)?

To address the first question, Experiment 1 examines whether infants interpret count noun labels as more informative than perceptual appearance when making inferences about categories of objects. Welder and Graham’s (2001) findings suggest that although infants are sensitive to count noun labels, they continue to rely on perceptual similarity when determining whether two objects belong to the same category. That is, although infants performed more target actions when objects were labeled with a novel noun than when they were not labeled, they generalized object properties more often to high similarity objects, than medium or low similarity objects. However, there is some evidence that toddlers and preschoolers treat count noun labels as indicators of object kinds (Gelman & Coley, 1990; Gelman & Markman, 1986). Objects that have been labeled with the same count noun are viewed as belonging to the same category, and as members of the same category, they may be thought to share important, underlying properties (Gelman, 2003). In this way, count noun labels can support inductive inferences between category members. In Experiment 1, our goal was to determine the role of count noun labels in infants’ inductive inferences, and specifically, whether novel count noun labels enable infants to override conflicting perceptual information.

The second question we address in Experiment 1 is whether count nouns are privileged over adjectives in guiding infants’ inductive inferences. One important feature of language is that words from different grammatical classes correspond to different kinds of concepts. The studies
presented here focus on contrasting nouns with adjectives. While count nouns label object categories, adjective meanings can be quite variable. They can refer to temporary states, such as sleepy and hungry, or permanent properties, such as brown and furry, or even properties that are essential to a category, such as alive. Because of this variability, adjectives may not necessarily support inferences between category members. Gelman & Coley (1990) found that adjective labels that describe transient properties (e.g. sleepy) do not support preschoolers’ inductive inferences in the same way as count noun labels. When adjectives, instead of nouns, were used to describe objects, children’s inferences did not differ from not using any labels at all. These results suggest that noun labels, but not adjective labels, are interpreted as conceptual markers of object kind, and facilitate children’s inferences between category members.

As early as 11 months, infants are beginning to identify nouns in fluent speech and link nouns to object categories (Waxman & Booth, 2003). An understanding of adjectives emerges slightly later: when hearing objects described with a novel adjective, infants at 14 months attend to both category or shape information as well as texture and color properties. For example, infants extend novel adjectives to other objects that belong to the same category (e.g. chairs) and other objects that share the same property (e.g. blue or fuzzy) (Waxman & Booth, 2001). By about 21 months, infants succeed in mapping novel adjectives to other objects with the same property (Waxman & Markow, 1998).

In the present study, if 18-month-old infants are able to differentiate nouns from adjectives not only syntactically but also semantically and conceptually, then we should expect that only count nouns would be interpreted as referring to object kinds. If this is the case, we would expect to observe that only count noun labels would enable infants to draw inferences between objects that may not be perceptually similar. At 18 months, infants may not yet be able to map a novel adjective onto a specific property, such as the demonstrated sound properties in these experiments. Infants who hear objects labeled with novel adjectives may rely on perceptual
similarity as a cue to category information, and their performance may be similar to those infants who do not hear any object labels.

The final question we focus on is the nature of infants' conceptions of artifact kinds and functions. One of the problems with the stimulus set used by Welder and Graham (2001) and Graham, et al. (2004) is that the objects may not have been perceived as real artifacts. The objects were somewhat shapeless and lacking in structure, and unlike real artifacts, they did not have functions that were related to their structure. For example, one object was a desk bell that was covered in fur. While the fur concealed the object to make the property nonobvious, it also gave the object the appearance of having a function that was unrelated to its structure. Additionally, the target actions that were required to elicit each object's nonobvious property (i.e., squeezing, shaking, hitting) are also actions that are commonly performed on other toys and may be familiar to infants. As a result, it remains unclear whether performing these actions actually indicates that infants had made inferences about the test objects.

The stimulus objects used in the present studies were designed to be real artifact kinds with structure-dependent functions. We define structure-dependent functions as functions that are specific to a particular kind of object, whereby the configuration of the object is not accidental. Without its structural configuration, the object would be unable to function in the way that it does. For example, the structure-dependent function of a shoe is to be worn on a foot. A shoe sole and the upper part of a shoe are shaped and configured such that a shoe can perform this function. An example of a stimulus object used in these studies is a toy castanet. Its structure consists of two hard, concave surfaces that are joined by an elastic hinge on one side, enabling the user to press together and release the two surfaces to produce its distinctive sound.

We contrasted structure-dependent functions with structure-independent functions, which we define as functions that are not closely linked to the structure of an object. Thus, the structural configuration of the object is not relevant to performing the function. As such,
structure-independent functions are not specific to a particular kind of object, and many other kinds of objects could potentially perform these functions. For example, a shoe could be used to wedge open a door. Although this function does depend on the fact that the shoe is solid and made of a material that allows it to be used to prop open a door, it does not depend on the parts of the shoe or the way in which the parts are configured. This function is clearly not specific to shoes and many other kinds of objects could be used to perform this function. The same castanet described above could also be used to ring a bell. This function does not depend on the configuration of the parts of the castanet, and other objects could easily be used to perform the same function, as long as they are made of a hard, solid material that can make contact with the bell.

In Experiment 2, we sought to determine whether infants in an inductive inference task are sensitive to the differences between structure-dependent and structure-independent functions. If infants conceive of structure-dependent functions as specific to particular artifact kinds, we would expect infants to generalize the function to other objects of the same kind, as determined by perceptual similarity or labeling. When the demonstrated function is structure-independent, it remains an open question whether linguistic labels or perceptual similarity would have any systematic effects. To our knowledge, there have not been any empirical studies contrasting infants' understanding of structure-dependent and structure-independent functions. One possibility is that if infants construe structure-independent functions as viable in many kinds of objects, they may draw generalizations to objects of the same kind, as well as objects of a different kind.

Experiment 1

The purpose of Experiment 1 was to examine the role of perceptual similarity and linguistic labels in guiding infants' inductive inferences about structure-dependent functions. The experimenter taught 18-month-old infants a structure-dependent function of a novel training
object, presented test objects that varied in perceptual similarity to the training object (high similarity, medium similarity and low similarity), and then observed whether infants generalized the function to the test object.

To determine whether infants' performance is due to an expectation they develop during the training, there were three types of trials: Predicted, Surprised, and Baseline (similar to those used by Baldwin et al., 1993, Welder & Graham, 2001, and Graham et al., 2004). In the Predicted trials, both the training and the test objects possess a sound property. In the Surprised trials, the training object possesses the sound property, but the test object does not. If infants perform the target action on the test object without the reinforcing sound, we can conclude that their actions reflect that they have drawn an inference from the training object to the test object. In the Baseline trials, no action is demonstrated for the infant, and the objects do not possess the sound property. As such, these trials provide a baseline measure of infants' exploratory actions with the objects.

Objects were labeled with a novel count noun, novel adjective, or the objects were not labeled. We expected to replicate previous findings in which infants in the No Label condition generalized the object functions to perceptually similar test objects. We predicted that infants in the Count Noun condition would perceive all objects as members of the same category and generalize the object functions more broadly, despite perceptual differences. Infants in the Adjective condition were not expected to draw category information from the adjective labels, and we predicted that infants in this condition would disregard the adjective label and rely on perceptual similarity for their inferences.

Method

Participants

Fifty-four 16- to 19-month-old infants completed the procedure successfully (25 girls, 29 boys; \( M = 18;0; \) \( SD = 27.17 \) days; range = 16;9–19;25). 22 additional infants were excluded from
the final sample due to excessive fussiness \((n = 5)\), parental interference \((n = 2)\), experimenter error \((n = 1)\), lack of performance of actions on training objects \((n = 4)\), and lack of performance of any target actions on any of the test objects in the Surprised trials (often due to shyness and/or hesitancy) \((n = 10)\). Infants were recruited by obtaining names from newspaper birth announcements, and contacting parents by mail and phone.

**Stimuli**

Infants were presented with nine training object - test object pairs, drawn from three sets of novel objects: an accordion set, a bell set, and a castanet set (see Figure 1). The sets were visually distinct from one another, and consisted of four objects: a training object, and three test objects. The test objects varied in perceptual similarity relative to the training object, and consisted of a high similarity object, a medium similarity object, and a low similarity object. The high similarity objects were constructed with the same shape and material as the test objects, but differed in colour. The medium similarity objects had the same shape as the test objects, but differed in material and colour, and the low similarity objects differed from the test objects in shape, material and colour.

**Adult Ratings.** To determine whether test objects could reliably be categorized as high, medium, or low in similarity in relation to the training object, 10 adults rated the overall similarity of each test object to the training object. The adults were presented with the objects one pair at a time in random order. The rating scale ranged from 1 (very similar) to 7 (very different). Adults rated the high, medium and low similarity objects as significantly different from one another, in the expected direction. The high similarity objects \((M = 1.8)\) were rated significantly more similar to the training object than the medium similarity objects \((M = 2.9)\), \(t(29) = 7.66, p < .001\), and the medium similarity objects were rated significantly more similar to the training object than the low similarity objects \((M = 5.67)\), \(t(29) = 14.97, p < .001\). However, the difference between ratings for the high and medium similarity objects \((M = 1.23)\) was
significantly smaller than difference between the ratings for the medium and low similarity objects \((M = 2.7), t(29) = 6.42, p < .001\). Thus, adult raters found a smaller difference between the high and medium similarity objects than between the medium and low similarity objects. The high and medium similarity objects shared the same shape, and differed in material and color, while the medium and low similarity objects differed in material, color and shape. Thus, shape appears to be an important cue that is weighted more heavily than color and material when judging perceptual similarity.

Each training and test object had two versions: one version produced a sound when the function was performed on it, and the second version had been modified to be incapable of producing the sound. The accordion set was constructed by modifying small toy accordions. Each accordion was covered in various materials to create high, medium and low similarity objects as described above. The sound was elicited in the functional objects by pushing together and pulling apart the two flat surfaces. The bell set was constructed by attaching a bicycle bell (a round knob connected to a bell) to a Styrofoam box. The boxes were covered with various materials, and the knobs were painted in different colors. The functional objects in this set produced a ringing sound when the knob was turned. The castanet set was made by modifying toy castanets that are available in different colors and shapes. The functional objects produced a clacking sound when the two surfaces connected by an elastic hinge were pressed together.

Two Sony digital video cameras were used to provide a visual and auditory record of the experimental session.

**Design**

**Groups.** Infants were randomly assigned to one of three groups: a Novel Noun group \((n = 18)\), a Novel Adjective group \((n = 18)\), and a No Label group \((n = 18)\).

**Trial type.** The training object - test object pairs were presented in one of three trial types: Predicted, Surprised or Baseline. In the Predicted trials, both the training and the test objects
possessed a sound property, which could be elicited by performing the demonstrated function. In the Surprised trials, the training object possessed the sound property, but the test object did not. These trials were of particular interest because infants’ actions on the test objects would indicate their expectations about shared properties. Specifically, if infants perceive the test object as belonging to the same category as the training object, they should expect it to possess the same property, and persist in attempting to perform the function if the object does not produce the expected sound. These trials were contrasted with the Baseline trials, where neither the training nor the test objects possessed the sound property. The Baseline trials provided a measure of infants’ exploratory actions with the objects without any demonstration, and allowed us to ensure that the target property was nonobvious on visual inspection. The Predicted trials allowed us to avoid the expectation that all test objects would not possess the sound property.

**Blocks.** The experiment consisted of 9 trials, and each trial consisted of a presentation of a training object followed by a test object (either high, medium, or low similarity). Each infant was presented with 3 blocks of 3 trials. The 3 trials from each block contained one trial of each item set (accordion, bell, castanet). For every infant, each item set was presented as either a Surprised, Baseline or Predicted trial. Trial type and item pairings were counterbalanced across infants, and order of presentation was randomized across blocks.

**Procedure**

The infants were tested individually in a quiet testing room. Infants sat in a high chair or on their parent’s lap at a table across from the experimenter. In the Predicted and Surprised trials, each test trial began with the experimenter presenting the training object to the infant and saying three phrases containing either a novel count noun (e.g. “Look, a fep!”), a novel adjective (e.g. “Look, a feppish one!”), or no label (e.g. “Look at this one!”). The experimenter then demonstrated the target action five times, and alternated each action with a labeling phrase. Thus, infants heard the training object labeled up to 8 times. Infants were allowed to play with
the training object for up to 10 seconds while the experimenter encouraged the child to imitate the action. If the infant was hesitant to manipulate the training object, the parent was asked to imitate the action and pass the object to the infant. Following the training phase, the training object was placed to one side of the table, out of reach of the infant but within view, and the infant was presented with a test object, which the experimenter labeled three times in the same way as the training object. The infant was allowed to play with the test object for up to 20 seconds. The Baseline trials followed the same procedure, with the exception that infants did not receive any demonstration with the training objects.

**Coding.** Two coders recorded the number of target actions that infants performed on the training and test objects. Coders followed a detailed scheme, which defined the target action for each object set. For the accordion set, each push inward of the two sides counted as one target action. For the bell set, the target action was defined as a hand motion that turns the knob in one direction, and for the castanet set, the target action was defined as each hand motion that presses the two surfaces together.

**Interrater reliability.** To establish interrater reliability, 75% of the data ($n = 42$ participants) were coded twice. Intraclass correlations (ICCs) were used to establish the level of agreement between the two coders. The ICC coefficient for test object frequency ratings was significant, $ICC(243) = .95, p < .001$. Thus, the two raters were in very close agreement, and any discrepancies were resolved by the two coders re-examining the videos together and coming to an agreement on the number of target actions performed by the infant.

**Results**

Before conducting the analyses, the data were screened for outliers. We eliminated trials where the infant did not imitate any target actions on the training object, as it would be difficult to interpret any actions performed on the test object ($n = 2$). Trials with frequency of target actions greater than 3 SDs above or below the mean in the Surprised or Baseline trials were
eliminated from the final analysis \((n = 4)\). We did not find any differences between boys and girls, thus we collapsed the data across these groups.

**Baseline trials**

A one-way analysis of variance (ANOVA) of the Baseline trials using number of target actions as the dependent measure revealed a main effect of similarity, \(F(2, 83.65) = 9.135, p < .001\). Pairwise comparisons within the Baseline trials indicated that infants performed significantly more actions with the high \((M = 1.49, SD = 2.93)\) than the low similarity objects \((M = 0.23, SD = 60)\), \(t(53) = 3.29, p < .001\), and more actions with the medium \((M = 1.29, SD = 2.35)\) than the low similarity objects \((M = 0.23, SD = 60)\), \(t(53) = 3.39, p < .01\). There were no significant differences between the high and medium similarity objects, \(t(53) = .82, p = .42\). Thus, to some extent, the target actions that infants performed in the baseline trials did seem to depend on perceptual similarity.

**Surprised vs. Baseline trials**

We next compared Surprised trials with Baseline trials, to determine whether the demonstrated properties were nonobvious and generally performed by infants only after they had received an initial demonstration with the training object. We conducted a three-way ANOVA with trial type (Baseline, Surprised) and similarity (high, medium, low) as within-subjects factors, and labeling (Count Noun, Adjective, No Label) as a between-subjects factor. There was a significant effect of trial type, \(F(1, 101) = 39.53, p < .001\). As predicted, infants performed more target actions in the Surprised trials \((M = 2.99, SD = 2.75)\) than in the Baseline trials \((M = 1.00, SD = 2.25)\).

To further analyze the differences between the Surprised and Baseline trials, we contrasted these trials at each level of shape similarity. These tests revealed that infants performed significantly more target actions in the Surprised trials than in the Baseline trials at each level of similarity. With the high similarity objects, infants performed significantly more
target actions in the Surprised ($M = 3.46, SD = 2.65$) than in the Baseline trials ($M = 1.49, SD = 2.93$), $t(53) = 4.45, p < .001$. Similarly, with the medium similarity objects, they performed more actions in the Surprised ($M = 3.94, SD = 3.09$) than in the Baseline trials ($M = 1.29, SD = 2.35$), $t(53) = 5.17, p < .001$, and again, with the low similarity objects, infants performed more actions in the Surprised ($M = 1.57, SD = 1.81$) than in the Baseline trials ($M = 0.23, SD = 0.60$), $t(53) = 5.38, p < .001$. Thus, at all levels of shape similarity, infants performed more target actions on test objects when they had first received the training.

**Surprised trials**

We next considered infants' performance within the Surprised trials, our main condition of interest. To test the effects of perceptual similarity and labeling on infants' target actions, we conducted a two-way ANOVA with similarity as the within-subjects factor and labeling as the between-subjects factor. The ANOVA yielded a significant main effect of similarity $F(2, 102) = 21.62, p < .001$, and a significant interaction between labeling and similarity, $F(4, 102) = 2.54, p < .05$ (see Table 1 and Figure 2).

Pairwise comparisons indicated that infants performed significantly fewer actions on the low similarity objects ($M = 1.58, SD = 1.81$) than on the high ($M = 3.46, SD = 2.65$), $t(51) = 4.89, p < .001$, or medium similarity objects ($M = 3.94, SD = 3.09$), $t(51) = 5.73, p < .001$ and there were no differences between the high ($M = 3.46, SD = 2.65$) and medium similarity objects ($M = 3.94, SD = 3.09$), $t(51) = 1.40, p = .17$.

We followed up the significant interaction by analyzing the effect of labeling at each level of similarity using planned comparisons. With the high similarity objects, there were no significant differences between the No Label ($M = 2.97, SD = 2.41$) and Noun groups ($M = 2.76, SD = 1.97$), $t(51) = 0.27, p = .79$, but infants performed significantly more actions in the Adjective group ($M = 4.64, SD = 3.16$) than the Noun group ($M = 2.76, SD = 1.97$), $t(51) = 2.21$, ...
p < .05. The difference between the Adjective (M = 4.64, SD = 3.16) and No Label groups (M = 2.97, SD = 2.41) was marginally significant t(51) = 1.94, p = .06.

With the medium similarity objects, we did not find significant differences between the No Label (M = 3.74, SD = 2.31) and Noun groups (M = 3.72, SD = 2.87), t(51) = 0.01, p = .99, the No Label (M = 3.74, SD = 2.31) and Adjective groups (M = 4.35, SD = 4.00), t(51) = 0.59, p = .56, or between the Noun (M = 3.72, SD = 2.87) and Adjective groups (M = 4.35, SD = 4.00), t(51) = 0.60, p = .55.

Finally, comparisons with the low similarity objects revealed a marginally significant difference between the Noun (M = 2.29, SD = 1.71) and No Label groups (M = 1.17, SD = 1.98), t(51) = 1.91, p = .06, as well as a marginally significant difference between the Noun (M = 2.29, SD = 1.71) and Adjective groups (M = 1.27, SD = 1.59), t(51) = 1.74, p = .09, and no differences between the No Label (M = 1.17, SD = 1.98) and Adjective groups (M = 1.27, SD = 1.59), t(51) = 0.17, p = .87. These results suggest that performance on the low similarity objects was affected by the way in which the object was labeled. Infants performed more target actions when the object was labeled with a count noun than when the object was not labeled, or labeled with an adjective.

An analysis of the items (accordion, bell, castanet) used in this study did not reveal a significant three-way (Item x Similarity x Labeling) interaction, F(8, 90) = 1.07, p = .39, indicating that the interaction between similarity and labeling did not depend on the type of item.

Discussion

The analyses from Experiment 1 revealed that infants performed more target actions in the Surprised trials, in which they had initially received training, and fewer actions in the Baseline trials, in which the action was not demonstrated. It seems that target actions were nonobvious and infants generally did not perform them unless they received the training first.
However, infants were affected by perceptual similarity in the Baseline trials. When the objects were more similar in appearance to the training object, the target action may have been more obvious to the infant. In an attempt to make the objects appear more like real artifact kinds with structure-dependent functions, the functions may not have been as nonobvious as the properties of the stimulus objects used by Welder & Graham (2001).

Infants were also affected by perceptual similarity in the Surprised trials, and performed more target actions on the high and medium similarity objects than on the low similarity objects. However, there was also an interaction between similarity and labeling. Infants performed more target actions on the high and medium similarity objects across the No Label, Noun and Adjective groups, but differences between labeling groups were found with the low similarity objects. Infants performed more target actions when the low similarity objects were labeled with a count noun than when they were not labeled, or labeled with an adjective. These results suggest that when objects were labeled with a count noun, infants overlooked perceptual information and treated high, medium and low similarity objects as members of the same category. In contrast, when objects were labeled with an adjective, or when they were not labeled, infants only treated perceptually similar objects as members of the same category. Infants in the Adjective group did not assume the adjective provided information about the object category, and instead, relied on perceptual similarity, when making inferences. These data suggest that count nouns are privileged in providing object category information and supporting inductive inferences about object properties.

The findings from Experiment 1 support the hypothesis that infants construe objects that have been labeled with the same count noun as members of the same kind. Moreover, infants expect objects of the same kind to function in similar ways and share properties that may not be perceptually obvious. The demonstrated functions in Experiment 1 were structure-dependent, where the functions were related to the structure of the objects. For example, infants were taught
to push together and pull apart the sides of an accordion to produce a sound. Infants generalized these functions systematically to new objects of the same kind, whether kind was based on perceptual similarity, or a shared count noun label. Perhaps infants are sensitive to structure-function relationships and perceive structure-dependent functions as relevant to particular kinds. For example, the demonstrated function of the accordion may be perceived as specific to accordions, because of the structural configuration that enables the function.

Alternatively, infants may not be sensitive to structure-function relationships, and may be willing to generalize any demonstrated function to other objects that are of the same kind.

To test this possibility, we investigated infants’ inferences about a different type of function. We taught infants structure-independent functions of training objects, or functions that are not closely linked to the structure of an object. As described earlier, structure-independent functions are not specific to a particular kind of object, and many other kinds of objects could potentially perform these functions. If infants are not sensitive to structure-function relationships, and are willing to generalize structure-independent functions to other category members, the results should not differ from Experiment 1. If, however, infants are sensitive to the differences between structure-dependent and structure-independent functions, a different pattern of results may emerge.

**Experiment 2**

The goal of Experiment 2 was to contrast the structure-dependent functions in Experiment 1 with structure-independent functions, to determine whether infants treat structure-dependent functions as intrinsic to particular artifact kinds. In Experiment 2, we demonstrated structure-independent functions using the same stimulus objects, design, and procedure from Experiment 1. If infants do not treat structure-independent functions as specific to a particular object kind, that is, if they view these functions as able to be performed by many other kinds of objects, then we would predict a different pattern of inferences from those of Experiment 1.
Infants may not generalize structure-independent functions on the basis of category membership, and as a result, linguistic labels and perceptual similarity might not influence their inductive generalizations.

Method

Participants

Thirty-six 16- to 19-month old infants completed the procedure successfully (20 girls, 16 boys; M = 18.0; SD = 24 days; range = 16.15-19.15). 7 additional infants were excluded from the final sample due to excessive fussiness (n = 4), or lack of performance on training objects (n = 3).

Stimuli

We used the same stimulus objects as Experiment 1, with minimal modifications. Each item set had an additional object that was used for demonstrating a structure-independent function. The structure-independent functions were designed to be similar to the structure-dependent functions in that they possessed an appealing sound property and were easy for infants to perform. The accordion was used to press upon a pillow to produce a squeaking sound, the bell was used to place inside a box that had a bell below the bottom surface (the weight of the bell inside the box would cause a ringing sound), and the castanet was used to tap against an ornament made up of a set of jingle bells. All of these functions could be performed by many other objects, and were not specific to any particular kind of object.

Procedure

The procedure was similar to that of Experiment 1, except that the demonstrated functions were structure-independent, as described above. All other labeling and trial type conditions remained the same.

Coding. Coders recorded the number of target actions that infants performed on the training and test objects. Similar to Experiment 1, coders followed a detailed scheme, which
defined the target action for each object set. For the accordion set, each press downward on the pillow using the accordion counted as one target action. For the bell set, the target action was defined as each instance the infant dropped the bell inside the box, or pressed the bell down inside the box. For the castanet set, the target action was defined as a tap on the ornament using the castanet.

*Interrater reliability.* To establish interrater reliability, all of the data \((n = 36\) participants) were coded twice. Intraclass correlations (ICCs) were used to establish the level of agreement between the two coders. The ICC coefficient for test object frequency ratings was significant, \(ICC(2|2) = .97, p < .001\). Thus, the two raters were in very close agreement, and any discrepancies were resolved by the two coders re-examining the videos together and coming to an agreement on the number of target actions performed by the infant.

**Results**

As in Experiment 1, the data were first screened for outliers. We eliminated trials where the infant did not imitate any target actions on the training object \((n = 3)\). We also excluded outlier trials where the frequency of target actions was greater than 3 SDs above or below the mean \((n = 4)\). We did not find any differences between boys and girls, thus we collapsed the data across these groups.

*Baseline trials*

The first analysis tested the effect of similarity in the Baseline trials. A one-way ANOVA revealed no effect of similarity, \(F(2, 42.69) = 1.95, p < .17\).

*Surprised vs. Baseline trials*

We next compared infants’ performance in the Surprised trials and Baseline trials. A 2 (Trial: Surprised, Baseline) x 3 (Similarity: High, Medium, Low) x 3 (Labeling: No Label, Count Noun, Adjective) ANOVA indicated that as predicted, infants performed significantly more target actions in the Surprised trials \((M = 2.42, SD = 2.55)\), than in the Baseline trials \((M = 0.32,\)
Paired $t$-tests compared Surprised and Baseline trials at each level of shape similarity, and found that infants performed significantly more target actions in the Surprised trials at all three levels of similarity. With the high similarity objects, infants performed significantly more target actions in the Surprised ($M = 2.94, SD = 2.91$) than in the Baseline trials ($M = 0.28, SD = 0.78$), $t(35) = 5.68, p < .001$. Similarly, with the medium similarity objects, they performed more actions in the Surprised ($M = 2.27, SD = 2.44$) than in the Baseline trials ($M = 0.50, SD = 1.75$), $t(35) = 4.02, p < .001$, and again, with the low similarity objects, infants performed more actions in the Surprised ($M = 2.06, SD = 2.24$) than in the Baseline trials ($M = 0.17, SD = 0.61$), $t(35) = 5.68, p < .001$. These analyses indicate that the target properties were indeed nonobvious, and infants did not perform the target actions unless they were demonstrated in the training period.

**Surprised trials**

The next set of analyses examined the effect of labeling and perceptual similarity within the Surprised trials only. A 3 (labeling group) x 3 (similarity) ANOVA found no effect of similarity, $F(2, 66) = 1.37, p = .26$, or labeling group, $F(2, 33) = 0.17, p = .84$, nor was there a significant labeling x similarity interaction, $F(4, 66) = 11.61, p = .18$ (see Figure 3).

We used planned comparisons to analyze the effect of labeling at each level of similarity. With the high similarity objects, there were no significant differences between the No Label ($M = 3.26, SD = 3.27$) and Noun groups ($M = 2.38, SD = 2.87$), $t(33) = 0.72, p = .47$, the No Label ($M = 3.26, SD = 3.27$) and Adjective groups ($M = 3.18, SD = 2.72$), $t(33) = 0.06, p = .95$ or the Noun ($M = 2.38, SD = 2.87$) and Adjective groups ($M = 3.18, SD = 2.72$), $t(33) = 0.66, p = .51$.

Again, with the medium similarity objects, we did not find significant differences between the No Label ($M = 3.08, SD = 2.97$) and Noun groups ($M = 2.06, SD = 2.26$), $t(51) = 1.03, p = .31$, the No Label ($M = 3.08, SD = 2.97$) and Adjective groups ($M = 1.65, SD = 1.97$),
Finally, comparisons with the low similarity objects revealed no significant differences between the No Label ($M = 1.08, SD = 1.31$) and Noun groups ($M = 2.17, SD = 1.64$), $t(51) = 1.22, p = .23$, or between the Noun ($M = 2.17, SD = 1.64$) and Adjective groups ($M = 2.92, SD = 3.12$), $t(51) = 0.85, p = .40$. However, there was a significant difference between the No Label ($M = 1.08, SD = 1.31$) and Adjective groups ($M = 2.92, SD = 3.12$), $t(51) = 2.07, p < .05$.

An analysis of the items (accordion, bell, castanet) used in this study did not reveal any significant item effects, $F(2, 27) = .183, p = .83$.

Cross-experiment comparisons

The results of Experiment 2 suggest a different pattern from Experiment 1. To test whether there were any statistical differences between structure-dependent functions in Experiment 1 and structure-independent functions in Experiment 2, we conducted an ANOVA using the data from Surprised trials, with type of function and labeling as a between-subjects factors, and similarity as a within-subjects factor. This analysis found no main effect of function, $F(1, 84) = 2.04, p = .16$, but found a significant Function x Similarity interaction, $F(2, 168) = 5.47, p < .01$. As we found in previous analyses, there was a significant effect of similarity with structure-dependent functions in Experiment 1, but not structure-independent functions in Experiment 2.

Discussion

These findings indicate that in Experiment 2, infants generalized structure-independent functions to other objects, but not in any systematic way. Infants did not appear to rely on perceptual similarity, count noun labels or adjective labels when generalizing the functions, and thus did not seem to generalize structure-independent functions on the basis of category membership. However, we did find that infants performed more actions on the low similarity
objects when they were labeled with an adjective than when they were not labeled. This finding suggests that adjective labels may have helped infants infer that the low similarity object could function in the same way as the training object. However, we need to exercise caution in interpreting these results because the number of participants in this study \((N = 36; 12 \text{ in each labeling condition})\) does not yet match Experiment 1 \((N = 54; 18 \text{ in each labeling condition})\). Continued testing to increase the number of participants in Experiment 2 is needed to obtain a clearer picture of the effects in this study.

This overall pattern of results in Experiment 2 is differs from Experiment 1, where infants learned structure-dependent functions of objects, and generalized them systematically to perceptually similar objects, or objects labeled with the same count noun. A cross-experiment comparison did not find that infants performed one type of function more than the other, allowing us to conclude that infants found both types of functions equally interesting. Thus, the differences between the two studies appear to be due to infants' sensitivity to structure-function relationships, and unlike Experiment 1, infants in Experiment 2 did not appear to link structure-independent functions with particular object kinds.

General Discussion

This research has focused on 18-month-old infants' reasoning about artifact kinds by examining the role of perceptual and labeling cues in infants' inductive inferences, as well as the kinds of functions that infants generalize to other category members. In Experiment 1, infants generalized structure-dependent functions to other objects that were perceived as members of the same category. When objects were not labeled or when they were labeled with an adjective, infants categorized objects on the basis of perceptual similarity and generalized the function to high and medium similarity objects. When objects were labeled with a count noun, infants categorized objects according to the count noun label, and generalized the function to high, medium, and low similarity objects. This study demonstrates that infants are able to use language
to guide their reasoning and infer that objects that are labeled with the same count noun are the same kind of thing. They also expect objects of the same kind to function in similar ways, despite perceptual differences.

In addition, these studies provide new evidence that 18-month-old infants are sensitive to structure-function relationships. In Experiment 1, infants generalized structure-dependent functions of the training objects systematically to other objects that were perceived as members of the same category. In contrast, infants in Experiment 2 were trained on structure-independent functions. Infants’ generalizations of these functions to the test objects were unrelated to labeling or perceptual similarity.

It is important to note that it was not the case that infants preferred one type of function to the other. A comparison of structure-dependent and structure-independent functions found no effect of type of function when the data from the two experiments were combined and collapsed over labeling groups and levels of similarity. This suggests that structure-dependent and structure-independent functions were equally interesting for infants.

One unexpected finding was the effect of similarity in the Baseline trials in Experiment 1, and lack thereof in Experiment 2. This finding suggests that structure-dependent functions are more likely to be discovered than structure-independent functions. One of the goals for these studies was to develop stimulus objects that were real artifact kinds with structure-dependent functions, rather than shapeless objects that one would be unlikely to encounter in the world. Perhaps a consequence of using artifacts with structure-dependent functions is that appearance can be a cue to the function, making the target actions more obvious than those in other studies. For example, the item with the highest number of target actions in the Baseline trials was the accordion. The accordions were constructed of two flat sides with protruding handles, connected by a flexible material in the middle. The structural configuration of this object may have caused the function to be more apparent to some infants. In contrast, infants in Experiment 2 very rarely
performed the structure-independent function without the initial training. Infants were more likely to perform structure-dependent functions than structure-independent functions in the Baseline trials, lending additional support to our hypothesis that infants are sensitive to structure-function relationships.

One caveat involves an important distinction between the structure-dependent and structure-independent functions used in this study. The structure-dependent functions involved the use of the stimulus object alone, and could be thought of as intrinsic functions of the objects. In contrast, the structure-independent functions involved a relation between two objects, where the target object is used to act upon another object. An alternative explanation for our results is that the differences between Experiments 1 and 2 were due to this relational factor, rather than the structure-function relationship. Structure-dependent functions involving a relation between two objects and structure-independent functions that are not relational seem to be much less common, perhaps due to the nature of these different kinds of functions. Further research is needed to clarify whether the relational aspect of the structure-independent functions in the present research was an important factor.

In the rest of the general discussion, each of the three questions identified at the beginning of this paper will be discussed: (a) do infants rely on perceptual information or shared labels when making inductive inferences? (b) are count noun labels privileged over words from other form classes for inductive inferences? and (c) what are infants’ patterns of inference when the demonstrated property is a structure-dependent function or a structure-independent function?

Perceptual similarity vs. Labeling

The first question asks whether 18-month-old infants rely on perceptual similarity or shared labels when determining category membership and making inductive inferences. Experiment 1 indicates that perceptual similarity can be a very important cue for infants when objects have not been labeled with a count noun. As predicted, infants perceived high and
medium similarity objects as members of the same category as the training object and expected these objects to function in the same way. The high and medium similarity objects were similar in shape to the training object, but differed from each other in color and material. The low similarity object was the only object that differed from the training object in shape. As such, shape seems to be the most relevant cue to category membership. Although Welder and Graham (2001) found that infants made a distinction between high and medium similarity objects, in their study, these objects differed in shape. Results from both studies suggest that shape is the most relevant perceptual cue that infants rely on when determining object kind. These findings are consistent with other studies that have demonstrated evidence of a “shape bias” – a bias to focus on shape over other perceptual cues when categorizing objects (Landau, Smith, & Jones, 1992; 1988). Moreover, we find that infants expect objects that share the same shape to share nonobvious properties as well, suggesting that shape is perceived as a reliable cue to kind membership. This expectation likely reflects infants’ experiences with objects in their environment, where objects of the same kind tend to be perceptually similar and function in similar ways.

18-month-old infants were also sensitive to count noun labels as a cue to object kind. By this age, infants interpret count noun labels as referring to an object’s category or kind (Waxman & Markow, 1998). This knowledge seems to be weighted more heavily than perceptual similarity. When objects were labeled with a count noun, infants performed similarly across the high, medium and low similarity objects. This finding demonstrates that infants perceived objects labeled with the same count noun as members of the same category, and expected them to function in the same way, despite perceptual differences. Thus, count noun labels appear to be perceived by 18-month-old infants as conceptual markers of object kind. Infants’ understanding of the semantics of count nouns may guide them in understanding that objects can be organized into categories, and enable them to draw inferences between category members.
Count nouns vs. adjective labels

The second question concerns whether infants interpret count noun labels and adjective labels in different ways. Results from Experiment 1 reveal a different pattern of responses between infants in the Count Noun group and Adjective group. Infants who heard objects labeled with a count noun performed similarly across the high, medium and low similarity objects, but those who heard objects labeled with an adjective performed fewer target actions on the low similarity objects. Much like the No Label group, infants in the Adjective group categorized objects on the basis of perceptual similarity, and did not expect low similarity objects to possess the same property as the training object. Thus, infants who heard the objects labeled with a novel adjective did not generalize the property based on the shared adjective. This finding suggests that infants did not perceive the adjective as labeling the category, or the demonstrated sound property of the object. These results are consistent with studies showing that by about 14 months, infants are able to distinguish count nouns from adjectives, but are not able to map novel adjectives to object properties until about 21 months (Waxman, 1999; Waxman & Markow, 1998).

These findings are also consistent with the linguistic differences between count nouns and adjectives. In general, there is a correspondence between different kinds of concepts and linguistic parts of speech. The basic function of a noun is to identify kinds of things that may share certain properties or features (Gentner, 1982). The use of an adjective, on the other hand, implies that one is referring to one property or feature among many that could be said about something. Adjectives can define an abstract class, such as red things, but they do not define a kind (Wierzbicka, 1986).

It has been proposed that nouns are special because they have a natural conceptual priority when it comes to word learning (Gentner, 1982; Macnamara, 1982; Markman & Hutchinson, 1984; Soja, Carey, & Spelke, 1991; Waxman & Markow, 1996). Infants growing up
in an English-speaking environment demonstrate a noun bias: infants acquire a disproportionate number of nouns in comparison to the proportion they hear spoken to them by their caregivers, and in comparison to the number of other kinds of words in the English language (Bates, Dale, & Thal, 1995). In addition, count nouns may be easier to learn because they often name cohesive whole objects, which can be relatively easy for the child to individuate and identify (Markman, 1990). In contrast, adjectives do not possess this conceptual priority, and as a result, can be more difficult to learn and are generally acquired later (Hall, Waxman, & Hurwitz, 1993).

The results from Experiment 1 show that by about 18 months, infants are sensitive to the syntactic differences between nouns and adjectives, and are able to differentiate nouns from other classes of words. They are also sensitive to the semantic differences between nouns and adjectives. Although infants are not yet able to map an adjective onto a property of an object, they understand that count noun labels refer to an object kind or category. An area of future exploration could examine whether adjectives guide infants’ inferences about functions that are related to the material of an object. For example, sandpaper is made of material that is related to its function. If adjectives are interpreted as referring to a material property (e.g. rough), infants may infer that other objects labeled with the same adjective would have the same function.

*Structure-function relationships*

The final question that we investigated concerns infants’ conceptions of artifact kinds and their functions. In Experiment 1, infants learned a structure-dependent function for each object and generalized these functions to other objects perceived as the same kind. In Experiment 2, infants were taught functions that were structure-independent, or unrelated to the structural configuration of the object. Infants did not generalize these functions systematically to other category members. Each experiment displayed a different pattern of results, suggesting that infants are sensitive to structure-function relationships in artifacts.
Structure-dependent functions and kinds. Why do infants’ patterns of inference differ with structure-dependent and structure-independent functions? We suggest that infants perceive structure-dependent functions as specific to particular kinds of objects, and they generalize these functions to other objects of the same kind. Prasada, Ferenz, and Haskell (2002) found similar differences between adults’ construals of structure-dependent and structure-independent functions. They suggest that an entity is construed as an object of some kind, rather than as an amount of solid stuff when a structure-dependent function has been demonstrated. For example, participants were shown a solid entity, such as a piece of plasterboard, and they viewed a demonstration of either a structure-dependent function or a structure-independent function. In the structure-dependent function, the piece of plasterboard fit into a slot in an additional apparatus to ring a bell. In the structure-independent function, the bell was rung by hitting it with the plasterboard. Participants were then asked to choose between a sentence that described the entity as an object using a novel count noun (e.g. There is a blicket in the tray) or a sentence that described the entity as an amount of stuff using a mass noun (e.g. There is a piece of zav in the tray). Participants who saw the structure-dependent function preferred the object / count noun description. In this way, the existence of a structure-dependent function provides a reason for thinking that the structure of the entity is nonarbitrary and bias participants toward construing the entity as an object. In the studies presented here, infants who were trained on a structure-dependent function may have been able to discern the relationship between the configuration of the various parts of the object and the way in which the object was used. Infants may have construed structure-dependent functions as a central property of a particular kind of object, and expected other objects of the same kind to share the same function. Structure-independent functions, however, may have been perceived quite differently. When infants observed the experimenter demonstrating the structure-independent function, they would not have been able to discern any important links between the structure of the object and the way in which the object
was being used. Because structure-independent functions can be performed by many other kinds of objects, infants may not have perceived these functions as specific to a particular kind of object, and therefore did not generalize these functions on the basis of perceptual similarity or a count noun label.

A number of other studies have also examined the links between categorization, naming, and object functions. Booth & Waxman (2002) provide evidence that infants as young as 14 months are able to categorize novel objects according to a novel noun label when the objects' functions are demonstrated. Object functions in this study were structure-dependent – for example, one set of objects had a loop that was used for hanging the object on a hook. When the object function was demonstrated, infants were more likely to attend to the relevant perceptual features, such as shape or parts, to categorize new objects. A count noun label provided an additional cue to the object's kind, and in this way, may have suggested the presence of a structure-dependent function. Booth and Waxman suggest that functions enhance infants' categorization by providing the noun with core meaning.

Along similar lines, Kemler Nelson, Russell, Duke & Jones (2000) found that two-year-olds generalize object labels based on information about an object's function. In their task, children were shown a novel object that was named (e.g. gidget), and the object's function was demonstrated. When presented with two new objects and asked to give the experimenter the gidget, children selected objects that possessed the same function but were not necessarily perceptually similar. They suggest that young children categorize objects according to function when the objects have transparent and compelling structure-function relations. They posit that the structure-function relation is compelling for young children when (a) aspects of the structure that enable the function are easily perceptible, (b) how the structure enables the function is easily understood, and (c) the relations between function and appearance are convincing. They show
that young children are able to override global perceptual similarity if they understand the structure-function relation and attend to structural properties that are functionally relevant.

In the current studies, the appearance of our stimulus objects provided some clues to the structure-dependent function. The functionally relevant structural features were visible, and may have appeared even more relevant after the function was demonstrated. In contrast, the appearance of the stimulus objects had little to do with structure-independent functions. While the objects could perform the function, functional affordance may not provide enough information for object categorization. Similarly, Diesendruck, Markson, and Bloom (2003) found that 3-year-olds are able to form function-based categories if they are told that an object is “made for” the demonstrated function, rather than if they are told the object “can do” the function. For example, children were shown a hollow cylinder made of cork, and were told that it is a wug. They were either told that it is made for holding coins, or that it can hold coins. Children were then presented with two new objects: a perceptually similar solid cylinder that cannot hold coins, and a wooden rectangular box that can hold coins. Children were then asked which object was also a wug. Children who were told that the object was made for its function preferred the object with the matching function, rather than the object with the matching shape. These findings suggest that preschoolers understand that intentional functions are relevant to an artifact kind, while functions that are described as incidental to an object are not construed as relevant to the artifact kind. Similarly, infants may be sensitive to the structural differences between functions that an object is made for versus a function that an object can do.

Conclusions

The present experiments complement and extend the body of knowledge about inductive inference in infancy by demonstrating that infants prioritize count noun labels, but not adjective labels, over perceptual similarity, when making inductive inferences. These studies also explore a new direction in studying infants’ understanding of object functions by comparing patterns of
inference with structure-dependent and structure-independent functions. The current results point to several lines of inquiry. Further testing of infants in Experiment 2 is needed to obtain an equal number of participants in both experiments, particularly to clarify the role of adjective labels in inferences about structure-independent functions. Future research is also needed to test structure-dependent functions that are relational, in order to investigate the importance of relational and non-relational object functions. In addition, future studies may also explore whether words from other form classes are useful in different kinds of inferences. For example, we may examine whether adjectives are informative when describing material related functions, or whether infants are able to make inferences about object functions that are described by verbs. Finally, it remains an open question how infants in these studies were able to translate category representations into performing actions on an object. It appears that infants were able form a category or kind representation, and infer that new category members would share the same functions, but we do not yet understand the mechanisms by which infants make this inference. Research into these issues will shed light on the precise mechanisms involved in inductive inferences in infancy.
References


Footnotes

1 The items analysis also revealed a significant interaction between item and similarity, $F(4, 90) = 23.96, p < .001$. This finding suggests that there may have been a stronger effect of similarity for some items.

2 Although adults' similarity ratings of the high, medium, and low similarity objects did not group the high and medium similarity objects together, we did find a smaller difference in ratings between the high and medium similarity objects than between the medium and low similarity objects. One might speculate that if adults had observed the functions being performed, they may have perceived a greater degree of similarity between the high and the medium similarity objects, like the infants.
Table 1

*Means and Standard Deviations of Target Actions on Test Objects*

<table>
<thead>
<tr>
<th></th>
<th>No Label</th>
<th>Noun</th>
<th>Adjective</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Experiment 1 – Structure-Dependent Functions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>2.97</td>
<td>2.41</td>
<td>2.76</td>
</tr>
<tr>
<td>Medium</td>
<td>3.74</td>
<td>2.31</td>
<td>3.72</td>
</tr>
<tr>
<td>Low</td>
<td>1.17</td>
<td>1.98</td>
<td>2.29</td>
</tr>
<tr>
<td><strong>Experiment 2 – Structure-Independent Functions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>3.26</td>
<td>3.27</td>
<td>2.38</td>
</tr>
<tr>
<td>Medium</td>
<td>3.08</td>
<td>2.97</td>
<td>2.06</td>
</tr>
<tr>
<td>Low</td>
<td>1.08</td>
<td>1.31</td>
<td>2.17</td>
</tr>
</tbody>
</table>
Figure 1. Stimulus object sets used in Experiments 1 and 2.

<table>
<thead>
<tr>
<th>Training</th>
<th>High</th>
<th>Medium</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accordions</td>
<td><img src="image1" alt="Accordions Training" /></td>
<td><img src="image2" alt="Accordions High" /></td>
<td><img src="image3" alt="Accordions Medium" /></td>
</tr>
<tr>
<td>Bells</td>
<td><img src="image5" alt="Bells Training" /></td>
<td><img src="image6" alt="Bells High" /></td>
<td><img src="image7" alt="Bells Medium" /></td>
</tr>
<tr>
<td>Castanets</td>
<td><img src="image9" alt="Castanets Training" /></td>
<td><img src="image10" alt="Castanets High" /></td>
<td><img src="image11" alt="Castanets Medium" /></td>
</tr>
</tbody>
</table>
Figure 2. Mean number of target actions performed on the test objects in the Surprised Trials in Experiment 1 (Structure-dependent functions). Vertical lines depict standard errors of the means.
Figure 3. Mean number of target actions performed on the test objects in the Surprised Trials in Experiment 2 (Structure-independent functions). Vertical lines depict standard errors of the means.