UNIVERSITY OF CALGARY

What Promotes Inductive Inferences in Infancy? Generalizing Knowledge About Nonobvious Object Properties

by

Andrea Natasha Welder

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

PROGRAM IN CLINICAL PSYCHOLOGY

DEPARTMENT OF PSYCHOLOGY

CALGARY, ALBERTA

JULY, 1999

0-612-48053-4

Abstract

The role of shape and labels in promoting inductive inferences in infancy was examined in two experiments. In both experiments, infants were presented with novel prototype and test objects in three within-subjects conditions: (a) both prototype and test objects possessed a target nonobvious property (predicted), (b) neither prototype nor test objects possessed the nonobvious property (interest control), and (c) only prototype objects possessed the nonobvious property (surprised). Test objects were high, medium, and low in shape similarity to a prototype. In Experiment 1, infants were either taught labels for prototype and test objects, or were not taught labels for the objects. In Experiment 2, the experimenter introduced objects' labels twice as often as in Experiment 1. As expected, infants made more inductive inferences when test objects were high in shape similarity to a prototype (Experiments 1 and 2), and when objects' labels were salient (Experiment 2).

Acknowledgements

First, I would like to thank my supervisor, Dr. Susan Graham, for her support, guidance, and mentoring during the past two years, and her tireless commitment to helping me hone my research skills. Her enthusiasm and dedication to this thesis and our research in general been an inspiration. She has taught me how to continue to challenge myself, and has offered me opportunities that have helped me grow both as a scientist and a clinical practitioner. I would also like to thank my thesis committee, Dr. Christopher Sears, Dr. Suzanne Hala, and Dr. Susan Rvachew for generously giving up their time to take part in my thesis process, and providing me with thoughtful feedback. In particular, I would like to thank Dr. Christopher Sears and Dr. Suzanne Hala for their research assistance and expertise and their support and encouragement, throughout the year.

I would like to thank the Natural Sciences and Engineering Council of Canada and the Alberta Heritage Foundation for Medical Research for their financial support during the completion of this thesis. I would also like to thank the parents and children who generously volunteered to take part in this research, along with the staff of the Calgary Regional Health Authority clinics in Midnapore, Ranchlands, Northill, and Thornhill, who were extremely giving of their time in handing out brochures. I would like to thank those in the Language and Cognitive Development Lab for their friendship, and their assistance in the preparation and presentation of this thesis: Alissa Pencer, Tamara Demke, Carlene Adamache, Jennifer Storms, and Joanna Card.

I would like to thank my family in Winfield: Marilyn and Lawrence Slivinski, my family in Cranbrook: Barbara and Gary Johnson, and my family in Kelowna: Irene and Jack Welder, for their continued love and support. A special thank you to Deborah Welder for her friendship, and opening up her home and heart to me. Finally, I would like to thank Effie Galway, my brothers and sisters, my extended family in Vancouver, and my colleagues from high school, S.F.U., and U. of C., for their endless support and friendship.

Dedication

This thesis is dedicated to my husband, Dean. Thank you for continuing to believe in me, for supporting me in my chosen career path, and for encouraging me through some difficult, uncertain, and lonely times. Thank you for being proud of me and for seeing what happiness the future holds.

List of Tables

List of Figures

What Promotes Inductive Inferences in Infancy?

Generalizing Knowledge About Nonobvious Object Properties

Introduction

Young children's ability to reason inductively using knowledge about object categories is a fundamental aspect of early conceptual development. Given the complexity of the task of induction, how do those with such limited experience in the world begin to acquire and utilize inductive strategies? The process begins with the development of concepts (Flavell, Miller, & Miller, 1993). According to Flavell et al., a concept can be defined as a "mental grouping of different entities into a single category on the basis of some underlying similarity--some way in which all the entities are alike, some common core that makes them all, in some sense, the 'same thing'" (p. 88). Concepts are not orthogonal entities, but are interconnected with many other concepts (Keil, 1989). They divide the world into useful categories, identifying areas of similarity in what would otherwise be unmanageable diversity (Flavell et al., 1993). A category has been defined by Flavell et al. as an extension of a concept, or all the instances of a concept. According to Shipley (1993), three psychological properties appear to characterize categories: (a) they have labels that are used to identify object members; (b) their members are believed to share a "deep" resemblance; and (c) they form the foundation for induction, as properties can be generalized on the basis of knowledge about categorical membership.

From Concepts and Categories to Induction

In simple terms, inductive reasoning involves: (a) the premise that some thing (e.g., an object property) is characteristic of a sample or population of things (e.g., balls), and (b) the conclusion that the same thing is characteristic, or holds true, across the entire (or most) of that class of things (e.g., all balls) (Moore & Parker, 1989). That is, inductive reasoning is based on the premise that things that are true for one member of a category will hold true for other members of the same category. These things may include nonobvious kinds of properties such as internal structure, chemical structure, behavior, and function (Gelman, 1988). For example, a child may begin with the specific premise that "Every ball I have seen has bounced" (characteristic of the sample), and generalize the property by concluding, "Therefore, all balls bounce" (characteristic of the entire population of balls). In inductive reasoning, then, the conclusion is constructed by the individual. Gaps in knowledge about the world motivate individuals to rely on the utilization of inductive reasoning strategies (Rips, 1975).

In contrast, in deductive reasoning, the conclusions one draws follow absolutely from the premises (Moore & Parker, 1989). For example, a child may begin with the premise that "All balls bounce," and conclude that "This is a ball, therefore it bounces." Thus, when using deductive reasoning, the conclusions one draws are based on knowledge about "all" of a population of something rather than just "some" of a population.

Given that young children have such limited knowledge about the world, the development of inductive reasoning is critical. Indeed, an ability to generalize knowledge to new instances and new situations is a fundamental aspect of adaptive intelligence (Shepard, 1987). An inductive ability is at least one of the mechanisms responsible for generative human thought and reasoning (Baldwin, Markman, & Melartin, 1993; Gelman, 1988), allowing humans to make predictions, set forth assumptions, and extend beliefs beyond the range of direct experience (Lopez, Gelman, Gutheil, & Smith, 1992). Inductive inference, therefore, allows for cognitive efficiency. As noted by Lopez et al., without an inductive ability, children would have to learn about every object in the world anew, unable to benefit from past experiences.

As the ability to reason inductively is such a fundamental aspect of cognition, three questions seem most pertinent to the study of early inductive development: (a) When and under what circumstances will children use conceptual information, as opposed to perceptual information, in making inductive inferences?; (b) Under what conditions will children constrain, or avoid making inductions that are too broad or meaningless?; and (c) When do children begin to acquire an inductive capacity? In the following sections, these questions will be examined with particular reference to the relevant theoretical proposals and empirical research.

Perceptual versus Conceptual Categorization

A prominent controversy in the conceptual development literature centers around how young children's categories are organized. Some researchers have argued that infants' earliest categories are perceptual, or appearance-based, and through maturity and experience in the world children eventually come to categorize on a conceptual level (e.g., Quinn & Eimas, 1986). Others have hypothesized that infants are able to represent, and therefore categorize, objects both perceptually and conceptually early on in infancy (e.g., Keil, 1991; Mandler, 1998). The question as to when and whether there is an actual perceptual-to-conceptual shift in representation and categorization has been labelled the "perceptual-conceptual debate" (e.g., Madole & Oakes, 1999). The perceptual-conceptual debate centers specifically around the issue of whether infants begin life with only perceptual representations of objects and move to conceptual representations when they have a more mature linguistic ability (i.e., preschool age), or whether conceptual representations begin early in infancy.

According to Madole and Oakes (1999), many influential theories tied to conceptual development have presupposed a dichotomy between perceptual and conceptual modes of categorization (for reviews see Madole & Oakes, 1999; Mandler, 1998). Whereas perceptual categories are based on what objects or events look like and how they are recognized (e.g., color, shape, and texture), conceptual categories (concepts) are thought to encompass underlying meaning, or essences of objects or events (e.g., unobservable properties such as internal structure or function). Mandler (1998) suggests that there are at least five main differences between perceptual and conceptual categories: (a) the type of information they encapsulate (appearances vs. meaning), (b) whether the information is accessible (she argues that perceptual ones are not), (c) the amount of information they contain (she suggests that perceptual categories are richer in the early stages of infant development), (d) how they are acquired (whether there is a perceptual-to-conceptual shift), and (e) their functions (e.g., she argues that perceptual categories are used to recognize objects and events, whereas conceptual categories are an important source of object meaning and may have greater inductive potential).

Mandler (1988, 1992, 1998) proposes that a qualitative perceptual-to-conceptual shift in representation does occur, but early on in infants' first year of life. She argues that the earliest representations in infancy are perceptual in nature, as stable perceptual displays first must be perceived in order for infants to analyze them on a conceptual level. However, once infants can form images, she argues, there is nothing to prevent them from engaging in conceptual representation. Mandler suggests that the conceptual system develops in parallel with the sensorimotor (perceptual) system, and that the conceptual system is not created, or derived from the perceptual system. Once infants are able to partake in the conceptual world, Mandler proposes, a separate, innate perceptually-based mechanism analyzes perceptual displays, producing meanings, or conceptual representations for the infant. Both perceptual and conceptual representation, according to Mandler, "feed off" perceptual displays, displays which include visual, auditory, and kinesthetic information. She suggests that this perceptual analysis mechanism allows meaning from percepts to be stored in an explicit, accessible, and interpretable conceptual representational format for the infant to use. Mandler indicates that both perceptual and conceptual information are important for the inductive process to occur. To begin with, perceptual similarity aids in determining whether things actually belong to the same conceptual category. Although objects which look alike are more likely to be the same kind of thing, if they are not perceived as such, generalizations from one thing to another thing will not occur (Mandler, 1998).

Keil (1989, 1991) argues that both perceptual and conceptual kinds of representation are in place from the very beginning of infancy. He has developed the term "doctrine of original sim," to describe the view that posits that before children develop concepts or theories about the world, they only make associations and generalizations on a perceptual basis. Keil disagrees with the proponents of this view (e.g., Flavell, 1985), and proposes that young infants possess a number of theoretical beliefs about the way the world works, along with isolated factual knowledge. One of Keil's main principles is that most concepts are actually partial theories, because they include explanations of relations among their parts and of their relations to other concepts. However, Keil argues that conceptual representations are not at the same level of sophistication during a child's development, as their depth, interconnectedness, and the frequency with which these representations are relied on all increase over time. In addition, concepts become more refined, and are capable of explaining a broader range of knowledge as a child matures (see Siegler, 1991).

Quinn and Eimas (1986) argue that perceptual and conceptual representations do not develop out of separate systems, and that there is no need for a special mechanism or qualitative shift to be involved in the ultimate formation of conceptual categories. Quinn and Eimas propose that infants' concepts form through the gradual accumulation of perceptual associations that eventually come to support conceptual representations, through maturity and experience. According to Eimas (1994), the development of conceptual representation is actually continuous in nature, and perceptually-based representations eventually merge into conceptually-based behavior. Eimas suggests that concepts are actually kinds of percepts, and that concept formation can be regarded as a perceptual enrichment of sorts (see Quinn & Eimas, 1986).

Although it seems likely that even young infants are capable of conceptual representation, it is difficult to determine whether, as Mandler (1988, 1992, 1998) argues, there is a distinct perceptual analysis transformation process that enables mature conceptual representation to form in parallel with the perceptual system, whether, as Keil (1989, 1991) proposes, conceptual representation is present from the beginning of development, or whether, as Quinn and Eimas (1986; see also Eimas, 1994) suggest, concepts eventually emerge from and merge with percepts through experience. Although the debate over the acquisition of conceptual representation and categorization continues, researchers have also begun investigating the conditions under which young children rely on either perceptual versus conceptual categories, particularly within the context of making inductive inferences about object properties.

Investigating Preschoolers' Inductive Abilities

For many years, it was not known whether preschool children had the ability to make inductive inferences conceptually (Carey, 1985). Widespread assumptions that children could not form mature scientific categories or look beyond obvious perceptual similarities were held in earlier research (e.g., Flavell, 1985; Tversky, 1985; see Keil, 1986, 1991; Mandler, 1998; Wellman & Gelman, 1988 for reviews). Researchers had argued that children required a mature scientific understanding in order to appreciate the inductive power of categories (Carey, 1985; Gelman, 1988). During the past decade, however, a number of important studies have provided evidence that preschoolers are indeed capable of categorizing conceptually and making inductive inferences on a conceptual basis (e.g., Gelman & Markman, 1986, 1987). Instead of simply asking children to categorize objects whichever way they saw fit (i.e., such as by color, shape, theme--see Tversky, 1985), in recent experimental procedures children are provided with more explicit cues as to the type of categorization that is intended by the researcher (e.g., by object labels--see Gelman & Markman, 1986). We now know that preschoolers expect categories to capture more than obvious perceptual features, and that they will make inductions on the basis of conceptually-based information (e.g., Gelman & Coley, 1990).

How does one investigate the inductive abilities of preschoolers? Rips (1975) outlined three components required to empirically examine individuals' inductive reasoning strategies: (a) a set of exemplars, (b) some property which could be consistently possessed by the instances, and (c) an initial specification of those exemplars known to have the property. With these components in place, Rips indicated that the researcher can then require a participant to make some judgment about those exemplars not already known to have the property. Utilizing these minimal requirements, Rips reasoned that a further restriction can be added: the researcher can limit the study to cases in which the set is composed of natural kinds. Natural kinds are naturally-occurring substances found in the world, and can be either biological (e.g., animals, plants), or non-biological (e.g., chalk, sand). Natural kinds have insides (or "deeper") properties which are not necessarily suggested by their outward appearances (e.g., a bat looks like a bird but in fact is a mammal, so it does not lay eggs).

Given that many nonobvious conceptual properties of natural kinds are unknown to young children, researchers can examine whether children will generalize properties of natural kinds on the basis of perceptual appearances, or information provided about conceptual identity. If a property to be generalized is one already known to a child, inferences made may simply reflect previous knowledge, and are not based on reasoning inductively about something known to something unknown. As older children and adults have more knowledge regarding the properties possessed by natural kinds, induction experiments sometimes require that they extend made-up nonobvious properties from one natural kind to another (e.g., birds have fluviam inside). Rips (1975) indicated accepting these properties for the duration of the inductive task is sometimes difficult for older children and adults. Preschoolers, on the other hand, may be more amenable to accepting made-up facts about natural kinds, as their real-world knowledge is more limited, and make-believe often plays a large role in their everyday interactions.

Gelman and Markman have conducted a highly influential series of studies which provide evidence that 4-year-olds (Gelman & Markman, 1986) and 3-year-olds (Gelman & Markman, 1987) are able to use conceptual information (i.e., object labels) to make inductive inferences about unobservable object properties. Gelman and Markman (1986) viewed natural kinds as a good starting point for investigating young children's inductive capabilities, as natural kinds usually have rich correlated structures (e.g., an animal with fins should swim), but they are not necessarily identified by simple perceptual features (e.g., a dolphin looks like a fish but is actually a mammal). Gelman and Markman reasoned that object labels provide the necessary information as to object identity, or kind. They argued that if children generalized a nonobvious property from one object to another object which shared the same label rather than the same appearance, then children were using conceptual identity as a basis for their inductions.

In Gelman and Markman's (1986) study, children were taught a novel fact about two pictured natural kinds (e.g., a tropical fish and a dolphin). A third target natural kind picture was then presented (e.g., a shark), which looked like one of the paired pictures (e.g., the dolphin) but was given the same category label as the other paired picture (e.g., "fish"). The child's task was to decide whether the target picture (e.g., the shark) shared a specific unobservable property with either: (a) the picture of the same appearance (e.g., the dolphin), or (b) the picture with the same category label (e.g., the tropical fish). For example, children were asked, "See this fish? Does it breathe underwater like this fish, or pop above the water to breathe like this dolphin?" Thus, the authors pitted category membership against perceptual appearances to examine whether preschoolers would overlook salient perceptual similarity if category labels (e.g., "fish" and "dolphin") were taught to them. As expected, Gelman and Markman found that children did make inductive inferences on the basis of similar labels rather than similar appearances (i.e., children said that the shark stayed underwater to breathe like the tropical fish). When children were not provided with labels for any of the pictured natural kinds, they made inductions on the basis of perceptual appearances (i.e., children said that the shark popped its head above the water to breathe like the dolphin). Thus, it was only when children knew of the underlying shared identity of two objects that they used these labels as a basis for inductive generalization of an unobservable property.

In Study 2, Gelman and Markman (1986) provided evidence that participants were not simply using identical labels to extend the property (i.e., mimicking the experimenter's labels), but were actually using the label as conceptual information, for inductive purposes. The researchers employed the same methodology as Study 1, but when children were asked whether the third target picture shared the same property as the first or second picture, this third picture was introduced with a synonymous (but nonidentical) label (e.g., "baby dog" instead of "puppy," as it would have been labelled in one of the first two pictures). Study 2 results replicated those of Study 1, indicating that the children in Study 1 were not simply answering in accord with identical labels, but were considering the meaning of the labels. In Study 3, the researchers demonstrated that children will use category information for inferences only when it is relevant to the task. Children were told that a certain color of sticker "goes with" a labelled natural kind (e.g., "dog") and that another color goes with another labelled natural kind (e.g., "fox"). The children were then shown a third target picture (e.g., something that looked like the fox but was labelled a dog), and they were asked which color sticker went with the third pictured animal. In this task, children did not show any preference for picking the color that went with either the dog or the fox (i.e., they chose each color equally often, at chance level). Thus, it could be concluded that the children used labels only when the labels were important for the task of making inductive inferences about properties, and not in tasks where it was unclear as to why perceptual versus conceptual information would be more relevant as an inductive base. In a subsequent study, Gelman and Markman (1987) found that 3-year-old children also used shared labels as a basis for making inductive inferences about object properties. Interestingly, children made inductive inferences on a conceptual basis even when no label was provided, if they already had some knowledge or experience with the natural kind object in question.

Gelman and Coley (1990) extended Gelman and Markman's (1986, 1987) studies to examine inductive inferences in 2 1/2-year-old children. Gelman and Coley were specifically interested in clarifying the age at which young children will use labels (conceptual identity) as an inductive base, rather than deceptive perceptual appearances. As in Gelman and Markman's (1986) study, physical appearances and label information of animal pictures were dissociated. In Study 1, children were taught a property (e.g., "lives in a nest") for a target object (e.g., a bluebird) and were then shown a series of four other pictures. One of the test pictures shared the same appearance and animal kind as the target (e.g., another bluebird), one shared only the same animal kind as the target (e.g., a dodo bird), one shared only the same appearance as the target (e.g., a pterodactyl), and one shared neither the same animal kind nor the same appearance as the target (e.g., a stegosaurus). In the label condition, category labels were provided for children (e.g., "bird" for the bluebird and dodo bird and "dinosaur" for the pterodactyl and stegosaurus), and in the no label condition, no category labels were provided for children. Participants were asked if each of the test pictures shared the same property as the target (e.g., whether each lives in a nest, like the target bluebird). As expected, Gelman and Coley found that in the absence of conceptual information (labels), children relied on perceptual information for making inductions (e.g., they indicated that the bluebird and pterodactyl lived in a nest, but the dodo and stegosaurus did not). However, when labels were provided, children overlooked salient appearances and extended object properties on the basis of category labels significantly more often than chance (e.g., they said that the bluebird and dodo lived in a nest and the pterodactyl and stegosaurus did not). Thus, it was concluded that even 2 1/2-year-olds' inductions are not inflexibly rooted in appearances, and that they will use conceptual information (labels) for making inferences about unobservable object properties.

In a second study, Gelman and Coley (1990) added an adjective control condition to rule out the possibility that the children in Study 1 were simply responding on the basis of objects with identical labels, and that children did not understand that the labels provided them with the conceptual information necessary for making inductions. Instead of providing children with labels (names) for natural kind objects, children in Study 2 were taught transient adjectival properties for them. For example, the target bluebird was introduced as, "This is wide awake. It lives in a nest." Children were then told that the bluebird and dodo bird were also wide awake, but that the stegosaurus and pterodactyl were sleepy. Children then had to decide whether each of the four test objects shared the same property as the target (e.g., live in a nest). As expected, the researchers found that children's performance in the adjective control condition was similar to that in no label condition of Study 1: inferences were made on a perceptual basis (e.g., they extended the property to the bluebird and pterodactyl only). Thus, children did not generalize the unobservable property on the basis of shared transient adjectival properties, which are not an informative basis for inductive generalization.

In summary, Gelman and colleagues' research has examined whether children as young as 2 1/2 years of age will use object labels to make inductive inferences on a conceptual level. These studies indicate that when children do not have knowledge about the properties of particular natural kind objects, they will use perceptual similarity as an inductive base. If, however, they are provided with information as to the underlying identity of something, then they will use that information to guide their inductions. Importantly, children do not merely mimic their responses in accordance with similar labels. Instead, they understand what types of object labels represent, and use label information for inductive generalization only when it is relevant and informative (i.e., when labels describe object kind).

What Factors Affect Preschoolers' Ability to Make Specific Inductions?

Following the Gelman and Markman (1986, 1987) studies, Gelman (1988) investigated whether preschool children are able to constrain, or restrict their inferences, and under what conditions they are more likely to do so. Gelman reasoned that without constraints, inductive reasoning would be too powerful and create too many overgeneralizations of object properties. In a series of studies, Gelman asked 4-year-old children to extend a property from either a natural kind or an artifact to other natural kinds and other artifacts. For example, the experimenter would show the child a picture of a rabbit and say, "This rabbit likes to eat alfalfa." Children were then shown another picture and were asked, "See this dog? Do you think it likes to eat alfalfa, like this rabbit?" Gelman examined children's yes-no responses to see whether they extended a target unobservable property to other categories of objects. She found that preschool children were able to place some constraints on their inferences (i.e., would not extend a property to other objects from different domains, such as from a rabbit to a telephone). However, more overgeneralizations were made from one natural kind to another (e.g., from a rabbit to a dog), than from one artifact object to another (e.g., from a football to a baseball). Gelman found that two factors influenced preschoolers' inductive generalizations: category homogeneity (i.e., the similarity of the members of the category), and property generalizability (whether the property could be true of other members as well). She argued that any theory developed within the context of preschoolers' inductive abilities must include: (a) attention to category homogeneity and property generalizability, (b) attention to category labels, and (c) attention to preschoolers' domain-specific knowledge or beliefs. Gelman concluded that a consideration of children's increasing scientific expertise becomes crucial for investigating how they learn to organize knowledge and make inductive generalizations.

In light of the previous findings by Gelman and her colleagues, Davidson and Gelman (1990) examined how young children, with such limited scientific expertise, come to understand the role of category labels as a basis for inductive inference. They reasoned that there are at least two explanations, characterized as representing two extremes on a continuum. The first proposal is that children's understanding of the rich structure of categories may be built up individually for each category. The second proposal is that children may hold a more general expectation about language that is independent of specific experiences with a particular category. Davidson and Gelman examined whether children require experience with a given category before treating it as a basis for induction, or whether children hold a more general expectation that objects with the same label share deeper nonobvious properties. The authors presented 4-year-old children with unfamiliar pictured objects that were given either novel or familiar labels, to examine whether children would generalize unobservable properties to other objects with similar labels, or those with similar appearances. As expected, they found that children made significantly more inferences when labels matched, even when the labels conflicted with perceptual appearances. Interestingly, when the same pictures were given familiar labels, children drew inferences on the basis of category labels and appearances. However, the authors also found that children did not consistently draw inferences on the basis of the novel label. Novel labels were only utilized as an inductive base when perceptual appearances of test objects were not extremely different from a given target object, and therefore test and target objects could be construed as plausibly connected in some way. The authors concluded that children may have a more general assumption about language, one positing that objects with the same label share deeper similarities (even objects with which they have had no prior experience).

Davidson and Gelman's (1990) results suggest that in tasks such as those used by Gelman and Markman (1986, 1987), children generalized an unobservable property from a target to a conceptually similar test object on the basis that the appearances of these two objects were deemed similar enough to warrant them sharing the same label, or identity. For example, children may have accepted that a shark could breathe underwater like a tropical fish, even though the shark looked somewhat different in size and appearance from the tropical fish, as it still had features (e.g., general body shape, fins) which made it possible to have similar unobservable properties as the tropical fish. In a similar study, Shipley (1993) also found that children extended an unobservable property on the basis of shared labels when there was a moderate, rather than large, discrepancy between appearances of target and test objects. Thus, children may not have a completely general assumption about the power of language in conveying the deeper identity of novel objects. Other factors, such as object appearances, category homogeneity, and property generalizability (as proposed by Gelman, 1988) may also play an integral role in determining whether children will extend a nonobvious property from one object to another.

As children learn about the world, they also experience the difficult task of deciding which category of many should be the focus for induction. Traditionally, it was thought that preschoolers were not capable of multiple classification for the following reasons: (a) preschoolers were assumed to have a lack of knowledge/experience about the world, (b) preschoolers were assumed to have a one object-one kind representation (i.e., they could only think about things in one way), and (c) preschoolers were thought to be too limited in their knowledge about probabilities of occurrences to decide whether one way of classifying would be more informative than another in a real-world situation (Kalish & Gelman, 1992). In a series of studies, Kalish and Gelman (1992) investigated preschoolers' ability to classify in multiple ways, particularly in the context of categories that are important for different types of inductive inferences. For example, if a child sees someone kissing their pet German Shepherd, it would not be a good strategy for her to induce that all German Shepherds can be kissed. However, it would be acceptable for her to conclude that all pets can be kissed. Results of these studies indicated that children are able to classify artifact objects in multiple ways. Children judged material kind (e.g., wood, cotton) rather than object kind (e.g., pillow, chair) as appropriate for making generalizations of object fragility (whether something will break) and object texture (whether something is hard or soft). However, when material kind was pitted against object kind (e.g., children were asked whether a pictured wooden pillow was hard or soft) children made significantly more appropriate inferences (e.g., that the wooden pillow was hard) when they were told that the pillow was wooden (label condition), as compared to when they were not informed of the pillow's material kind (no label condition). Thus, Kalish and Gelman concluded that when young children are not able to classify in multiple ways, the reason may be that they simply do not have enough knowledge/experience with particular kinds of objects to know, for example, their constitution (especially from a picture).

Indeed, Déak and Bauer (1996) indicated that the two-dimensional line drawings utilized in many induction and categorization studies with children are typically perceptually impoverished, potentially biasing children's responses and affecting the outcome of tasks which involve complex decision making and reasoning. The authors indicated that three-dimensional object stimuli are perceptually more informative, making experimental tasks more representative of those types of tasks children are faced with in real-world situations. In a series of studies, Déak and Bauer found that when objects were labelled, 4-year-old children made significantly more inferences from one <u>line drawing</u> to another (e.g., two pictures of lightbulbs) than from one <u>object model</u> to another (e.g., two actual lightbulbs). In contrast, when objects were not labelled, children made significantly more inferences from one <u>object model</u> to another than from one <u>line drawing</u> to another. Déak and Bauer concluded that a lack of physical information present in the line drawings perhaps compelled children to attend more exclusively to the labels. They argued that the type of task, the context, children's knowledge, and combinations of these all affect children's categorization decisions and inductive inferences generated. Thus, it is important to consider the quality of the perceptual information present in particular stimuli, before drawing broad conclusions about children's ability or inability to make inductions in a particular way.

Déak and Bauer (1996) also discussed the problem of the "either or" mentality in the literature regarding the perceptual versus conceptual nature of children's categorizations and inductions. They argued that the assumption that perception and cognition can be separated is difficult to support, and that Gelman and colleagues' studies are based on an artificial distinction. Indeed, when children are provided with information about category membership, they are often likely to have information about perceptual similarity and typical form-function correspondences (McCarrell & Callanan, 1995). Déak and Bauer indicated that a dichotomous framework does not readily accommodate evidence that children use many kinds of information for categorization, and that instead of asking "when and how preschoolers overcome perceptual boundedness," researchers should focus on delineating the conditions under which young children of various ages use different kinds of information to make categorization decisions. Déak and Bauer further argued that the labels used by Gelman and colleagues to constitute and isolate "conceptual identity" actually contain perceptual information as well, alerting children to look for physical details that support inferences about nonobvious similarities. Although this may be true (e.g., see Davidson & Gelman, 1990), Gelman and colleagues' research has been highly influential in challenging the traditional notion that preschool children are inflexibly perceptually-bound. These studies have shown that even when one test object is more perceptually dissimilar to a target object than another test object, preschoolers will generalize an unobservable property from the target to the perceptually dissimilar object if they share the same label (e.g., see

Gelman & Markman, 1986). Thus, preschoolers rely on more than just physical details when making inductive inferences.

Conclusions: Preschoolers' Inductive Abilities

The preceding sections have covered a number of key studies in inductive inferences with preschoolers. To summarize, research to date indicates that: (a) preschoolers will use perceptual similarity as an inductive base in situations where they have no knowledge of objects' shared underlying (conceptual) similarity (e.g., Gelman & Coley, 1990; Gelman & Markman, 1986, 1987; Kalish & Gelman, 1992); (b) preschoolers will use conceptual information (i.e., shared object labels) as the basis for their inductive inferences, even when objects' appearances are dissimilar (e.g., Gelman & Coley, 1990; Gelman & Markman, 1986, 1987; Kalish & Gelman, 1992), but still plausibly related in some way (e.g., Davidson & Gelman, 1990); (c) preschoolers are able to classify and make inductions about objects in multiple ways, but they require knowledge about and experience with unobservable properties in order to do so (e.g., Gelman, 1988; Kalish & Gelman, 1992); (d) preschoolers are able to limit, or constrain the inferences they make, depending on factors such as knowledge about a particular category, property generalizability, and category homogeneity (e.g., Gelman, 1988); and (e) label information appears to be more salient and utilized more often for generalizing object properties when stimuli are pictured line drawings, rather than three-dimensional objects (e.g., Déak & Bauer, 1996).

Research over the past decade has certainly increased awareness and knowledge about the inductive abilities of preschoolers. However, there are still gaps in the context of understanding when and how young children develop and utilize specific inductive strategies in everyday life. As Déak and Bauer (1996) pointed out, researchers now need to focus on examining the differential circumstances in which children of various ages and levels of cognitive development draw upon certain kinds of information and experiences to generate inductive inferences. A variety of tasks and types of stimuli (e.g., three-dimensional objects) should be used in future inductive experiments to provide more support for current research findings and interpretations. In addition, researchers need to investigate when young children first begin to use inductive strategies, and the kinds of perceptual and conceptual knowledge that infants rely on for answering questions of an inductive nature.

Investigating Inductive Abilities In Infancy

Although a number of induction experiments have been conducted over the past decade with preschool-age children, only recently have researchers begun to examine inductive strategies in infancy. According to Mandler (1998), it is surprising, given the amount of theoretical speculation in the literature, that there is a lack of research regarding when infants first begin to make inductive generalizations, or on what bases they make these inductions. However, recent research has shed light on infants' concepts and categorization abilities. For example, Mandler and McDonough (1998a) have shown that even before twelve months of age, infants have already formed concepts of animals, vehicles, plants, and even kitchen utensils and furniture (and are able to distinguish among them!).

Given the recency of Mandler and McDonough's research on infants' categorization abilities, there is still much to investigate in terms of investigating the inductive potential of various kinds of perceptual and conceptual categories in infancy. Mandler (1998) suggested that until recently, there has been no theory about how the first concepts are represented because it is easier to study early concepts by the words that newly verbal children use. However, Mandler argued that concepts and words are not equivalent, and that assessing concepts only through children's newly developed linguistic system will cloud our knowledge regarding what the earliest concepts are really like. Although induction studies with preschoolers have all required verbal responses as evidence for inductions, nonverbal techniques are used with infants, whose verbal abilities (if any) may be unreliable.

How does one examine the inductive abilities of infants, when their linguistic skills are limited? According to Oakes, Madole, and Cohen (1991), categorization studies with young infants have traditionally relied on visual habituation methods, whereby category recognition is inferred by differential looking at novel and recognizable stimuli. However, more active exploratory play procedures can often be utilized in categorization tasks with older infants (i.e., at least six months of age) who are much better at controlling motor responses and manipulating objects in some way to indicate their categorization judgements. In the object examining procedure, for example, infants actively examine test stimuli for a certain period of time, and habituation (and thus categorization) is inferred by disinterest in stimulus play (e.g., see Mandler & McDonough, 1993, 1998a). Although object examining is a valid measure of infants' ability to discriminate among objects, evidence for inductive generalization cannot come through object examining. Object examining only indicates whether children treat objects as members of the same category--it does not provide information regarding how they generalize nonobvious properties. Imitation task methods, however, allow the researcher to investigate whether infants, when shown a particular object property on a target object, will generalize the property to other test objects (i.e., make inductive inferences).

In the imitation task procedure, a specific target action is modelled by the experimenter, and infants' imitative target actions on particular test objects are recorded.

The rationale behind this procedure is outlined in Mandler and McDonough (1998b). The use of the imitation task procedure is based on two facts about infant behavior: (a) infants imitate spontaneously and do not need instructions to do so, and (b) their imitations are determined by what they have understood from their observations. Indeed, Mandler and McDonough (1996) found that 14-month-old infants did not imitate everything the experimenter showed them, and they generally did not imitate actions thought to be incorrect (e.g., infants would imitate putting a cat to sleep but would not imitate putting a car to sleep). Thus, infant imitation can be a viable measure of inductive generalization in infancy.

Do Infants Make Specific Inductive Inferences?

To date, only a few studies have been conducted in the area of infants' inductive abilities (e.g., Baldwin et al., 1993; Mandler & McDonough, 1996, 1998b). Mandler and McDonough (1996) utilized the imitation task technique to explore 14-month-olds' inductive generalizations at the superordinate level. The authors examined whether infants would generalize only those actions appropriate to animals (e.g., eating) to other animals and those actions appropriate to vehicles (e.g., keying a vehicle) to other vehicles, even when appearances were pitted against perceptual similarity (e.g., an test airplane model was more perceptually similar to a target bird model than a test cat model). In Experiment 1, actions appropriate to either the animal or vehicle domain were modelled on a target animal or vehicle (e.g., a dog was shown drinking from a cup). Infants were then shown another exemplar from the same domain (e.g., a cat), as well as a distracter from the other domain (e.g., an airplane). Infants' actions were coded for imitation of the appropriate action on the test model. It was found that when the infants were shown appropriate actions on a model of an animal or vehicle, they generalized the appropriate action to other exemplars in the same domain, even when objects from the other domain were perceptually similar to the target (e.g., infants would imitate the cat drinking but not the airplane).

In Experiment 2, Mandler and McDonough (1996) confirmed that the infants had not learned the appropriate actions from previous experience with the stimuli; similar results were found with atypical animal and vehicle exemplars (e.g., aardvark and crane). In Experiment 3, appropriate (e.g., showing a bird drinking from a cup) as well as inappropriate actions (e.g., showing a car drinking from a cup) were modelled on animal and vehicle exemplars, and infants were tested 24 hours later in order to examine whether they would perform only appropriate actions on animal and vehicle models. Mandler and McDonough found that even when inappropriate actions were modelled on target objects, infants performed appropriate actions on test exemplars significantly more often than the inappropriate actions that they had been shown (e.g., infants would imitate the cat drinking from a cup, but not the car). Thus, the authors concluded that even without the use of labels, 14-month-old infants constrained cross-domain inductive generalizations by using their knowledge of conceptual categories of animals and vehicles. Importantly, because infants did not imitate inappropriate actions on test models, (even when they were perceptually similar to the target), they were not merely mimicking everything they were shown. Mandler and McDonough's findings suggest that infants use imitation purposely, with the intent of expressing knowledge about the world.

Following from their previous work, Mandler and McDonough (1998b) examined whether 14-month-old infants would distinguish between generalizing actions appropriate for models from only one domain (either animal or vehicle), and actions that were potentially appropriate for models from both domains (e.g., only a horse can be put to sleep but both a car and a horse can be washed) (Experiment 1). Thus, the authors were interested in whether infants would generalize domain specific actions versus domain neutral ones. In particular, they examined whether infants would choose a test exemplar from the same domain for all domain neutral action imitations, or, as they expected, infants would also imitate the neutral actions on the test exemplar from the other domain. As predicted, Mandler and McDonough found that infants limited inductive imitations to specific domain-appropriate models, but generalized imitations of neutral actions to models from both domains (e.g., they would put the horse to sleep but not the car, and would wash both the horse and the car).

Mandler and McDonough (1998b) also conducted a series of follow-up experiments in which they examined 14-month-old infants' ability to generalize: (a) domain specific properties to animal and vehicle objects from the same basic-level category (e.g., dog to cat), (b) domain specific properties from land to air objects (e.g., dog to bird), (c) domain specific properties from familiar to unfamiliar objects (e.g., dog to anteater), and (d) specific actions from one basic-level category to (incorrectly) another (e.g., dog chewing a bone to goose chewing a bone). The authors' general purpose was to determine whether infants first begin with more global-level discriminations and generalizations early in life, and throughout their development start to make more basic level inferences as they gain experience and knowledge about the world. Mandler and McDonough found that infants were much better at limiting generalizations at the global level (e.g., they only imitated actions appropriate to animals vs. vehicles), rather than the basic level (e.g., they incorrectly generalized actions only appropriate to a dog on a goose). However, older infants (20 months of age) showed a greater ability to make finer basic-level distinctions.

Although Mandler and McDonough's (1996, 1998b) research on inductive inferences in infancy has focused primarily on the question of domain (animal vs. vehicle) and level (basic vs. global) specificity of infants' generalizations, Baldwin et al. (1993) focused on the question of whether infants will form general expectations about object properties, discernible through the observation of variations in the way objects are explored/manipulated. Baldwin et al. argued that researchers should examine infants' ability to draw inferences about nonobvious properties, because these properties provide the clearest case for the involvement of inductive capabilities, because they cannot be directly experienced and must be inferred from surface features (see also Rips, 1975). Thus, Mandler and McDonough's induction research with infants differs from Baldwin et al.'s (1993) in that Mandler and McDonough's studies were based on the assumption that children came into the testing session with a certain level of knowledge about object categories, and the researchers were interested in examining what that knowledge was and how it affected infants' inferences. In contrast, Baldwin et al. were interested in investigating whether infants generated specific expectations about nonobvious object properties, based on knowledge gained and expectations formed solely during the experimental situation.

In Baldwin et al. (1993), 9- to-16-month-old infants were shown a series of six objects in three within-subject conditions. In the violated condition, infants were first given an target object which possessed an interesting property (e.g., a "wailing" can), and were then given another similar-looking test object that was disabled so it did not possess the interesting property (e.g., did not "wail"). This condition was intended to provide a strong case for the argument that infants developed expectations about the properties of particular objects. Infants were not aware of the fact that test objects were disabled in the violated condition. Thus, if they persisted in trying to make the test object work, they would be demonstrating that they generalized the particular property from the target object. In the interest control condition, neither the target nor the test object possessed the interesting property (e.g., neither produced a wailing sound). This condition was important as it provided a baseline of infants' exploratory actions. It could then be determined whether infants' target actions on test objects in the violated condition were based on expectations from experience with a given target, or whether the test objects in themselves suggested the property, without any previous knowledge or experience with the target required. If test objects' properties were obvious, then infants should perform a similar number of target actions on objects in the interest control condition and the violated condition. In the fulfilled condition, both target and test object possessed the target interesting property (e.g., both cans could "wail" when tipped). This condition was included so that infants would not develop expectations that every test object was disabled (as all objects in the violated expectation and interest control conditions were), and become disinterested or frustrated with the stimuli. Infants' expectations were measured using: (a) the frequency of target actions performed on test objects (e.g., the number of times they tipped the can to make it wail), and (b) the latency to the first target action performed (i.e., how quickly infants tried to reproduce the property on test objects). Baldwin et al. hypothesized that if infants generalized the property from target to test objects, then in the violated condition they would perform higher frequencies of target actions and have shorter latencies to the first target action performed than in the interest control condition.

In Experiment 1, Baldwin et al. (1993) presented infants with an object that possessed an interesting property, to explore for thirty seconds. While the target object was left in view, a similar-looking test object was presented for infants to explore for thirty seconds. There was a higher frequency of target actions performed and a shorter latency time to the first target action performed on test objects in the violated condition than the interest control condition. It was concluded that infants developed expectations about test objects on the basis of just a thirty second experience with the target object. Data on the fulfilled condition were not provided or examined. Interestingly, the researchers found no main effect of age, no interaction between age and frequency of target actions performed on test objects, and no interaction between age and latency to the first target action performed on test objects.

In Experiment 2, Baldwin et al. (1993) examined infants' expectations about test objects that were perceptually different from a target. The purpose of the experiment was to ensure that infants' target actions were specific to particular test objects (i.e., the perceptually similar ones), and that infants were not just performing target actions on any kind of object stimuli. As expected, it was found that infants performed significantly fewer target actions on perceptually different test objects than perceptually similar test objects in the violated condition. In addition, infants' latencies were longer for perceptually different test objects than perceptually similar ones in the violated condition. In fact, frequency and latency scores for perceptually different test objects in the violated condition were similar to those scores for perceptually similar objects in the interest control condition, suggesting that shared perceptual similarity is an important basis for infants' generalizations of nonobvious properties.

Conclusions: Infants' Inductive Abilities

Taken together, data from Mandler and McDonough (1996, 1998b) and Baldwin et al. (1993) indicate that: (a) infants are capable of demonstrating specific expectations about object properties; (b) infants will constrain their generalizations when they deem it appropriate to do so, indicating that their imitative actions are purposeful; (c) infants use perceptual similarity as a basis for inductive generalization when there is no conceptual information provided, and will use their knowledge of conceptual similarity when it matters to the inductive task; and (d) earlier in development, infants seem to base their inductions on more global level categorical distinctions, gradually making inductions on the basis of basic-level categorical distinctions with experience and maturity.

Although the few studies that have been conducted in the area of infants' inductive abilities are an important starting point, there is still a great deal to learn about the status of infants' inductive knowledge. Mandler (1998) concluded that what is understood at this time is that early conceptual categories are very different in kind from the perceptual ones that have traditionally been thought to underlie the beginnings of cognitive development. It is still not known whether infants can use conceptual information (e.g., object labels) to guide their inductions, or whether infants will differentially constrain their inferences depending on the type of task, stimuli, and object properties. Further research is needed to fully understand infants' inductive capabilities, and to examine whether specific strategies are utilized differentially by infants at various ages and stages of linguistic development.

Given that there are gaps in our understanding of the bases on which infants make inductive inferences about object properties, the present experiments were specifically designed to examine some of the perceptual and conceptual properties that promote inductive inferences in infancy. No research to date has examined how the degree of object shape similarity affects infants' inductive generalizations of object properties. Previous studies have only utilized test objects that are similar and those that are completely dissimilar in appearance to a target. Thus, it is unclear how much shape similarity is required for infants to perceive things as similar enough so that they will generalize a property from one object to another. In addition, no inductive development studies to date have used novel stimuli with infants. It is important to examine the kinds of inferences infants make when they have no previous experience with particular objects and their corresponding nonobvious properties. Finally, research examining whether children will use shared object labels to promote inductive inferences has only been conducted with preschoolers. It is still not known whether infants will extend a property on the basis of similar labels, particularly when a test object is of dissimilar appearance to a target.

The Present Experiments

In two experiments, infants' ability to make inductive inferences about nonobvious object properties was investigated. More specifically, the role of object labels and object appearances in promoting these inductions was examined. As demonstrated by Gelman and colleagues (e.g., Gelman & Markman, 1986, 1987), young children are able to use object labels as a clue to object identity (object kind), and, therefore, as a conceptual inductive base. However, Gelman and colleagues did not study the effect of using object labels in inductive tasks with children younger than 2 1/2. Moreover, Gelman and colleagues' studies did not examine the extent of children's inferences when three-dimensional objects are used (but see Déak & Bauer, 1996).

The question of whether shape similarity guides infant's inductive inferences about nonobvious properties of novel objects was also investigated in the present experiments. Previous research suggests that for 16- to 22-month-old infants, shape (but not color) is highly predictive of object category membership (Graham & Poulin-Dubois, in press). However, a gap exists in the research with regard to understanding the extent of infants' dependence on shape similarity for making inductive inferences about novel objects. To date, no studies have examined the amount of perceptual similarity required for infants to see objects as "the same thing" or "in the same category." Categorization/induction studies which have investigated the role of object physical appearance in infants' judgements have dichotomized object stimuli into those that are either perceptually similar or perceptually different (e.g., Baldwin et al., 1993; Mandler & McDonough, 1996). These studies have generally found that infants use perceptual similarity as an inductive base (e.g., Baldwin et al., 1993), but will also use their knowledge about object categories to make inductions on a conceptual level, even when two objects look practically identical in appearance (e.g., see Mandler & McDonough, 1996). Because infants typically have many kinds of information to draw from for making inductions, the present experiments were conducted to explore the extent to which infants will use object labels and object appearances to guide their inductive inferences about nonobvious properties of novel three-dimensional objects.

With the goal of investigating whether object shape and label similarity promote inductive inferences in infancy, it was necessary to utilize an imitation task procedure to examine infants' expectations about nonobvious object properties. More specifically, infants were shown target actions on prototype objects, and their imitation of these target actions on test objects of varying degrees of shape similarity to the prototypes were recorded. In addition, infants were either taught novel labels for prototype and test objects when objects were introduced (Experiments 1 and 2) or they were not provided with labels when objects were introduced (Experiment 1 only). This allowed us to determine whether infants imitated more target actions when test objects were labelled than when they were not. As the properties of the objects were intended to be nonobvious, the extent to which infants used their knowledge about a particular prototype to decide whether the test objects shared the same property would also be assessed.

In Experiment 1, infants between 16 and 21 months of age were shown a series of novel objects from three object sets ("bell" set, "ball" set, and "rattle" set). Each set was comprised of a prototype object and four test objects. One group of infants (label condition) was taught a novel label (e.g., "Look at this flum!") for a prototype object (e.g., a round object with "tail-like" extensions on either side). Another group of infants (no label condition) was not provided with a label to name the identity of the object (e.g., "Look at this one!"). Only one word (i.e., either the label "X" (label condition) or the word "one" (no label condition) distinguished the dialogue used to draw infants' attention towards the prototype. Both groups of infants were allowed to explore the prototype for 10 seconds. Afterwards, the prototype was left in sight but out of reach and either a high shape, medium shape, low shape, or dissimilar object was given to participants to explore for 20 seconds. This object was either introduced with the same label as the prototype (label condition), or with no label (no label condition). With the assumption that infants will imitate spontaneously and purposely (see Mandler & McDonough, 1998b), it was hypothesized that if infants perceived the test object as belonging to the same category as the target prototype, then they would make the induction that the specific property of the prototype would generalize to the test object as well. That is, infants would be expected to imitate a target action on a test object if they saw the test object as "the same kind of thing" as the prototype. If, however, infants did not see the test object as the same kind of thing as the prototype, then they would be less likely to attempt to imitate a particular target action on a test object.

Infants examined prototype and test objects from each of three object sets in three within-subjects expectation conditions (see Table 1). In the surprised condition, infants were presented with a target novel object which possessed an interesting property (e.g., "squeaked" when squeezed), and were then presented with either a high shape, medium shape, low shape, or dissimilar object, which was disabled so it did not possess the interesting property (e.g., did not "squeak" when squeezed). This condition was of particular interest, as it would provide evidence that infants developed expectations about test objects and extended a specific property from a prototype to a test object. Importantly, because test objects in the surprised condition were disabled, if infants persisted in attempting to make the test object work, they would be demonstrating the use of inductive reasoning. Infants would indicate, through their attempts, how much they believed the property should generalize from the prototype to the test object.

The second condition was the interest control condition. In this condition, neither the target nor the test object possessed the interesting property (e.g., neither produced a squeak sound). This condition was an essential baseline of infants' exploratory actions when there were no specific expectations made about an object. A comparison of infants'

Table 1

Description of Within-Subjects Expectation Conditions.

Expectation Condition	Prototype Functional?	Test Object Functional?
Surprised	Yes	No
Interest Control	No	No
Predicted	Yes	Yes

performance in the interest control condition to either the surprised or predicted condition (to be described below) would indicate whether the nonobvious properties of the objects were, in fact, nonobvious. That is, if infants attempted to make target actions on the test objects only when they had knowledge about the functionality of the prototype, then it can be concluded that the test objects, in themselves, do not suggest the property.

The third condition was the predicted condition, in which the target and test object both possessed the interesting property. This condition was necessary as infants could come to expect that all test objects were disabled (as all test objects in both the surprised and interest control condition were), and lack the desire to try any actions on test objects or become quickly bored or frustrated with the stimuli. In addition, the predicted condition would provide information regarding how infants respond to test objects that have reinforcing properties (i.e., those that function as expected).

Although the three within-subjects expectation conditions of the present experiments are similar to those in Baldwin et al. (1993), Baldwin et al. only used test objects of high and different perceptual similarity to a target, and did not investigate the effect of labelling the objects for participants. Moreover, Baldwin et al. argued that infants developed expectations based only on experience with the stimuli during the test procedure. However, some of the stimuli, though not identical in appearance to items the participants had seen before (e.g., stimuli such as cubes, dolls, castanets), may have been similar enough to other objects in infants' environment that previous expectations may have developed before the actual testing procedure began. Thus, it may be difficult to distinguish expectations infants developed about test objects' properties based on the testing session from those due to previous exposure. It is also not fully clear whether the particular properties of some of

Baldwin et al.'s stimuli were actually nonobvious, so that generalization of the property could be used as evidence for expectations generated only from the experimental conditions in the laboratory. For example, one set of stimuli involved a doll whose head separated from its body. From the appearance of the doll, and previous exposure to objects with parts, infants may have been cued to the possibility that the head and body would come apart. In the present experiments, however, great effort was made to ensure the properties of objects were nonobvious, so that they could not be accidentally discovered or known from previous exposure or experience. All experimental stimuli were created to be novel and distinct in appearance, possessing the nonobvious property of making an interesting sound.

Infants' expectations were examined using (a) the frequency of target actions performed on test objects, and (b) the latency to the first target action performed (i.e., how quickly infants tried to imitate the prototype's property on the test object). If infants were developing specific expectations about object properties, then they would evidence those expectations by performing particular actions on test objects that they had seen or performed themselves on prototype objects. More specifically, it was hypothesized that in comparison to the interest control condition, in the predicted condition infants would attempt more target actions on test objects. This comparison would provide important evidence regarding whether the objects themselves suggested the target action that was required to make the interesting sound. That is, it could be determined whether the properties were really nonobvious, as they were intended. It was also expected that in comparison to the interest control condition, in the surprised condition infants would attempt more target actions on test objects, and have shorter response latencies to the first target action performed on test objects.

No hypotheses were made with regard to comparing infants' performance in the surprised and predicted condition. It was unclear as to whether infants would perform more target actions on test objects in the predicted condition than the surprised condition because test objects performed as expected and the sound was interesting and independently reinforcing, or whether they would make less inferences because test objects performed as expected and the sound was quickly boring. As Baldwin et al. (1993) did not provide any information on their analogous fulfilled condition, hypotheses for infants' performance in the predicted condition could not come from previous research.

Another set of hypotheses focused on the role of shape similarity in promoting infants' inductive inferences. It was expected that infants would compare the shape similarity of a given test object to its prototype, in order to develop expectations and decide whether to generalize the nonobvious property of the prototype to a test object. Specifically, it was hypothesized that the higher the shape similarity between a test object and a prototype, the higher the frequency of target actions performed on test objects and the shorter the latency to the first target action performed on test objects. The dissimilar object in each set was included as a control. These objects were completely dissimilar in appearance compared to their corresponding prototypes in material (texture), as well as shape. In contrast, the material (texture) of prototype and high shape, medium shape, and low shape test objects in each object set were held constant, and were contrasted among the three object sets. The dissimilar object was necessary to ensure that infants' inductive generalizations were specific to objects that they perceived as belonging to the same category (i.e., the high, medium, and low shape objects), and to account for the possibility that participants were merely imitating the experimenter's actions on any object, regardless of whether an expectation was generated and the property was generalized from a prototype. It was expected that when infants were confronted with the dissimilar object, they would not imitate the target action performed on the prototype from the same toy set. More specifically, it was hypothesized that there would be significantly less target actions performed on dissimilar objects than high, medium, and low shape objects.

The final set of hypotheses centered around the role of object labels in promoting inductive inferences. It was expected that if infants used labels as an inductive base, then after being shown a nonobvious property of a novel labelled target object (label condition) they would expect test objects also introduced with the same label to share the nonobvious property (i.e., they would consider them to be of the same kind, or in the same category). In comparison to those in the no label condition, it was hypothesized that infants in the label condition would make significantly more inductive generalizations. Specifically, it was expected that: (a) there would be a higher frequency of target actions performed on test objects in the label condition and (b) there would be a shorter latency to the first target action performed on test objects in the label condition. However, if infants do not understand the role of linguistic labels, or if these labels do not promote inductive inferences, then infants in the label and no label condition would be expected to perform similarly in the inductive task. Based on previous research which indicates that children will generalize a property from one object to another when there is a shared conceptual identity, even when appearances are dissimilar (e.g., Gelman & Coley, 1990), a shape by condition interaction was also expected. That is, although infants may not perform as many target actions on medium or low shape test objects in the no label condition, it was predicted that in the label condition infants would use shared object labels to infer that even though the medium and low shape objects were not similar to their prototype in shape, they shared an underlying conceptual identity with the prototype (and thus would perform more target actions on these test objects).

Finally, the impact of infants' age and productive vocabulary size on their inductive abilities was also investigated. Research has shown that there is tremendous variability in 16- to 21-month-olds' productive vocabulary size (e.g., Poulin-Dubois, Graham, & Sippola, 1995). Specifically, in the present experiments it was predicted that infants with a larger productive vocabulary size would perform more target actions on test objects than those with a smaller productive vocabulary size when novel objects were labelled, as those with more words in their vocabularies would be perhaps more attuned to and knowledgeable about label information. In addition, age differences in inductive performance were investigated. Although Baldwin et al. (1993) did not find any significant age differences among four groups of infants: (9- to 10-month-olds, 11- to 12-month-olds, 13- to 14-month-olds, and 15- to 16-month-olds), the authors did report that older groups of infants performed more target actions on test objects. However, because the sample of infants in the present experiments were older than those in Baldwin et al., it was expected that perhaps there would be age differences in inductive performance, particularly when infants were taught labels for the stimuli.

To summarize, it was hypothesized that infants would perform significantly more target actions on test objects and have shorter latencies to the first target action performed on test objects when: (a) there was higher shape similarity between prototype and test objects, (b) prototype and test objects were labelled, (c) infants were older, and (d) infants had a larger productive vocabulary.

Experiment 1

Method

Participants

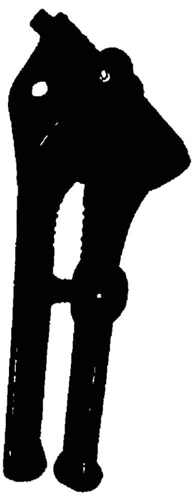
Sixty-four infants, 33 girls and 31 boys, (M = 18.83 months; SD = 1.42; range = 15.79 to 21.46 months) were tested in this experiment. Children were randomly assigned to one of two conditions: label (M = 18.84 months; SD = 1.51; range = 15.79 to 21.46 months), and no label (M = 18.81 months; SD = 1.33; range = 16.26 to 21.10 months). Seven additional infants were tested but were excluded from the sample due to fussiness (e.g., crying, trying to leave the room). Eight infants were excluded from the frequency and latency data analyses, as they did not perform target actions any of the high, medium, and low shape test objects in the surprised condition. Lack of any imitation on any of these test objects meant that children had no latency or frequency scores in the surprised condition. Because it could not be determined whether children who did not perform any imitation in the surprised condition were disinterested, did not understand what they were to do, or were perhaps shy, they were excluded from all the frequency and latency analyses involving the surprised, predicted, or interest control condition data. However, they were still included in the object transfer data analyses (described in data analyses section to follow).

Parents of participants were recruited through newspaper advertisements and brochures in four Calgary Regional Health Authority medical clinics. All parents gave informed consent and were debriefed after testing, in accordance with the University of Calgary Ethical Guidelines. The MacArthur Communicative Development Inventory: Toddlers Version (MCDI; Fenson et al., 1991) was filled out by parents to assess infants' productive vocabulary. Infants were presented with a small toy and certificate after testing, thanking them for their participation.

Materials

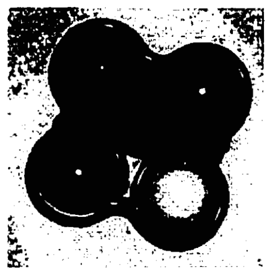
Stimuli. Three objects with interesting properties were used for the warm-up trials: a garlic press which opened and closed, a roller ball which could roll along the table, and a clicking clock with a knob which made a clicking sound when turned (see Figure 1). For test trials, three object sets consisting of five stimuli, each constructed to be novel, interesting, and visually distinctive from stimuli in other sets were used (see Figures 2 to 4). In each object set, one object was designated the prototype, and three other objects were

All stimuli within a given object set (except for the dissimilar object) were crafted from the same type of fabric, but they varied in color. In the first set ("ball" set), 7 cm hollow rubber balls were covered with pleated silky rayon and shaped in various ways with string and/or sponge. In the functional set (i.e., prototype in the surprised and predicted condition, test objects in the predicted condition) objects could squeak when they were squeezed. For the nonfunctional set (i.e., prototype in the interest control condition, test objects in the surprised and interest control condition), the air holes which made the squeaking noise possible were glued shut. In the second set ("bell" set), 7 cm bells were placed inside a Styrofoam shape and were covered with faux-fur material. In the functional set, objects could ring when they were tapped. For the nonfunctional set, the bells were taped closed to prevent vibration, and thus the objects could not ring. In the third set ("rattle" set), 7 cm (rattle portion) by 4 cm (handle portion) rattles were covered with felt and shaped in various ways with sponge, or by removing portions of the rattle frame. In the functional set, objects could produce a rattling noise when shaken. For the



Garlic Press

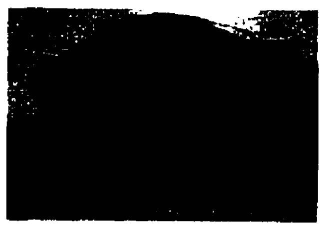
Clicking Clock



Roller Ball

Figure 1. Warm-up objects.

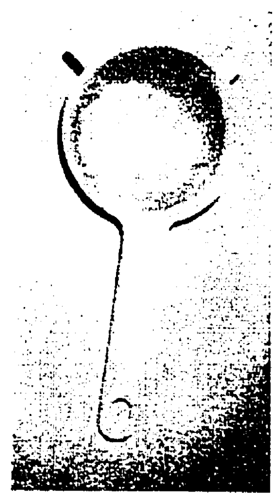
Prototype



High Shape

Medium Shape

Low Shape



Dissimilar Object

Figure 2. Bell set.

Prototype

High Shape

Medium Shape

Low Shape

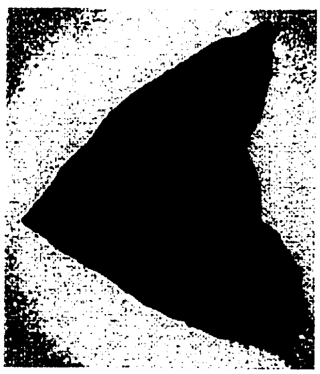
Dissimilar Object

Figure 3. Ball set.

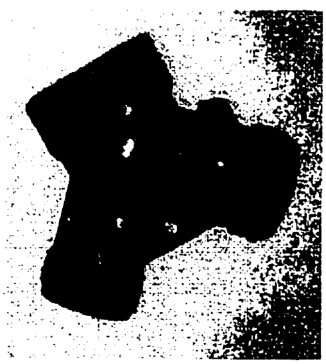
Prototype

High Shape

Medium Shape



Low Shape



Dissimilar Object

Figure 4. Rattle set.

nonfunctional set, the beads in the objects were removed, and thus the objects could not rattle. The dissimilar object in each set was: a plastic orange file (ball set), a small white strainer (bell set), and a plastic green hose splitter (rattle set). Parents were asked whether their infant had ever seen or experienced objects similar to any of the objects used in this experiment, to verify that the objects were indeed novel (none had).

Equipment. A videocamera capable of generating a time-stamped video record and a videocassette recorder equipped with frame by frame replay provided a visual record of infants' visual and tactile exploration of the stimuli. Both the videocamera and the videocassette recorder were placed behind a one-way mirror. A hidden microphone was placed on a shelf in the testing room, providing a continuous audiotaped record. Infants were seated in their parents' laps at a rectangular table across from the experimenter. The experimenter used a small stopwatch to time the 10- and 20-second intervals in which the stimuli were left on the table for infants to explore. Coders examined the video replays on a 27-inch television monitor, and used a stopwatch with a millisecond timer to record the latency to the first target action performed on each object.

Adult Ratings

It was important to ensure that, for each object set, test objects could reliably be categorized as being high shape, medium shape and low shape as compared to a target prototype, as intended. Thus, we asked adults to participate in a stimuli selection rating task. Fifteen University of Calgary adult volunteers (9 females and 6 males, all over age 18) participated. All were tested individually in a quiet room. They were shown nine pairs of stimuli, each pair consisting of a prototype and corresponding high, medium, or low similarity test object. Each rater was shown the pairs of objects in a different order from all other raters, with the stipulation that object pairs from each of the three object sets had to be presented one after another in the same order throughout the session. For example, if a participant was shown the rattle prototype and the high shape rattle test object first, the bell prototype and the medium shape bell object second, and the ball prototype and the low shape ball object third, he/she would be shown the remaining pairs of objects in the same order (i.e., rattle pair, bell pair, ball pair). However, the order of shape was presented randomly for each participant. That is, if a participant saw the high shape rattle object first, the experimenter randomly chose either the medium or low shape object for that object set the next time. Participants were asked to rate the shape similarity of each test object compared to the prototype, one pair at a time. They were instructed to focus solely on shape similarity, and that any other sources of similarity between a test object and prototype (i.e., color) should be ignored. Previous research (e.g., Graham & Poulin-Dubois, in press) indicates that infants do not use color as a meaningful way to categorize objects. Thus, it was not expected that infants would use color as an inductive base. Because adults might assume that the task involved a color comparison (i.e., demand characteristics of the task), it was decided that adults should be made aware that color was not important to their ratings. The rating scale ranged from 1 (not at all similar) to 7 (very similar) with the middle as 4 (somewhat similar). Participants did not physically handle any of the test stimuli.

Adults' ratings followed the expected pattern. That is, all high, medium, and low shape test objects in each object set were perceived as significantly different in shape from one another (all t-tests: $p < .05$), in the directionality intended. In other words, the high similarity test objects were rated significantly higher in shape similarity to their prototypes than the medium similarity objects, which were rated significantly higher in shape similarity to their prototypes than the low similarity objects (see Appendix A for mean ratings and t-tests).

Design and Procedure

Infants were brought into the laboratory and seated in their parent's lap. The experimenter sat across a table from the infant. Parents either filled out the MCDI before testing, if their infant required some time to become comfortable with the laboratory environment and the experimenter, or filled it out after testing or at home. All MCDI's that were completed at home were mailed back within a one-week period, except for two which were not returned (these infants' data were still included in all analyses except for the productive vocabulary analyses). Before testing began, parents were instructed to interact with their infant as little as possible, and not to mention, point to, or give objects to their infant during the testing session. However, they were told they could make "ooh-ahh" sounds and repeat their infant's name if he/she started to become fussy. Also, if infants dropped one of the objects on the floor near the parent, or if infants gave an object to parents, parents were instructed to silently put the object back on the table within their infant's reach.

In the warm-up phase, infants were shown three objects (the garlic press, roller ball, and clicking clock) with interesting properties. Infants were presented with the stimuli one at a time by the experimenter, who said in an excited voice, "Hi (infant's name)! Look what I'm doing! Look what I'm doing! Can you do this like me? Can you do this?" while the property was being demonstrated (e.g., the roller ball was rolled on the table) four times. Parents were then instructed to demonstrate the property (twice), and then present the object to their infant. Regardless of whether participants were able to imitate the experimenter's actions, all three warm-up trials were performed and then testing began. The purpose of the warm-up task was to make infants feel comfortable in the testing situation, practice imitating the experimenter's actions when appropriate, and understand that they are to give back the object in exchange for a new one when the experimenter gestured for the object and said, "Okay, let's try another one!".

In the testing phase, infants were presented with a prototype object from one of the three object sets. The experimenter made several excited remarks while demonstrating the prototype's property five times [e.g., "Hi (infant's name)! (shake rattle) Look at this one! (shake rattle) Look at this one! (shake rattle) See! (shake) Look at this one!" (shake rattle); no label condition]. The prototype was then passed to the parent who demonstrated the property twice (without saying anything), and then the parent presented the prototype to their infant. Participants were allowed to explore the prototype for 10 seconds, after which time it was placed closer to the experimenter, away from their reach but within their view. Then the experimenter presented participants with either the corresponding high shape, medium shape, low shape, or dissimilar test object from the same object set, and said, "Okay (infant's name)! Look at this one! Look at this one!" (no label condition). In the label condition, the label flum (ball set), zas (bell set), or blint (rattle set) was exchanged for the word one in the no label condition (e.g., "Look at this flum!"). Participants were allowed to explore a test object for 20 seconds, after which both the prototype and test object were placed back in the box beside the experimenter and a prototype and test object from another object set were demonstrated in the same fashion. If an object was dropped off the table or passed/thrown out of the infants' reach during the session, the object was quickly placed back within their reach. Time lost due to these actions was not compensated for, as they were considered to be intentional actions of frustration or disinterest (see Oakes et al., 1991). As the prototype from each object set was reintroduced to infants each time before a new test object was introduced, parents demonstrated the property only the first time (to help their infant feel comfortable performing the target action). The experimenter continued to demonstrate with the prototype throughout the testing phase. Only the properties of prototype objects in the surprised and predicted conditions were demonstrated (as there was nothing to demonstrate with the prototype objects in the interest control condition).

For each participant, one object set was designated the surprised condition, one set was designated interest control condition, and one set was designated the predicted condition (counterbalanced among participants), and these designations was maintained throughout testing. Thus, if an infant was shown the rattle set in the surprised condition, each time the rattle prototype was introduced it was functional (i.e., made a rattling sound when shaken), but corresponding rattle test objects were always disabled (i.e., they did not produce the rattling sound). However, if the rattle set was shown in the interest control condition for another infant, neither the rattle prototype nor any of the rattle test objects were functional. The prototype in the interest control condition was introduced with the same dialogue as in the surprised or predicted condition, but no property was demonstrated by either the experimenter or the parent. Each expectation condition occurred one after the other in the same order of presentation (e.g., surprised, interest control, predicted; surprised, interest control, predicted; etc.), four times. Each time, a new test object was introduced but the prototype was always the same. Order of presentation of expectation condition (e.g., whether the surprised, interest control, or predicted condition was designated first, second, or third), and order of presentation of object sets (e.g., whether the ball, bell, or rattle set were designated first, second, or third) were counterbalanced across infants. Presentation of the test objects (e.g., whether the high shape, medium shape, low shape, or dissimilar object was shown first, second, third, or fourth) was randomized among infants.

Coding

Coders were blind to the hypotheses of the experiment, and all coding was done with the volume on the monitor turned off. Thus, coders could not tell whether objects actually made sounds when target actions were performed by infants, and so the three expectation conditions could not be distinguished from one another. Coders recorded the frequency of instances a target action (i.e., rattle object was shaken, ball object was squeezed, bell object was tapped) was performed on the prototype and test objects, during their respective 10 and 20 second encounters. In addition, a latency measure was calculated for each prototype and test object through the use of the time-stamped video record (to determine the length of the 20-second interval) and an accompanying stopwatch with a millisecond timer (to determine the length of time elapsed from the moment the object was passed to the infant to the moment the infant performed a first target action). Coders also recorded any instances of "object transfer," where a target action appropriate to one object set was performed on an object from another object set (e.g., a bell object was squeezed or shaken instead of tapped). Object transfer was important to examine, in order to ensure that infants' imitative actions were specific to those objects within the same object set, and were not carry-over actions from other object sets.

Detecting target actions sometimes required coders to make complex decisions regarding subtle motor movement. The target action for the ball set was defined by a squeezing motion, in which the infant gripped and then compressed his/her fingers together on the object (not tapping the ball, hitting the ball on the table, shaking the ball, or gripping it to look at it or pass/throw it to the experimenter or parent). A release of the muscles of the fingers after squeezing the ball was not counted as a second target action, but every time the muscles contracted in a squeezing motion an action was counted. If the ball was squeezed with two hands together, one action was counted, unless the squeezing occurred at two separate points in time (i.e., one after the other), which was then counted separately.

The target action for the bell set was defined by a tapping, hitting, or patting motion, in which the infant made contact with the object (not squeezing the bell, hitting it on the table, shaking it, or gripping it to look at it or pass/throw it to the experimenter or parent). A downward motion making contact with the object was considered one action, but an upward motion to bring the hand or finger back from the object was not counted as a second action. Touching/stroking the bell gently with the hand or finger to poke it or feel its texture was not counted unless it was a swift "tapping" action. If the bell object was tapped with two hands, a target action was only counted once unless they occurred at two separate points in time (i.e., one after the other), which was then counted separately.

The target action for the rattle set was defined by a shaking motion with the wrist and/or whole arm in a back/forth or up/down motion (not tapping the rattle, squeezing it, hitting the table or a body part with it, or gripping it to look at it or pass/throw it to the experimenter or parent). An "up" shaking motion, for example, would be counted as one action and a "down" shaking motion as another action, only if there was a pause between them (i.e., the motion was not continuous or a rebound effect of moving the wrist/arm one way, but a true separate attempt to shake the object in another direction). If the infant performed a fluid shaking movement, then only one target action was counted. If the rattle object was shaken with two hands together, the same criteria as outlined above applied.

In order to establish inter-rater reliability, 20% of the data (n = 15 participants) was coded twice. Intraclass correlations (ICCs) were used to establish the level of agreement between the two coders. In this method, the pattern of agreement (i.e., whether rater A and B both rated one participant's frequency, latency, and object transfer scores as higher than another participant's scores), as well as the level of agreement (i.e., whether rater A and B were close in their exact scores for frequency, latency, and object transfer data) are considered (Sattler, 1992). Thus, the ICC coefficient considers the extent to which both raters meant exactly the same thing by their judgments, and is thus a more conservative measure of assessing inter-rater reliability than a traditional Pearson correlation. ICC coefficients for prototype and test object frequency ratings were both significant, $ICC\ (360) = 1.00$, $p < .001$, and $ICC\ (360) = .99$, $p < .001$, respectively. The ICC coefficient for test transfer frequency ratings was significant as well, $ICC\ (360) = 1.00$, $p < .001$. Thus, the two raters were in almost perfect agreement on all frequency ratings. ICC coefficients for prototype and test object latency ratings were also both significant, $ICC\ (360) = 1.00$, $p < .001$, and $ICC\ (360) = .96$, $p < .001$, respectively. Because of the amount of missing latency data coded by both raters, the ICC calculation may have been too liberal. Thus, it was recalculated excluding the trials in which both raters coded an absent latency score. This ICC coefficient was also significant for both prototype and test object latency ratings, $ICC\ (99) = .99$, $p < .001$, and $ICC\ (99) = .99$, $p < .001$, respectively.

Results

The data were analyzed in a number of steps. First, the data were screened for frequency of target action and latency to first target action outliers, which were eliminated from the final analyses. Second, the frequency of target action data for the prototype object and test object trials were analyzed. Third, the object transfer data were examined. Fourth, the latency to first target action data for test object trials were analyzed.

Data Screening

The frequency of target action and latency to first target action data were examined separately for the presence of outliers. Individuals with standard scores greater than 3 standard deviations above or below the mean were eliminated from the final analyses, under the assumption that there would be no loss of generalizability of the results (Tabachnick & Fidell, 1996). Inspection of the frequency of target action data for univariate outliers revealed seven cases with scores greater than 3 standard deviations above the mean. Thus, approximately 11% of the cases were eliminated from the final frequency analyses. Inspection of the latency to first target action data for univariate outliers revealed four cases with scores greater than 3 standard deviations above the mean. Thus, approximately 6% of the cases were eliminated from the final latency analyses (see Appendix B for a description of outliers).

The frequency of target action and latency to first target action data were analyzed individually. Thus, only frequency of target action outliers were eliminated from frequency analyses, and only latency to first target action outliers were eliminated from latency analyses. Although latency to first target action scores are dependent on the presence of at least one target action performed (frequency score), frequency and latency scores were analyzed separately, as each variable was considered an independently valid measure of infants' inferences. For example, regardless of whether a participant squeezes a ball set object three times or six times (frequency), the time to first attempt of the target action (latency) can vary between less than 1 second to a total of 20 seconds, and provides information regarding how quickly inductions were made, rather than the strength of infants expectations that test objects should perform in a particular way (which the frequency measure provides).

In the warm-up trials, only three participants failed to imitate target actions on at least one of the three objects. Because these participants did perform target actions on test objects in the experimental task, their data were not eliminated from any subsequent analyses.

Frequency of Target Action Data Analyses

Fifty-seven infants, 31 girls and 26 boys, (mean age = 18.77 months; SD = 1.47; range = 15.79 to 21.46 months) were considered in the analysis of the frequency data. There were 16 girls and 14 boys (mean age = 18.76 months; SD = 1.56; range = 15.79 to 21.46 months) in the label condition, and 15 girls and 12 boys (mean age = 18.77 months; SD = 1.39; range = 16.26 to 21.10 months) in the no label condition.

The aim of the first set of analyses was to establish that infants knew how the functional prototypes (i.e., in the surprised and predicted conditions) worked to produce a target sound. These analyses helped to establish that infants generalized properties from prototype objects to test objects on the basis of knowledge about a functional prototype's nonobvious property. Infants performed at least one target action on functional prototypes in 347 of the 456 total possible prototype trials in the surprised and predicted condition

(76% response rate). Thus, because participants actually tried the prototype and found out for themselves that it performed the interesting property on over three-quarters of the trials, those trials where target actions were not performed on the prototypes were not excluded. Arguably, infants knew about the nonobvious properties of prototypes, because the experimenter showed them. Thus, infants who may not have been interested in trying out the prototype were not excluded from the proceeding analyses, because they still could form valid expectations about test objects based on what they had observed from the experimenter's demonstration.

The second analysis was conducted to examine whether the target property (sound ability) of object stimuli were nonobvious to infants. It was expected that the appearance of the objects would not suggest the sound that objects could produce when a particular target action was performed. It was important to establish that the sound property was nonobvious, in order to conclude that infants were making inferences by utilizing information learned in the present experiment, rather than responding to the stimuli on the basis of previous knowledge or experience from outside the laboratory. If infants only generated expectations about test objects' properties and performed target actions on test objects on the basis of the preceding trial with a functional corresponding prototype, it was expected that they would perform significantly more target actions on test objects in the predicted condition than the interest control condition.

A univariate 2 (labelling) x 2 (expectation) x 4 (shape) mixed model analysis of variance (ANOVA) was performed on the data, with labelling (label, no label) as a between-subjects factor, and expectation (predicted, interest control) and shape (high, medium, low, dissimilar) as within-subjects factors. An alpha level of .05 was used for all omnibus analyses. As Mauchley's test of sphericity was significant for the effect of shape, $\underline{W} = .79$, $\underline{p} < .05$, and the Expectation x Shape interaction, $\underline{W} = .81$, $\underline{p} < .05$, the Greenhouse-Geisser adjustment to degrees of freedom was made in determining the critical F-values reported in the following omnibus analysis. The ANOVA revealed a significant main effect of shape, $\underline{F}$ (2.66, 165) = 26.65, $\underline{MSE}$ = 125.71, $\underline{p} < .001$, and expectation, $\underline{F}$ (1, 55) = 76.91, $\underline{MSE}$ = 879.24, $\underline{p} < .001$, but not labelling, $\underline{F}$ (1,55) = .43, $\underline{MSE}$ = 4.65, $\underline{p}$ = .52. The significant main effects were qualified by a significant Expectation x Shape interaction, $\underline{F}$ (2.68, 165) = 26.64, $\underline{MSE}$ = 127.00, $\underline{p} < .001$. The Labelling x Shape interaction was not significant, $\underline{F}$ (2.66, 165) = .56, $\underline{MSE}$ = 2.65, $\underline{p}$ = .62, nor was the Labelling x Expectation interaction, $\underline{F}$ (1, 55) = .59, $\underline{MSE}$ = 6.77, $\underline{p}$ = .45. The 3-way Labelling x Expectation x Shape interaction was not significant, $\underline{F}$ (2.68, 165) = .67, $\underline{MSE}$ = 3.19, $\underline{p}$ = .56. Thus, infants' frequency scores did not differ significantly depending on whether or not they were in the label or no label group. However, infants performed significantly more target actions on test objects in the predicted condition than in the interest control condition, as expected.

In order to follow up the Expectation x Shape interaction, simple effects tests comparing the frequency of target actions performed in the predicted condition to the interest control condition were conducted, across all levels of shape. A Bonferroni correction was applied to these comparisons, at an alpha level of .05/4 = .0125. The means for each condition are presented in Figure 5. As expected, the frequency of target actions performed on test objects was significantly higher in the predicted than in the interest control condition for the high shape objects [$\underline{M}$ = 5.09, $\underline{SD}$ = 4.76 (predicted), $\underline{M}$ = 0.02, $\underline{SD}$ = 0.13 (interest control)], $\underline{F}$ (1, 56) = 64.12, $\underline{MSE}$ = 732.64, $\underline{p} < .001$; medium shape objects [$\underline{M}$ = 3.49, $\underline{SD}$ = 3.99 (predicted), $\underline{M}$ = 0.11, $\underline{SD}$ = 0.41 (interest control)], $\underline{F}$ (1, 56) = 40.33,

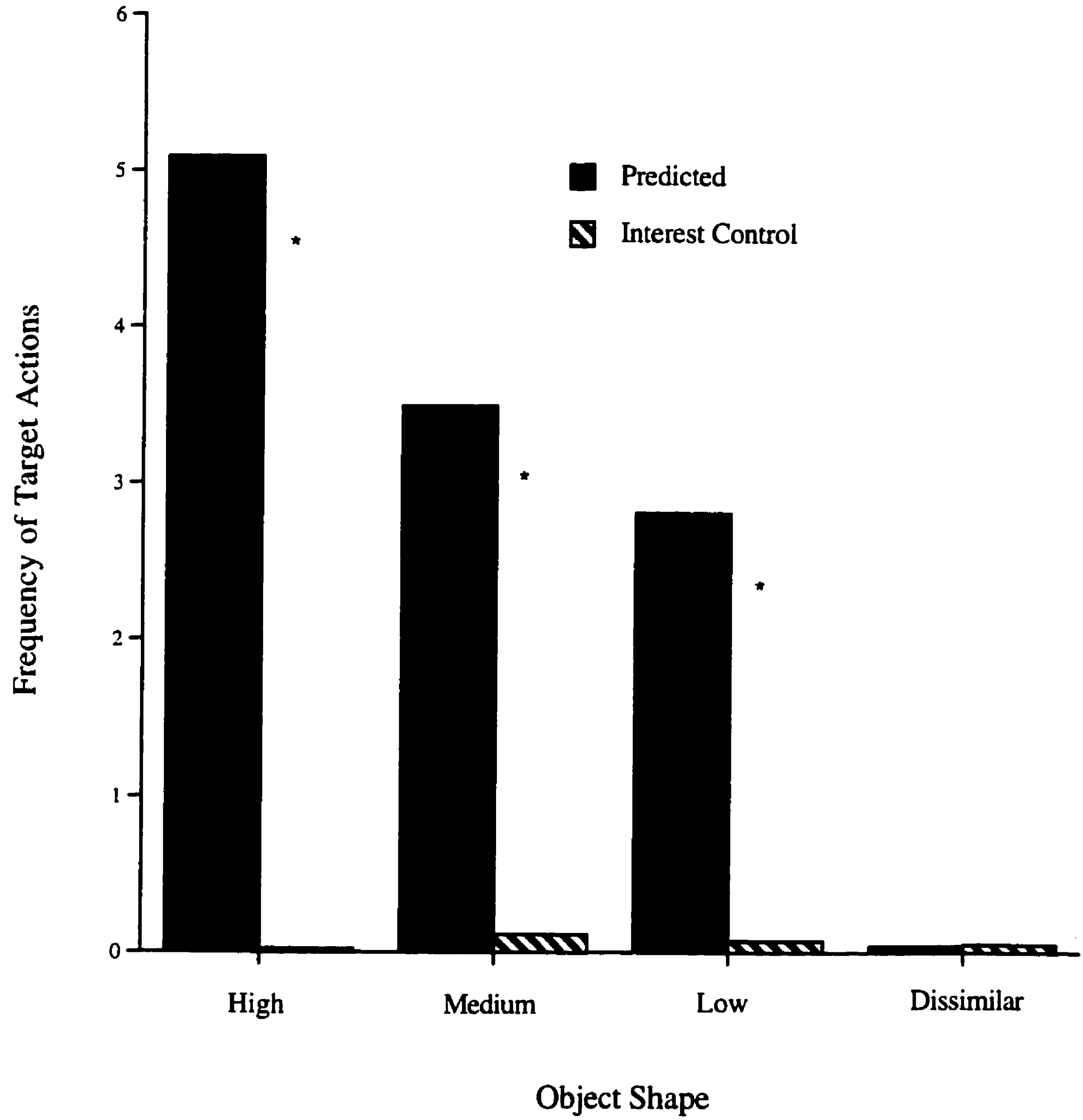


Figure 5. Mean frequency of target actions performed on high shape, medium shape, low shape, and dissimilar test objects in the predicted and interest control condition, collapsed across the label and no label condition (Experiment 1).

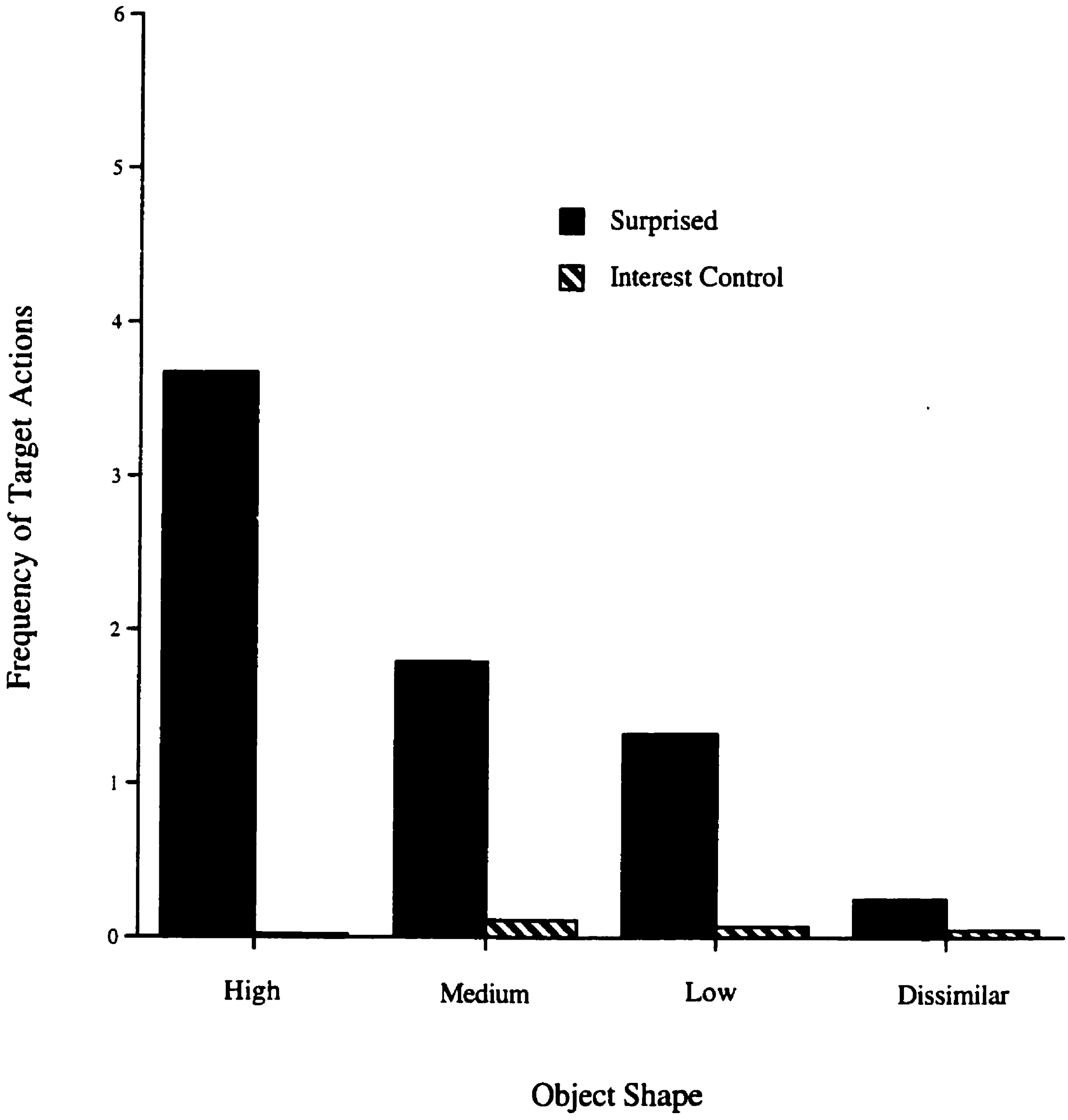
Note. * Indicates a significant difference between the predicted and interest control condition, $p < .001$.

MSE = 326.75, p < .001; and low shape objects [M = 2.81, SD = 3.38 (predicted), M = 0.07, SD = 0.26 (interest control)], F (1, 56) = 35.95, MSE = 213.47, p < .001. However, the frequency of target actions performed on the dissimilar object did not differ significantly between the predicted and interest control condition [M = 0.04, SD = 0.19 (predicted), M = 0.05, SD = 0.40 (interest control)], F (1, 56) = 0 .09, MSE = 0.01, p = .77. As it was expected that infants would rarely perform target actions on dissimilar test objects in any condition, a nonsignificant difference between the predicted and interest control condition was not surprising. Overall then, these analyses indicated that the properties of the test objects were indeed nonobvious. Infants did not expect the objects to have particular sound properties because of the way these objects looked or because of prior experience with similar kinds of objects outside of the laboratory. Rather, infants performed target actions on test objects only after they had been exposed to the properties of particular functional prototype objects during the testing session.

Next, infants' specific expectations about object properties and their inductive generalizations of target properties were examined. In order for target actions performed on test objects to be clearly indicative of induction, test objects had to be disabled so that they could not perform the target property (i.e., make the sound). If infants expected test objects to share the same property as their corresponding prototypes, then they should have persisted in trying to get these test objects to perform. Thus, it was hypothesized that there would be significantly higher frequencies of target actions performed in the surprised condition than the interest control condition, particularly under two conditions: (a) when a label was provided for prototype and test objects, and (b) when the test object was higher in shape similarity to its corresponding prototype. It was predicted that when infants were provided with labels for test objects, they would perform significantly more target actions than when they were not provided with labels for the objects, as they could use the labels to induce that the prototype and test object were of the same kind and thus, shared the same underlying property. In addition, it was predicted that at lower levels of shape similarity (i.e., with the low shape object), object label information would significantly increase participants' persistence in trying to get the test object to perform, as shape information would not be as salient for promoting inferences and infants could rely on labels as a basis for making inductions.

A univariate 2 (labelling) x 2 (expectation) x 4 (shape) mixed-model ANOVA was performed on the data, with labelling (label, no label) as the between-subjects factor, and expectation (surprised, interest control) and shape (high, medium, low, dissimilar) as within-subjects factors. An alpha level of .05 was used for all omnibus analyses. As Mauchley's test of sphericity was significant for the effect of shape, $\underline{W} = .56$, $\underline{p} < .001$, and the Expectation x Shape interaction, $\underline{W} = .54$, $\underline{p} < .001$, the Greenhouse-Geisser adjustment to degrees of freedom was made in determining the critical F-values reported in the following omnibus analysis. The ANOVA revealed a significant main effect of shape, $\underline{F}$ (2.16, 165) = 25.25, $\underline{MSE}$ = 58.14, $\underline{p} < .001$, and expectation, $\underline{F}$ (1, 55) = 76.44, $\underline{MSE}$ = 328.91, $\underline{p} < .001$, but not labelling, $\underline{F}$ (1, 55) = .93, $\underline{MSE}$ = 4.22, $\underline{p}$ = .34. Neither the Labelling x Shape interaction, nor the Labelling x Expectation interaction was significant, $\underline{F}$ (2.16, 165) = 1.81, $\underline{MSE}$ = 4.17, $\underline{p}$ = .17, and $\underline{F}$ (1, 55) = .60, $\underline{MSE}$ = 2.59, $\underline{p}$ = .44, respectively. The Expectation x Shape interaction was significant, $\underline{F}$ (2.11, 165) = 25.23, $\underline{MSE}$ = 60.77, $\underline{p} <$ .001. The 3-way Labelling x Expectation x Shape interaction was not significant, $\underline{F}$ (2.11, 165) = 1.75, $\underline{MSE}$ = 4.22, $\underline{p}$ = .18. Thus, unexpectedly, infants in the label and no label condition performed equally well. Although it was found that infants performed significantly more target actions on objects in the surprised condition than in the interest control condition, the frequency of target actions performed on test objects also depended on object shape similarity.

To examine the effect of shape, planned comparisons were conducted within the surprised condition using a Bonferroni correction (alpha level of .05/6 = .008). The means for each of the objects in the surprised (and interest control condition) are presented in Figure 6. As expected, these comparisons revealed that infants performed significantly more target actions on the high shape objects (M = 3.67, SD = 3.64) than the medium shape objects (M = 1.79, SD = 2.02), F (1, 56) = 18.74, MSE = 100.43, $p < .001$, low shape objects (M = 1.32, SD = 1.92), F (1, 56) = 24.33, MSE = 157.51, $p < .001$, and dissimilar objects (M = 0.25, SD = 1.31), F (1, 56) = 41.33, MSE = 333.55, $p < .001$. In addition, as expected, infants performed significantly fewer target actions on dissimilar object than on both medium shape objects, F (1, 56) = 28.37, MSE = 67.93, $p < .001$, and low shape objects, F (1, 56) = 12.04, MSE = 32.64, $p < .01$. Surprisingly, however, there were no significant differences between the frequency of target actions performed on medium and low shape objects, F (1, 56) = 2.06, MSE = 6.39, p = .16. Thus, high shape similarity was utilized more often as the basis for inductive generalization of target nonobvious properties than the other levels of shape similarity, within the surprised condition. No follow-up tests were planned or conducted within the interest control condition, as these comparisons were not particularly relevant or interesting, and few target actions were performed on interest control objects among all levels of shape.



<u>Figure 6.</u> Mean frequency of target actions performed on high shape, medium shape, low shape, and dissimilar test objects in the surprised and interest control condition, collapsed across the label and no label condition (Experiment 1).

To clarify the robustness and generality of the frequency results across object sets used in the study, a univariate 2 (expectation) x 3 (shape) x 3 (object set) mixed model ANOVA was performed on the data, with expectation (surprised, interest control) and shape (high, medium, low) as the within-subjects factor, and object set (ball, bell, rattle) as the between-subjects factor. The dissimilar objects were not included in this analysis, as it had already been established infants did not use them as an inductive base. In this analysis, infants' target actions performed on high, medium, and low shape test objects across the three object sets were analyzed. An alpha level of .05 was used for all omnibus analyses. As Mauchley's test of sphericity was significant for the effect of shape, $W = .86$, $p < .05$, and the Expectation x Shape interaction, $W = .78$, $p < .01$, the Greenhouse-Geisser adjustment to degrees of freedom was made in determining the critical F-values reported in the following omnibus analysis. As expected, there was no main effect of object set, F (2, 54) = .19, MSE = 1.15, $p = .82$, but there was a main effect of expectation, F (1, 54) = 70.11, MSE = 395.12, $p < .001$, and shape, F (1.76, 108) = 17.07, MSE = 42.16, $p < .001$. The main effects of expectation and shape were qualified by a significant Expectation x Shape interaction, F (1.64, 108) = 18.62, MSE = 47.92, $p < .001$. The Object Set x Shape interaction was not significant, F (3.52, 108) = .69, MSE = 1.71, $p = .58$, and neither was the Object Set x Expectation interaction, F (2, 54) = .09, MSE = .50, $p = .92$, nor the 3-way Object Set x Expectation x Shape interaction, F (3.29, 108) = 1.22, MSE = 3.15, $p = .31$. Importantly, there were no object set main effects or interactions, indicating that infants did not treat the objects differently, depending on which set they were from. Thus, the frequency data were collapsed across all three object sets for all analyses. No other tests were conducted to follow-up the Expectation x Shape interaction, as these data were examined in previous analyses.

Although there was no significant effect of the label in any of the above omnibus analyses, it was possible that older infants or those with a larger productive vocabulary size may have used the label to promote inductions. The MCDI revealed a wide range of productive vocabulary size among participants ($\underline{M}$ = 106.55 words, $\underline{SD}$ = 127.67; range = 3 to 498 words). Using a median split for both age of participants (19.07 months), and vocabulary level (35 words), infants' performance in the surprised condition was examined separately in a two univariate 2 (labelling) x 2 (age/vocabulary) x 3 (shape) mixed model ANOVAs, with labelling (label, no label) and age/vocabulary (under median, over median) as between-subjects factors, and shape (high, medium, and low) as the within-subjects factor.

First, the effect of age was examined. As Mauchley's test of sphericity was significant for the effect of shape, $\underline{W} = .88$, $\underline{p} < .05$, the Greenhouse-Geisser adjustment to degrees of freedom was made in determining the critical F-values reported for the shape effect in the following age median split omnibus analysis. The ANOVA revealed a significant main effect of shape, $\underline{F}$ (1.77, 102) = 20.30, $\underline{MSE}$ = 94.53, $\underline{p} < .001$, but not labelling or age, $\underline{F}$ (1, 53) = .78, $\underline{MSE}$ = 8.90, $\underline{p}$ = .38, and $\underline{F}$ (1, 53) = .71, $\underline{MSE}$ = 8.06, $\underline{p}$ = .40, respectively. Neither the Labelling x Age interaction nor the Labelling x Shape interaction was significant, $\underline{F}$ (1, 53) = .08, $\underline{MSE}$ = .92, $\underline{p}$ = .78, and $\underline{F}$ (1.77, 106) = 2.78, $\underline{MSE}$ = 12.94, $\underline{p}$ = .07. However, the Age x Shape interaction was significant, $\underline{F}$ (1.77, 106) = 3.36, $\underline{MSE}$ = 15.64, $\underline{p} < .05$. The 3-way Age x Labelling x Shape interaction was not significant, $\underline{F}$ (1.77, 106) = .98, $\underline{MSE}$ = 4.54, $\underline{p}$ = .37.

The Age x Shape interaction was followed up by comparing the under median and over median age group at each level of shape. The means were as follows: for high shape objects, $M = 4.48$, $SD = 4.23$ (under median), and $M = 2.93$, $SD = 2.85$ (over median), for medium shape objects, $M = 1.78$, $SD = 2.06$ (under median), and $M = 1.80$, $SD = 2.02$ (over median), and for low shape objects, $M = 1.15$, $SD = 2.18$ (under median), and $M = 1.47$, $SD = 1.67$ (over median). No significant differences between these comparisons were found (all p's $> .05$). Because there were only 27 participants in the under median group and 30 participants in the over median group, there may not have been enough power to detect a difference in means for the high shape objects in particular.

Next, the effect of productive vocabulary size was examined. As Mauchley's test of sphericity was significant for the effect of shape, $W = .88$, $p < .05$, the Greenhouse-Geisser adjustment to degrees of freedom was made in determining the critical F-values reported for the shape effect in the following vocabulary median split omnibus analysis. The ANOVA revealed a significant main effect of shape, $F(1.78, 102) = 16.01$, $MSE = 79.17$, $p < .001$, but not labelling or vocabulary, $F(1, 51) = .88$, $MSE = 10.34$, $p = .35$, and $F(1, 51) = .14$, $MSE = 1.61$, $p = .71$. None of the Labelling x Shape, Vocabulary x Shape, and Vocabulary x Labelling interactions were significant, $F(1.78, 102) = 1.97$, $MSE = 9.73$, $p = .15$; $F(1.78, 102) = .14$, $MSE = 0.67$, $p = .85$; and $F(1, 51) = .06$, $MSE = 0.67$, $p = .81$, respectively. In addition, the 3-way Vocabulary x Labelling x Shape interaction was not significant, $F(1.78, 102) = .90$, $MSE = 4.44$, $p = .40$.

As no significant age or productive vocabulary effects were found through the use of median splits, a series of Pearson correlations were calculated between: (a) age and the frequency of target actions performed in the surprised condition (see Table 2), and (b)

MCDI scores and the frequency of target actions performed in the surprised condition (see Table 3). None of these correlations reached significance (all p's > .05).

Object Transfer Data Analyses

Another question of interest in this experiment was whether infants performed instances of object transfer (e.g., when a bell set object was shaken or squeezed instead of tapped). Those participants with test frequency outliers (n = 7; see data screening section) were not included in the object transfer analyses, as it was assumed that if participants were engaged in performing a very high frequency of appropriate target actions, they would not be producing meaningful object transfer. However, even if participants did not perform any appropriate target actions on high, medium, or low shape test objects in the surprised condition, they were still included in the transfer analyses. Thus, the final transfer analyses sample included 65 infants. In the surprised condition, instances of transfer were observed with 25 participants (38%), and in 58 out of 195 total possible trials (30%). Of the 58 trials in which infants performed object transfer actions in the surprised condition, 16 were high shape object trials, 17 were medium shape object trials, and 25 were low shape object trials.

A univariate 2 (labelling) x 2 (expectation) x 3 (shape) mixed-model ANOVA was performed on the transfer data, with labelling (label, no label) as the between-subjects factor, and expectation (surprised, interest control) and shape (high, medium, and low) as within-subjects factors. An alpha level of .05 was used for all omnibus analyses. As Mauchley's test of sphericity was significant for the effect of shape, $W = .66$, $p < .001$, the Greenhouse-Geisser adjustment to degrees of freedom was made in determining the critical F-values reported for the main effect of shape. The ANOVA revealed a significant main effect of expectation [$M = 0.78$, $SD = 1.93$ (interest control), $M = 0.44$ $SD = 1.21$

Table 2

Correlations Between Frequency of Target Actions and Age in the Surprised Condition in Experiment 1 (Label and No Label Condition), and Experiment 2 (Salient Label Condition).

	Object Shape			
Labelling Condition	High	Medium	Low	All Objects
Label	.06 (30)	.08 (30)	.31 (30)	.18 (30)
No Label	-.03 (27)	.31 (27)	.12 (27)	.11 (27)
Label + No Label Combined	.01 (57)	.18 (57)	.19 (57)	.13 (57)
Salient Label	-.13 (29)	.34 (29)	-.10 (29)	.03 (29)
Salient Label + No Label Combined	.03 (56)	.36 (56)	.08 (56)	.18 (56)

Note. Numbers in parentheses indicate number of participants.

All p's > .05.

Table 3

Correlations Between Frequency of Target Actions and Vocabulary in the Surprised Condition in Experiment 1 (Label and No Label Condition), and Experiment 2 (Salient Label Condition).

	Object Shape			
Labelling Condition	High	Medium	Low	All Objects
Label	.00 (30)	-.02 (30)	.06 (30)	.01 (30)
No Label	-.20 (25)	-.11 (25)	-.08 (25)	-.18 (25)
Label + No Label Combined	-.10 (55)	-.06 (55)	-.04 (55)	-.09 (55)
Salient Label	.14 (28)	.23 (28)	.42 (28)	.37 (28)
Salient Label + No Label Combined	-.02 (53)	.10 (53)	.21 (53)	.11 (53)

Note. Numbers in parentheses indicate number of participants.

All p's > .05.

(surprised)], F (1, 63) = 6.72, MSE = 14.82, $p < .05$, but not labelling, F (1, 63) = .02, MSE = .10, p = .88, or shape, F (2.30, 189) = 1.97, MSE = 4.43, p = .14. Neither the Labelling x Shape interaction, the Labelling x Expectation interaction, nor the Expectation x Shape interaction was significant, F (3, 189) = .68, MSE = 1.52, p = .57; F (1, 63) = .14, MSE = .31, p = .71; and F (3, 189) = .85, MSE = 1.77, p = .47, respectively. In addition, the 3-way Labelling x Expectation x Shape interaction was not significant, F (3, 189) = .12, MSE = .26, p = .95. Thus, although there were rare instances of object transfer, there were no significant differences in transfer performance as a function of shape similarity or object labelling. There were, however, more overall incidences of object transfer in the interest control condition than in the surprised condition.

Latency to First Target Action Data Analyses

Sixty infants, 31 girls and 29 boys, (mean age = 18.81 months; SD = 1.46; range = 15.79 to 21.46 months) were considered in the analysis of the data. The label condition was comprised of 18 girls and 15 boys (mean age = 18.83 months; SD = 1.53; range = 15.79 to 21.46 months), and the no label condition was comprised of 13 girls and 14 boys (mean age = 18.79 months; SD = 1.40; range = 16.26 to 21.10 months).

Inspection of the latency to first target action data indicated that it could not be analyzed using standard parametric statistics, as there were many trials in which participants had no latency scores. For example, if a participant made no attempt to perform any target actions on interest control test objects, coders could not code latency scores for any of these objects. Latencies could not be coded as "0" as with the frequency data, as that would have meant infants performed a target action immediately, with no time lag between the experimenter releasing the object and the infant performing the first target action. Due to the amount of missing latency data in the surprised, predicted, and interest control condition, parametric inferential statistics could not be performed (see Table 4 for mean latency scores within the surprised condition). Instead, latency scores for high, medium, and low shape objects within the surprised condition were compared separately for the label and no label groups, using the Friedman's Rank Test for related samples. This test is the non-parametric, distribution-free analogue to the one-way repeated-measures ANOVA, in which scores are ranked and analyses are performed on the ranks rather than the raw scores (Howell, 1999).

The first Friedman's Rank analysis involved comparing the ranks of latency scores among high shape, medium shape, and low shape objects within the label condition. Only those participants who had a latency score for all three high, medium, and low shape objects were included in the analysis (N = 13). Of those who had missing latency scores (N = 20), 3 did not have latency scores for high shape objects, 15 did not have latency scores for medium shape objects, and 10 did not have latency scores for low shape objects. For each of the remaining 13 participants, high, medium, and low shape object latency scores were ranked from lowest to highest (i.e., 1 to 3; lowest rank = shortest latency), and these ranks were summed across participants, for each level of shape. The sums of the ranks were as follows: high shape = 23, medium shape = 26, and low shape = 28. The Friedman's Rank Test indicated that the high, medium, and low shape objects' rankings were not significantly different from one another, $\chi^2_F(2, \underline{N} = 13) = -3.05$, $\underline{p} > .05$.

The second Friedman's Rank analysis involved comparing the ranks of latency scores among high shape, medium shape, and low shape objects within the no label condition. Only those participants who had a latency score for all three high, medium, and low shape objects were included in the analysis (N = 6). Of those who had missing latency

Table 4

Mean Latency to First Target Action Scores for Infants in Experiment 1 (Label and No Label Condition) and Experiment 2 (Salient Label Condition) within the Surprised Condition.

	Label			No Label			Salient Label		
Object Shape	n	M	SD	n	M	SD	n	M	SD
High	30	1.36	1.31	23	1.71	1.80	28	2.05	3.10
Medium	18	1.71	2.01	17	2.09	2.01	20	0.82	3.10
Low	23	1.92	1.91	10	2.35	3.26	14	1.43	1.17
Dissimilar	02	0.60	0.09	01	2.84				

Note. Latency scores were in seconds.

scores (N = 21), 4 did not have latency scores for high shape objects, 10 did not have latency scores for medium shape objects, and 17 did not have latency scores for low shape objects. For each of the remaining 6 participants, high, medium, and low shape object latency scores were ranked from lowest to highest (i.e., 1 to 3; lowest rank = shortest latency) and these ranks were summed across participants, for each level of shape. The sums of the ranks were as follows: high shape = 9, medium shape = 15, and low shape = 12. The Friedman's Rank Test indicated that the ranks were not significantly different from one another, $\chi^2_F(2, \underline{N} = 6) = 3.01$, $\underline{p} > .05$.

The third Friedman's Rank analysis involved comparing the ranks of latency scores among high shape, medium shape, and low shape objects within the label and no label condition together. Only those participants who had a latency score for all three shape objects were included in the analysis (N = 19). High, medium, and low shape object latency scores were ranked from highest to lowest (i.e., 1 to 3; lowest rank = shortest latency), and these ranks were summed across all participants, for each level of shape. The sums of the ranks were as follows: high shape = 32, medium shape = 41, and low shape = 40. The Friedman's Rank Test indicated that the high, medium, and low shape objects' rankings were not significantly different from one another, $\chi^2_F(2, \underline{N} = 19) = -1.56$, $\underline{p} > .05$. Thus, in Experiment 1, the relative rankings of the latency scores did not distinguish infants' inductive generalizations.

Discussion

The goal of Experiment 1 was to examine the role of object labels and object shape similarity in promoting inductive inferences in infancy. Novel objects were presented to infants in one of two between-subjects conditions. In the label condition, infants were taught novel labels for novel objects (e.g., "Look at this zas!"), whereas in the no label condition infants were not provided with any labels for the objects (e.g., "Look at this one!"). Objects were presented to infants in three within-subject expectation conditions. In the surprised condition, prototype (target) objects were functional (i.e., they produced a target nonobvious sound) but test objects were disabled so that they could not produce the sound. In the interest control condition, neither prototype objects nor test objects were functional. In the predicted condition, both prototype and test objects were functional. Test objects were high shape, medium shape, or low shape similarity to, or dissimilar in both shape and texture to, a corresponding prototype. The frequency of target actions performed on test objects, the latency to first target action performed on test objects, and frequency of test object transfer were examined in the context of investigating infants' inductive generalizations.

Two main conclusions can be drawn from the results of Experiment 1. First, infants will generalize a specific nonobvious property from a prototype object to other test objects perceived as "in-category" members, when they have knowledge about the particular property the prototype possesses. Infants performed significantly more target actions on test objects in the predicted and surprised condition than the interest control condition. This finding indicates that: (a) test objects' properties were nonobvious, (b) infants used an inductive reasoning strategy to conclude that other test objects should share the same nonobvious sound property as their prototype, and (c) infants formed specific expectations about the nonobvious properties of test objects from knowledge gained during the testing session about the functionality of prototype exemplars. It was also found that infants rarely performed target actions on dissimilar objects in either the surprised or predicted condition.

This finding suggests that infants only generalized object properties when test objects minimally shared the same texture as their prototype. Importantly, the lack of imitation on dissimilar objects indicates that infants were not merely mimicking the experimenter's actions on any object they were handed; they were particular about when they would generalize object properties. Thus, infants were not just making associations between imitations and objects in general--they only generalized the nonobvious property when it was considered appropriate. In addition, because incidences of object transfer were rare, it can be concluded that infants only performed target actions specifically intended for a particular object set; they were not simply imitating any imitative action they had learned previously from other object sets.

The second conclusion that can be drawn from Experiment 1 is that infants will generalize nonobvious properties on the basis of high shape similarity between objects. Infants performed significantly more target actions on high shape test objects than medium shape, low shape, or dissimilar test objects. This suggests that infants expect objects that share the same shape to also share nonobvious properties, an issue that will be discussed further in the General Discussion. Interestingly, infants did not treat medium and low shape objects significantly differently in terms of the number of target actions performed on them. One possible explanation for this finding is that infants did not perceive the medium and low shape objects as different in shape, even though adults rated them as such. Another possibility is that if objects were not perceived as highly similar in shape, infants tended to overlook shape differences in favor of salient shared texture, or material kind, and thus treated both low and medium shape objects as being approximately equivalent in terms of perceived similarity to a prototype. This possibility seems likely, as infants performed significantly more target actions on low and medium shape objects than dissimilar test objects, and dissimilar test objects did not possess the same texture as any of the other objects. Thus, infants may have considered a texture change important for deciding whether particular objects belong to the same category or not.

A purely descriptive comparison between the surprised and predicted condition indicates that infants' tendency to perform target actions on test objects was more pronounced when test objects performed in the manner expected (i.e., more target actions were performed in the predicted than in the surprised condition). However, it becomes problematic to interpret the results of this comparison, as it is unclear why infants continued to perform target actions on functional test objects (in the predicted condition). It is likely that the interesting sound produced by functional test objects had a reinforcing quality, and so infants continued to perform target actions on these objects. Thus, how does one distinguish those target actions performed as a result of expectations formed from knowledge about prototype objects, from those performed as a result of the reinforcing nature of the sound property? Because of the difficulty in interpreting infants' motivations for performance in the predicted condition, inferential comparisons between the surprised and predicted conditions were neither planned nor conducted. Importantly, however, the same pattern of shape data was found in the predicted and surprised condition, indicating that the functionality (and thus reinforcing quality) of all test objects in the predicted condition still did not result in as many target actions being performed on medium and low shape objects as on high shape objects. This suggests that in the predicted condition, both shape similarity as well as objects' reinforcing sound quality influenced the number of target actions performed on test objects.

Some unexpected null effects were found in Experiment 1. There were no differences in inductive performance as a function of the provision of object labels, even though it was expected that object labels would promote inductive inferences. Perhaps the shape and texture properties of prototype objects were more salient for infants than object labels. Thus, infants may have overlooked the label (conceptual) information, focusing instead on the perceptual information provided to them, when making inductive inferences. In addition, each time a prototype object's label was introduced in either the predicted or surprised condition of Experiment 1, it was immediately followed by a demonstration of the interesting sound property. Thus, infants may have attended primarily to the interesting sound made by a prototype object and missed the accompanying label information.

No differences in inductive performance were found as a function of infants' age, suggesting that developmental maturity was not related to inductive ability, at least in this experiment, and within the age range of 16- to 21 months. These results are consistent with Baldwin et al. (1993), who also did not find significant age differences (in terms of frequency and latency scores) in their sample of 9- to 16-month-olds. In addition, no differences in inductive performance were found as a function of infants' productive vocabulary size. This suggests that verbal competence may not be significantly tied to the cognitive strategies involved in inductive generalization, at least within infants' range of linguistic abilities in this experiment. These results are consistent with other research which has found that infants approximately the same age as those in the present experiment extend novel words on the basis of shape similarity, regardless of infants' productive vocabulary size or composition (e.g., Graham & Poulin-Dubois, in press; Poulin-Dubois, Frank, Graham, & Elkin, 1999). Though it may be the case that a measure of receptive vocabulary is a more accurate reflection of infants' abilities to use labels and make inferences on the basis of shared labels, numerous studies have shown that productive and receptive vocabulary are highly correlated (e.g., see Bloom, 1998, for a review). Therefore, it is likely that infants' inductive abilities would not be distinguishable by a comparison of receptive vocabulary size.

Interestingly, no differences were found when ranked latency scores in the label and no label conditions were compared across levels of shape in the surprised condition. This suggests that, in this experiment, latencies may not have been a valid measure for examining whether object shape similarity and object labels influenced infants' inductive generalizations. When infants expected test objects to share the same property as the prototype, they generally performed the first target action immediately (within the first few seconds). Thus, the lack of variance in infants' latency scores as well as the lack of latency data due to nonperformance of target actions was likely responsible for the lack of a shape effect in this data set.

To summarize, the results of Experiment 1 indicate that 16- to 21-month-olds use high shape similarity, but not labels, in promoting inductive inferences about nonobvious object properties. However, it is possible that infants can use label information as an inductive base, but that participants in Experiment 1 did not attend to the object labels because of the salience of other object features, such as the interesting sound produced by prototype objects (in the surprised and predicted conditions). In order to examine this possibility, Experiment 2 was conducted utilizing a "salient label" condition, in which object labels were introduced without the accompanying sound property demonstration. If infants simply did not attend to the labels in Experiment 1 because they were more interested in the sound objects' produced, then by isolating the introduction of the label information from the sound demonstration, the question of whether infants will use label information to promote inductive inferences can be answered with more confidence.

Experiment 2

Method

Participants

Thirty-two infants, 18 girls and 14 boys, (M = 19.46 months; SD = 1.03; range = 17.49 to 20.85 months) were tested in this experiment. Eight additional infants were tested but were excluded from the sample for the following reasons: fussiness (n = 3), parental cueing (n = 1), and experimenter error (n = 1). Four infants were excluded from the frequency and latency data analyses as they did not perform target actions on any of the high, medium, and low shape test objects in the surprised condition. These four infants were not excluded from the object transfer analyses. Parents of participants were recruited through newspaper advertisements and brochures in four Calgary Regional Health Authority medical clinics. All parents gave informed consent and were debriefed after testing, in accordance with the University of Calgary Ethical Guidelines. Infants were presented with a small toy and certificate after testing, thanking them for their participation. None of the infants who participated in Experiment 1 participated in Experiment 2.

Materials

Stimuli. The same objects were used in Experiment 2 as Experiment 1, however the dissimilar objects (i.e., strainer, hose splitter, and file) were not included. The purpose for using dissimilar test objects was to examine whether infants would imitate target actions indiscriminately on any object they were given. It was demonstrated in Experiment 1 that, indeed, infants did not generally imitate the prototype's action on dissimilar test objects. Thus, these objects were eliminated from Experiment 2.

Equipment. Same as Experiment 1.

Design and procedure

The design and procedure were similar to Experiment 1, with two exceptions: (a) lack of a dissimilar object, and (b) labels for prototype objects were taught to infants three times without the accompanying property (sound) demonstration, and three times in conjunction with the property (sound) demonstration. Labels for test objects were repeated four times. Thus, the object labels in Experiment 2 were repeated twice as often as in Experiment 1. In an excited voice, the experimenter showed infants a prototype object and said, "Hi (infant's name)! Look at this (X)! This is a (X)! Yes this is a (X)! (property demonstration). See this (X)! (demonstration twice) Look at this (X)! (demonstration twice) Yes, this is a (X)!" The experimenter showed infants the corresponding prototype and said, "Okay (infant's name)! Look at this (X)! Yes this is a (X)! Look at this (X)! Yes this is a (X)!" Thus, in the surprised and predicted conditions, the prototype's property was demonstrated the same number of times as in Experiment 1 (five times), but the infants were given more opportunity to attend to the label before perhaps becoming distracted by or more interested in the sounds produced by these objects. As in Experiment 1, the first time each prototype was shown to infants, parents also demonstrated the property twice, and no properties were demonstrated on the prototype in the interest control condition.

As in Experiment 1, all MCDI's that were completed at home were mailed back within a one-week period, except for one which were not returned (this infant's data was still included in all analyses except for the productive vocabulary analyses).

Coding

The target actions for the ball, bell, and rattle object sets were coded according to the same criteria as Experiment 1.

As in Experiment 1, approximately 20% of the data (n = 8 participants) were coded a second time to assess inter-rater reliability. The Intraclass correlation (ICC) statistic was computed to assess the level and pattern of agreement in frequency and latency ratings. ICC coefficients for prototype and test object frequency ratings were both significant, ICC (144) = 1.00, $p < .001$, and ICC (144) = .99, $p < .001$, respectively. The ICC coefficient for test transfer frequency ratings was significant as well, ICC (144) = 1.00, $p < .001$. Thus, the two raters were in almost perfect agreement on all frequency ratings. ICC coefficients for prototype and test object latency ratings were also both significant, ICC (144) = 1.00, $p < .001$, and ICC (144) = .96, $p < .001$, respectively. Because of the amount of latency data coded as missing by both raters, the ICC calculation may have been too liberal. Thus, it was recalculated excluding the trials in which both raters coded an absent latency score. This ICC coefficient was also significant for both prototype and test object latency ratings, ICC (40) = .97, $p < .001$, and ICC (40) = .96, $p < .001$, respectively.

Results

As in Experiment 1, the data from Experiment 2 were analyzed in a number of steps. First, the data were screened for frequency of target action and latency to first target action outliers, which were eliminated from the final analyses. Second, the frequency of target action data for the prototype object and test object trials were analyzed. Third, the object transfer data were examined. Fourth, the latency to first target action data for test object trials were analyzed.

Data Screening

Frequency of target action and latency to first target action data were examined separately for the presence of univariate outliers. Z-scores greater than 3 standard deviations above or below the mean were eliminated from the final analyses. Inspection of the frequency data for outliers revealed three cases with scores greater than 3 standard deviations above the mean. Thus, approximately 9% of the cases were removed from the final frequency analyses. Inspection of the latency data for univariate outliers revealed two cases with scores greater than 3 standard deviations above the mean. Thus, approximately 6% of the cases were eliminated from the final latency analyses (see Appendix B for a description of outliers). As in Experiment 1, frequency of target actions and latency to first target action data were analyzed separately.

In the warm-up trials, only one participant failed to imitate target actions on at least one of the three objects. Because the participant did perform target actions on test objects in the experimental task, this participant's data were not eliminated from any subsequent analyses.

Frequency of Target Action Data Analyses

Twenty-nine infants, 16 girls and 13 boys, ($\underline{M}$ = 19.43 months; $\underline{SD}$ = 1.01; range = 17.49 to 20.85 months) were considered in the analyses of the frequency data.

As with Experiment 1, it was important to examine whether infants knew how the functional prototypes (i.e., in the surprised and predicted conditions) performed the target sound, in order to draw the conclusion that infants either did or did not perform target actions on test objects on the basis of expectations generated from knowledge about the nonobvious properties of functional prototypes. Infants performed at least one target action on functional prototypes in 153 of the 174 total possible prototype trials in the surprised and predicted condition (88% response rate). As in Experiment 1, those trials where target actions were not performed on the prototypes were not excluded. Participants tried the prototype and found out for themselves that it performed the interesting property on the majority of the trials, and arguably, infants knew about the nonobvious property of functional prototypes because the experimenter demonstrated the property for them.

As the general procedure, stimuli, and experimenter were the same in Experiments 1 and 2, the no label condition of Experiment 1 was compared with the salient label condition (Experiment 2). One limitation in drawing conclusions about the results of this comparison is that the experimenter gave more attention to objects when they were introduced in the salient label condition (Experiment 2) than in the no label condition (Experiment 1). However, this comparison is an important preliminary investigation into whether infants will use salient labels to promote inductive inferences.

To examine whether infants perceived the properties of test objects as nonobvious, a univariate 2 (labelling) x 2 (expectation) x 3 (shape) mixed model analysis of variance (ANOVA) was performed on the data, with labelling (salient label, no label) as a between-subjects factor, and expectation (predicted, interest control) and shape (high, medium, and low) as within-subjects factors. An alpha level of .05 was used for all omnibus analyses. As Mauchley's test of sphericity was not significant for any effects, no degrees of freedom adjustments were made in the omnibus analyses. The ANOVA revealed a significant main effect of labelling, $\underline{F}$ (1, 54) = 5.80, $\underline{MSE}$ = 198.09, $\underline{p} < .05$, expectation, $\underline{F}$ (1, 54) = 58.29, $\underline{MSE}$ = 1970.55, $\underline{p} < .001$, and shape, $\underline{F}$ (2, 108) = 6.08, $\underline{MSE}$ = 477.23, $\underline{p} < .01$. Although the Labelling x Shape interaction was not significant, $\underline{F}$ (2, 108) = .12, $\underline{MSE}$ = 1.58, $\underline{p}$ = .88, both the Labelling x Expectation interaction and the Expectation x Shape interaction reached significance, F (1, 54) = 5.32, MSE = 179.76, p <.05, and F (2, 108) = 5.75, MSE = 71.59, p < .01, respectively. The 3-way Labelling x Expectation x Shape interaction was not significant, F (2, 108) = .02, MSE = 0.25, p = .98. Thus, infants' frequency of target action scores were significantly higher in the salient label condition than in the no label condition of Experiment 1. As expected, frequency scores were higher overall in the predicted condition than in the interest control condition.

In order to follow up the Expectation x Shape interaction, simple effects tests were conducted to compare the frequency of target actions performed in the predicted condition to the interest control condition, across all levels of shape. A Bonferroni correction was applied for all comparisons, at an alpha level of .05/3 = .0167. As expected, the frequency of target actions performed on each object was significantly higher in the predicted than in the interest control condition for the high shape objects [M = 8.35, SD = 9.15 (predicted), M = 0.28, SD = 0.65 (interest control)], F (1,28) = 22.38, MSE = 944.07, p < .001, medium shape objects [M = 6.38, SD = 6.93 (predicted), M = 0.21, SD = 0.62 (interest control)], F (1, 28) = 24.51, MSE = 552.43, p < .001, and low shape objects [M = 4.76, SD = 6.89 (predicted), M = 0.07, SD = 0.26 (interest control)], F (1, 28) = 13.27, MSE = 318.90, p < .01. Overall then, these analyses indicated that the properties of the test objects were perceived as nonobvious by participants.

The Labelling x Expectation interaction was followed up with an examination of the labelling condition (no label, salient label) separately for the predicted and interest control condition. A Bonferroni correction was applied for these comparisons, at an alpha level of .05/2 = .025. Within the predicted condition, planned comparisons revealed a significantly higher frequency of target actions performed in the salient label condition (M = 6.50, SD = 5.99), than in the no label condition (M = 3.49, SD = 2.85), $F(1, 54) = 5.58$, MSE = 125.88, $p < .025$. Within the interest control condition, planned comparisons revealed no significant differences between the frequency of target actions performed in the salient label condition (M = 0.18, SD = 0.33) and in the no label condition (M = 0.11, SD = 0.29), $F(1, 54) = .76$, MSE = 0.07, $p = .39$.

Next, infants' inductive generalizations were compared in the surprised and interest control condition. It was hypothesized that there would be significantly higher frequencies of target actions performed on test objects in the surprised condition than in the interest control condition, particularly when salient labels were provided for prototype and test objects (salient label condition, Experiment 2), compared to when infants were not provided with labels (no label condition, Experiment 1). In addition, higher frequencies of target actions performed on test objects were expected when test objects were more similar in shape to their corresponding prototypes.

A univariate 2 (labelling) x 2 (expectation) x 3 (shape) mixed-model ANOVA was performed on the data, with labelling (salient label, no label) as the between-subjects factor, and expectation (surprised, interest control) and shape (high, medium, and low) as the within-subjects factors. An alpha level of .05 was used for all omnibus analyses. As Mauchley's test of sphericity was not significant for any effects, no degrees of freedom adjustments were made in the omnibus analyses. The ANOVA revealed a significant main effect of shape, $F(2, 108) = 31.65$, MSE = 124.94, $p < .001$, expectation, $F(1, 54) = 111.07$, MSE = 981.71, $p < .001$, and labelling $F(1, 54) = 12.81$, MSE = 105.85, $p < .01$. Both the Labelling x Expectation interaction and the Expectation x Shape interaction were significant, $F(1, 54) = 10.47$, $MSE = 92.57$, $p < .01$, and $F(2, 108) = 27.84$, $MSE = 120.83$, $p < .001$, respectively. The Labelling x Shape interaction was not significant, $F(2, 108) = 1.51$, $MSE = 5.98$, $p = .23$, nor was the 3-way Labelling x Expectation x Shape interaction, $F(2, 108) = .66$, $MSE = 2.86$, $p = .52$.

The Labelling x Expectation interaction was followed up with planned comparisons of labelling condition (no label, salient label) separately for the surprised and interest control condition. A Bonferroni correction was applied for these comparisons, at an alpha level of .05/2 = .025. Within the surprised condition, the comparison revealed a significantly higher frequency of target actions performed in the salient label condition ($M = 4.66$, $SD = 2.41$), than in the no label condition ($M = 2.49$, $SD = 2.32$), $F(1, 54) = 11.79$, $MSE = 66.06$, $p < .01$. Within the interest control condition, the comparison revealed no significant differences between the frequency of target actions performed in the salient label condition ($M = 0.18$, $SD = 0.33$) and the no label condition ($M = 0.11$, $SD = 0.29$), $F(1, 54) = 0.76$, $MSE = 0.07$, $p = .39$.

To examine the effect of shape, planned comparisons were conducted within the surprised condition using a Bonferroni correction (alpha level of .05/3 = .0167). The frequency means for each of the objects in the surprised and interest control condition are presented in Figures 7 and 8, respectively. As expected, these comparisons revealed that infants performed significantly more target actions on high shape objects than medium shape objects, $F(1, 28) = 17.37$, $MSE = 216.28$, $p < .001$, and low shape objects, $F(1, 28) = 32.29$, $MSE = 309.59$, $p < .001$. However, there were no significant differences between frequency of target actions performed on medium and low shape objects, $F(1, 28) = 0.80$, $MSE = 8.34$, $p = .38$. Thus, high shape similarity was utilized more often as the basis for

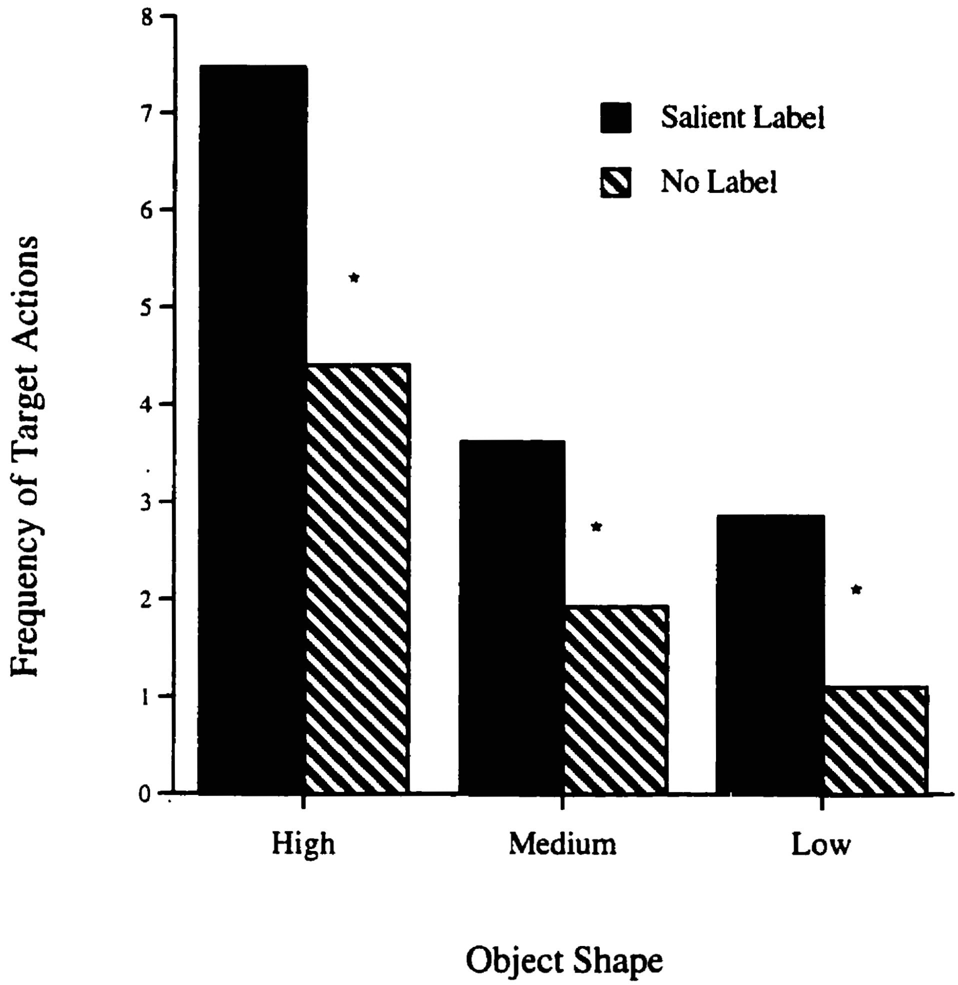


Figure 7. Mean frequency of target actions performed on high shape, medium shape, and low shape objects in the surprised condition, for the salient label (Experiment 2) and no label (Experiment 1) condition.

Note. * Indicates a significant difference between the salient label and no label condition, $p < .05$.

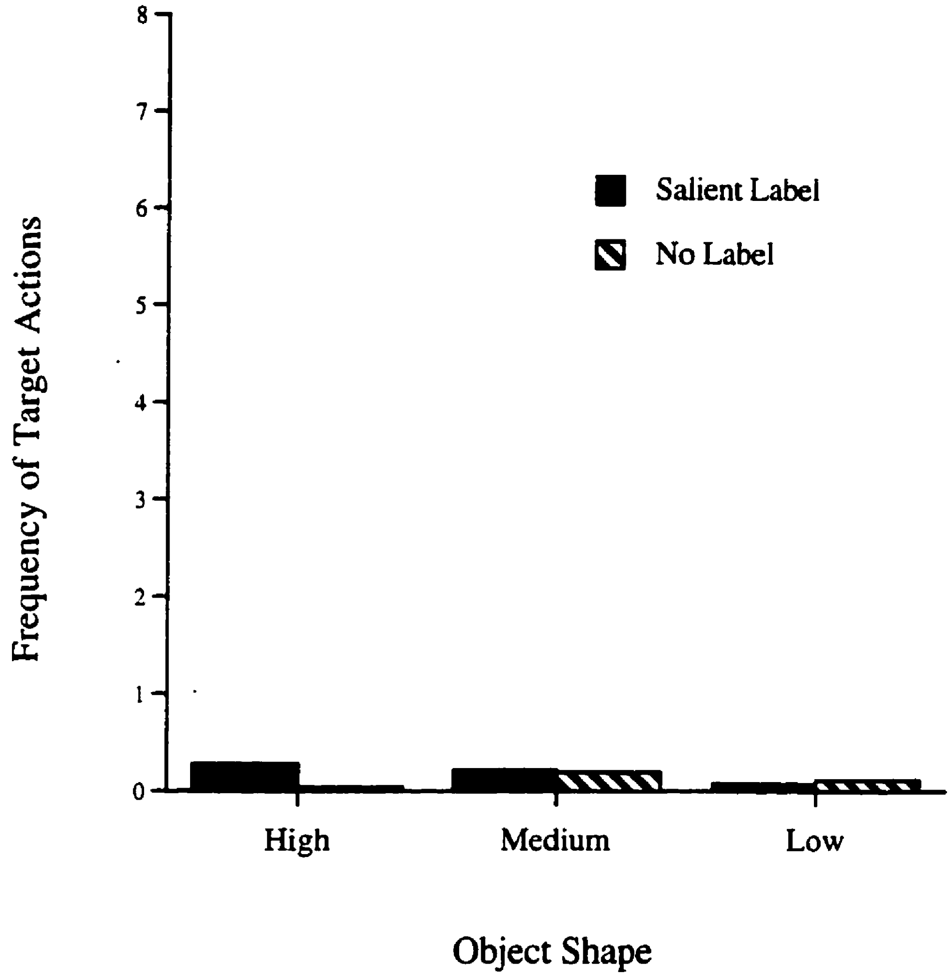


Figure 8. Mean frequency of target actions performed on high shape, medium shape, and low shape objects in the interest control condition, for the salient label (Experiment 2) and no label (Experiment 1) condition.

inductive generalization of the nonobvious properties than the other levels of shape similarity, within the surprised condition. No follow-up tests were planned or conducted within the interest control condition, as these comparisons were not particularly relevant or interesting, and few target actions were performed on interest control objects among all levels of shape.

Although the Labelling x Shape interaction in the omnibus analyses above was not significant, separate planned analyses were conducted to examine whether infants' inductive performance significantly differed in the no label condition (Experiment 1) and the salient label condition (Experiment 2), for high, medium, and low shape test objects within the surprised expectation condition (these means are presented in Figure 7). The alpha level remained at .05, as these comparisons were planned and were not omnibus follow-up tests. For high shape objects, infants performed significantly more target actions in the salient label condition ($M = 7.48$, $SD = 4.11$) than in the no label condition ($M = 4.41$, $SD = 4.12$), $F(1,54) = 7.81$, $MSE = 132.24$, $p < .05$. This was also found for medium shape objects, as infants performed significantly more target actions in the salient label condition ($M = 3.62$, $SD = 3.26$) than in the no label condition ($M = 1.93$, $SD = 2.00$), $F(1, 54) = 5.41$, $MSE = 40.16$, $p < .05$, as well as for low shape objects, as infants performed significantly more target actions in the salient label condition ($M = 2.86$, $SD = 3.40$), than the no label condition ($M = 1.11$, $SD = 2.36$), $F(1, 54) = 4.95$, $MSE = 42.87$, $p < .05$. Thus, infants in the salient label condition generalized target nonobvious object properties significantly more often than infants in the no label condition, at all levels of shape. Importantly, the presence of a labelling effect suggests that infants will use shared label information to guide their inductions about novel objects and their nonobviousproperties, when the labels are salient.

As in Experiment 1, to clarify the robustness and generality of the frequency results across object sets used in Experiment 2, a univariate 2 (expectation) x 3 (shape) x 3 (object set) mixed model ANOVA was performed on the data, with expectation (surprised, interest control) and shape (high, medium, low) as the within-subjects factor, and object set (ball, bell, rattle) as the between-subjects factor. It was expected that, as in Experiment 1, there would be no effect or interactions with type of object set. An alpha level of .05 was used for all omnibus analyses. As Mauchley's test of sphericity was not significant for any effects, no degrees of freedom adjustments were made in the omnibus analyses. There was no main effect of object set, F (2, 26) = 1.09, MSE = 8.85, p = .35, but there was a main effect of expectation, F (1, 26) = 87.46, MSE = 858.10, p < .001, and shape, F (2, 52) = 19.91, MSE = 96.47, p < .001. The main effects of expectation and shape were qualified by a significant Expectation x Shape interaction, F (2, 52) = 14.93, MSE = 83.55, p < .001. The Object Set x Shape interaction was not significant, F (4, 52) = 2.28, MSE = 11.04, p = .07, and neither was the Object Set x Expectation interaction, F (2, 26) = .55, MSE = 5.37, p = .59, nor the 3-way Object Set x Expectation x Shape interaction, F (4, 52) = 1.58, MSE = 8.85, p = .19. Importantly, there were no object set main effects or interactions, indicating that infants did not treat the objects differently, depending on the set to which each belonged. Thus, no follow-up tests of interest were performed.

As a significant labelling effect was found when the salient label condition was compared to the no label condition, it was of interest to examine whether there were any significant interactions between labelling condition and infants' age, and labelling condition and infants' productive vocabulary size. The MCDI revealed a wide range of productive vocabulary size among participants, M = 136 words, SD = 126.74; range = 4 to 441 words.

Using a median split for both age of participants (19.25 months), and vocabulary level (59 words), infants' performance in the surprised condition was examined separately in a two univariate 2 (labelling) by 2 (age/vocabulary) by 3 (shape) mixed model ANOVAs, with labelling (label, no label) and age/vocabulary (under median, over median) as between-subjects factors, and shape (high, medium, low) as the within-subjects factor.

First, the effect of age was examined. As Mauchley's test of sphericity was not significant for any effects, no degrees of freedom adjustments were made in the age mean split omnibus analyses. The ANOVA revealed a significant main effect of shape, $\underline{F}$ (2, 104) = 31.91, $\underline{MSE}$ = 246, $\underline{p}$ < .001, as well as labelling, $\underline{F}$ (1, 52) = 11.62, $\underline{MSE}$ = 199.92, $\underline{p}$ < .01, but not age, $\underline{F}$ (1, 53) = .20, $\underline{MSE}$ = 3.43, $\underline{p}$ = .66. Neither the Labelling x Age interaction nor the Labelling x Shape interaction was significant, $\underline{F}$ (1, 52) = .59, $\underline{MSE}$ = 10.23, $\underline{p}$ = .44, and $\underline{F}$ (2, 104) = 1.25, $\underline{MSE}$ = 9.62, $\underline{p}$ = .29. However, the Age x Shape interaction was significant, $\underline{F}$ (2, 104) = 4.04, $\underline{MSE}$ = 31.15, $\underline{p}$ < .05. The 3-way Age x Labelling x Shape interaction was not significant either, $\underline{F}$ (2, 104) = .59, $\underline{MSE}$ = 4.58, $\underline{p}$ = .55.

The Age x Shape interaction was followed up by comparing the under median and over median age group at each level of shape. The means were as follows: for high shape objects, $\underline{M}$ = 6.86, $\underline{SD}$ = 4.52 (under median), and $\underline{M}$ = 5.14, $\underline{SD}$ = 4.10 (over median), for medium shape objects, $\underline{M}$ = 2.18, $\underline{SD}$ = 2.67 (under median), and $\underline{M}$ = 3.43, $\underline{SD}$ = 2.90 (over median), and for low shape objects, $\underline{M}$ = 2.10, $\underline{SD}$ = 3.27 (under median), and $\underline{M}$ = 2.00, $\underline{SD}$ = 2.87 (over median). No significant differences were found in these comparisons (all $\underline{p's}$ > .05). As in Experiment 1, these comparisons may have been significant with a larger N, allowing for more power to detect significant differences in means.

The age range of infants was slightly broader in the no label condition (Experiment 1) than the salient label condition (Experiment 2). Thus, as a check to ensure these conditions were equally comparable for the variable of age, those infants in the no label condition whose ages were higher or lower than the age range of infants in the salient label condition ($n = 3$) were excluded from the analyses. An ANOVA was then performed using an age median split (19.40 months). As Mauchley's test of sphericity was not significant for any effects, no degrees of freedom adjustments were made in the omnibus analyses. Infants' performance was examined in a 2 (labelling) x 2 (age) x 3 (shape) mixed model ANOVA, with labelling (salient label, no label) and age (under median, over median) as between-subjects factors, and shape (high, medium, and low) as the within-subjects factor. The ANOVA revealed a significant main effect of shape, $F(2, 88) = 22.19$, $MSE = 201.72$, $p < .001$, and labelling, $F(1, 44) = 5.79$, $MSE = 100.96$, $p < .05$, but not age, $F(1, 44) = .36$, $MSE = 6.70$, $p = .55$. Neither the Labelling x Age interaction nor the Labelling x Shape interaction was significant, $F(1, 44) = 2.07$, $MSE = 38.77$, $p = .16$, and $F(2, 88) = .68$, $MSE = 6.17$, $p = .51$, respectively. The Age x Shape interaction was not significant either, $F(2, 88) = 1.94$, $MSE = 17.59$, $p = .15$. In addition, the 3-way Age x Labelling x Shape interaction was not significant, $F(2, 88) = .26$, $MSE = 2.38$, $p = .77$. Thus, a similar pattern of data were found in this set of age analyses as in the previous set of age analyses.

Next, the effect of productive vocabulary size was examined. As Mauchley's test of sphericity was not significant for any effects, no degrees of freedom adjustments were made in the vocabulary median split omnibus analyses. The ANOVA revealed a significant main effect of shape, $F(2, 98) = 25.18$, $MSE = 211.06$, $p < .001$, and labelling, $F(1,49) = 8.66$, $MSE = 155.87$, $p < .01$, but not vocabulary, $F(1, 49) = .29$, $MSE = 5.27$, $p = .59$. Neither the Labelling x Shape, Vocabulary x Shape, nor the Vocabulary x Labelling interaction was significant, $F(2, 98) = 1.41$, MSE = 11.84, $p = .25$; $F(2, 98) = .79$, MSE = 6.66, $p = .46$; and $F(1, 49) = .54$, MSE = 9.71, $p = .47$, respectively. In addition, the 3-way Vocabulary x Labelling x Shape interaction was not significant, $F(2, 98) = .02$, MSE = 0.16, $p = .98$.

As no significant age or productive vocabulary effects were found through the use of median splits, a series of Pearson correlations were calculated between: (a) age and the frequency of target actions performed in the surprised condition (see Table 2), and (b) MCDI scores and the frequency of target actions performed in the surprised condition (see Table 3). None of these correlations reached significance, using an omnibus alpha level of .05.

Object Transfer Data Analyses

As in Experiment 1, instances of object transfer were examined. Those participants in Experiment 2 with test frequency outliers (n = 4; see data screening section) were not included in the object transfer analyses, as it was assumed that if participants were engaging in performing a very high frequency of appropriate target actions, they would not be engaging in meaningful object transfer. As in Experiment 1, even if participants did not perform target actions on any of the high, medium, or low test objects in the surprised condition, they were still included in the transfer analyses. Thus, the final sample size included 32 infants. In the surprised condition, instances of transfer were observed with 16 participants (50%), and in 23 out of 96 total possible trials (24%). Of the 23 trials in which object transfer occurred in the surprised condition, 9 were high shape object trials, 5 were medium shape object trials, and 9 were low shape object trials.

Incidences of object transfer in the no label condition (Experiment 1) were compared to incidences of object transfer the salient label condition (Experiment 2). A univariate 2 (labelling) x 2 (expectation) x 3 (shape) mixed-model ANOVA was performed on the data, with labelling (salient label, no label) as the between-subjects factor, and expectation (surprised, interest control) and shape (high, medium, low) as the within-subjects factors. An alpha level of .05 was used for all omnibus analyses. As Mauchley's test of sphericity was significant for the effect of shape, $\underline{W} = .80$, $\underline{p} < .01$, the Greenhouse-Geisser adjustment to degrees of freedom was made in determining the critical F-values reported for the main effect of shape. The ANOVA yielded a nonsignificant main effect of expectation, $\underline{F}$ (1, 63) = 3.47, $\underline{MSE}$ = 16.74, $\underline{p}$ = .07, and labelling, $\underline{F}$ (1, 63) = 1.63, $\underline{MSE}$ = 6.85, $\underline{p}$ = .21, but yielded a significant main effect of shape, $\underline{F}$ (1.67, 126) = 4.49, $\underline{MSE}$ = 8.92, $\underline{p} < .05$. The Labelling x Shape interaction was also significant, $\underline{F}$ (2, 126) = 5.20, $\underline{MSE}$ = 10.32, $\underline{p} < .01$. However, neither the Labelling x Expectation interaction nor the Shape x Expectation interaction was significant, $\underline{F}$ (1, 63) = .07, $\underline{MSE}$ = .35, $\underline{p}$ = .79; $\underline{F}$ (2, 126) = .14, $\underline{MSE}$ = .37, $\underline{p}$ = .87. The 3-way Labelling x Expectation x Shape interaction was not significant either, $\underline{F}$ (2, 126) = 2.08, $\underline{MSE}$ = 5.35, $\underline{p}$ = .13.

To follow-up the statistically significant Labelling x Shape interaction, a series of simple effects tests were conducted, comparing infants' performance in the no label to the salient label condition at each level of shape (high, medium, low). The alpha level was .05/3 = .0167. For the high shape object, there was no significant difference between transfer actions performed in the salient label condition ($\underline{M}$ = 0.88, $\underline{SD}$ = 1.60) and the no label condition ($\underline{M}$ = 0.27, $\underline{SD}$ = 1.3), $\underline{F}$ (1, 63) = 2.85, $\underline{MSE}$ = 5.89, $\underline{p}$ = .10. This was also the case for the medium shape object, with no significant difference between infants' performance in the salient label condition (M = .34, SD = .90) and the no label condition (M = 0.36, SD =1.20), $F(1, 63) = .01$, MSE = .01, $p = .94$, , as well as for the low shape object, with no significant differences between infants' performance in the salient label condition (M = 1.00, SD = 2.03), and the no label condition (M = 0.61, SD = 1.1), $F(1, 63) = .96$, MSE = 2.52, $p = .33$. Thus, at each level of shape, there were no significant differences between the number of transfer actions performed in the salient label and no label condition.

Latency to First Target Action Data Analyses

Thirty infants, 17 girls and 13 boys, (M = 19.47 months; SD = 1.04; range = 17.49 to 20.85 months) were considered in the frequency analyses.

As in Experiment 1, a Friedman's Rank Test was performed on the latency data for objects in the surprised condition in Experiment 2 (see Table 4 for mean latency scores). Ranks of latency scores were compared among high shape, medium shape, and low shape objects within the salient label condition. Only those participants who had a latency score for all three high, medium, and low shape objects were included in the analysis (N = 9). Of those participants with missing latency data (N = 21), 2 did not have latency scores for high shape objects, 10 did not have latency scores for medium shape objects, and 16 did not have latency scores for low shape objects. For each of the other 9 remaining participants, high, medium, and low shape object latency scores were ranked from lowest to highest (i.e., 1 to 3; lowest rank = shortest latency), and these ranks were summed across participants, for each level of shape. The sums of the ranks were as follows: high shape = 14, medium shape = 16, and low shape = 18. The Friedman's Rank Test results indicated that high, medium, and low shape objects' rankings were not significantly different from one another, $\chi^2_F(2,$ N

= 9) = -43.36, $p > .05$. Thus, as in Experiment 1, the relative rankings of the latency scores did not distinguish infants' inductive generalizations .

Discussion

The purpose of Experiment 2 was to examine whether infants will use label information to promote inductive inferences about nonobvious object properties when the label is made more salient in the experimental procedure. In Experiment 2, infants were presented with novel prototype objects that were labelled (e.g., "Look at this zas!") three times without the accompanying sound demonstration, as well as three times with the sound demonstration. Thus, the label was introduced twice as often as compared to the label condition in Experiment 1. In addition, test objects were introduced with a label four times instead of two times, as in Experiment 1. The differences between Experiment 1 and 2 in terms of the amount of attention given to objects may limit the strength of the conclusions that can be drawn from the results comparing the no label condition (Experiment 1) to the salient label condition (Experiment 2). However, the results of Experiment 2 do provide important preliminary evidence regarding whether infants can and will use label information as an inductive base.

Three main conclusions can be drawn from the results of Experiment 2. First, consistent with Experiment 1, infants in Experiment 2 performed significantly more target actions on test objects in the surprised and predicted condition than in the interest control condition, indicating that they made inductive inferences about test objects' properties on the basis of knowledge about the functionality of prototype objects.

Second, and also consistent with Experiment 1, when prototype objects were functional, high shape similarity promoted significantly more inductive generalizations of nonobvious object properties than either medium or low shape similarity. No differences in infants' inductive performance were found when the number of target actions attempted on medium and low shape objects were compared.

Third, a markedly different effect of object labels was found when the frequency of target action results of Experiment 2 (salient label condition) was compared to those of Experiment 1 (no label condition), within the surprised condition. At each level of shape (high, medium, low), infants generalized nonobvious object properties significantly more often in the salient label condition than in the no label condition. Importantly however, the results of Experiment 2 suggest that the presence of object labels did not lead infants to overlook shape information when making inductive generalizations. Infants still performed significantly more target actions on high shape objects than on either medium or low shape objects. Thus, object shape similarity appears to play a key role in infants' inductive strategies, even when conceptual information is available. However, as a no label condition was not present in Experiment 2, planned future work in which a no label condition, similar in all respects to the salient label condition except for the label, will likely provide incremental evidence that object labels promote inductive inferences in 16- to 21-month old infants.

As in Experiment 1, no differences in inductive performance were found as a function of infants' age or productive vocabulary size in Experiment 2. Also consistent with Experiment 1, no differences in inductive performance were found by comparing the ranks of the latency scores among levels of shape similarity within the surprised condition.

In sum, the results of Experiment 2 indicate that 16- to 21-month-old infants will use object label similarity to promote inductive inferences about nonobvious object properties, but not to the same extent as they will use object shape similarity in this endeavor. Experiment 2 provides an important replication and extension of the results of Experiment 1, in that when infants were taught novel labels for novel objects in the absence of a sound demonstration, it appears that they were more able to focus on the label information and use it as an inductive base. Regardless of whether infants were provided with information as to objects' shared conceptual identity, however, they still considered high shape similarity to play a critical role in the induction process.

General Discussion

The purpose of the present studies was to examine what drives inductive inferences in infancy. To this end, the role of object shape and object labels in promoting infants' inductive generalizations about nonobvious object properties was investigated. Taken together, the results of Experiments 1 and 2 produced three main findings. First, by 16 months of age, infants will form expectations about shared properties of novel objects after only a 10 second experience with a functional prototype exemplar, and infants will extend a specific nonobvious property from a prototype exemplar to other objects perceived as "in-category." Second, high shape similarity promotes inductive inferences about nonobvious object properties. Third, object labels influence infants' inductive judgements about shared nonobvious properties. Each of these three main conclusions will be discussed in turn, followed by an examination of their links to previous induction research with preschoolers and infants, and their implications for future work in this area.

To begin, Experiments 1 and 2 have demonstrated that infants will make specific inductive inferences about nonobvious properties of novel objects by 16 months of age. When infants were presented with prototype objects that were functional (i.e., in the surprised and predicted condition), and were then presented with test objects that were either: (a) also functional (predicted condition), or (b) nonfunctional (surprised condition), they performed significantly more target actions on test objects than when both prototype and test objects were nonfunctional (i.e., in the interest control condition).

These results provide important evidence that the sound properties possessed by functional objects in the present experiments were indeed nonobvious to infants; there was nothing in their appearances that inherently suggested the target nonobvious properties. These results also suggest that infants can make inductive inferences about objects which they have had no previous experience with outside of the laboratory. The results of the present experiments are consistent with those of Baldwin et al. (1993), and Mandler and McDonough (1996, 1998b), who have also shown that infants as young as 9 months, and 14 months, respectively, will make inductive inferences about object properties. However, the results of the present experiments go further than these previous studies, as inductive inferences about nonobvious properties of novel objects were examined. Mandler and McDonough's (1996, 1998b) research was based on the assumption that at least some of their stimuli were familiar to infants. Baldwin et al. also used stimuli which infants may have been already familiar with before they came in to the laboratory (e.g., a doll), and therefore these infants may have developed pre-existing expectations about the functionality of the stimuli used. In addition, some of Baldwin et al.'s nonobvious object properties (e.g., the doll's head and body come apart) may have actually been obvious, or expected by infants who were already familiar with the stimuli. However, the frequency data in the present experiments followed the same general pattern as Baldwin et al.'s, suggesting that an inferential ability extends to unfamiliar objects and their unobservable properties as well.

Although Baldwin et al. found that latency data differentiated infants' expectations about object properties, latency data in the present experiments could not do so because of the amount of data missing, and because of the lack of variance in the remaining latency data. These difficulties with the latency data in the present experiments are related to the fact that: (a) often the lower shape similarity objects (i.e., medium shape and low shape) did not promote inferences, and (b) infants were able to make rapid judgements about whether or not to generalize the nonobvious property (generally within 2 seconds).

The results of Experiments 1 and 2 also demonstrate that object shape is important in promoting inductive inferences. Infants were more apt to generalize a nonobvious object property if test objects were high in shape similarity to a prototype, than when test objects were medium or low in their degree of shape similarity to a prototype. These results are an important extension of Baldwin et al.'s (1993) study, providing incremental evidence that infants use high shape similarity as an inductive base, and also demonstrating that degree of similarity of a test object to a prototype object has an important influence on infants' inductive inferences. The results of the present experiments are consistent with a large body of empirical research which has shown that preschoolers and even adults will demonstrate a reliance on shape in various kinds of word learning situations (e.g., Graham & Poulin-Dubois, in press; Landau, Smith, & Jones, 1988, 1992, 1998; Smith, Jones, & Landau, 1992). A number of research studies have also shown that preschoolers and adults will use object appearances (shape) to make inductive inferences about objects' conceptual identities (e.g., Davidson & Gelman, 1990; Florian, 1994; Gelman & Markman, 1986, 1987; Lopez et al., 1992). Perceptual similarity may also facilitate infants' expectations that particular objects are of the same underlying kind. Indeed, the development of perceptual categories may help with ongoing acquisition of conceptual knowledge (Oakes, Coppage, & Dingel, 1997). Thus, shape should not simply be considered a superficial basis for inductive generalization, as it has the potential to provide important information about many kinds of other object characteristics, such as object kind and object functionality.

Given that the results of the present experiments are consistent with previous research which has demonstrated preschoolers' reliance on shape for promoting inductive inferences, it is not surprising that infants will also use high shape similarity, or the property of shape in general, to promote inductive inferences about nonobvious object properties. Shape is an easily detectable and visible object property. Unlike other perceptual properties which may not be as salient (e.g., texture, color, height, weight, density), shape is a prominent feature of any object. Shape also often does not vary among objects to the same extent as other perceptual properties (e.g., color, size) (Graham, Williams, & Huber, in press). Moreover, shape generally conveys information about functional features of an object, object identity, and underlying basic level category membership (McCarrell & Callanan, 1995; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). That is, things that share the same shape are often also the same kind of thing. For example, if something looks round, it is more likely to roll than to glide, and also more likely to be a ball than an airplane.

The results from Experiments 1 and 2 also indicate that infants can use object labels to promote inferences about nonobvious object properties. Infants extended a nonobvious property to other objects significantly more often when salient labels were provided for prototype and corresponding test objects (Experiment 2) than when no labels were provided for any objects (Experiment 1, no label condition). This is an important finding, as few studies have examined whether infants are capable of making conceptually-based inferences (but see Mandler & McDonough, 1996, 1998b). The present research extends Mandler and McDonough's (1996, 1998b), as infants in their studies were not taught labels for object stimuli, and thus had to rely on their previous knowledge about conceptual categories in order to make inferences on the basis of shared conceptual similarity. However, it is important to note that object labels did not allow infants to overlook perceptual appearances in the present experiments. That is, even when infants were taught labels for objects, they performed more target actions on the high shape objects than the medium and low shape objects (Experiments 1 and 2).

The finding that infants will not overlook perceptual similarity when making inductive inferences is inconsistent with the results of induction studies conducted with preschoolers. For example, Gelman and Markman (1986, 1987) and Gelman and Coley (1990) found that preschoolers made inductive inferences on the basis of shared labels, rather than shared perceptual appearances, when appearances and label information were pitted against each other. For example, in Gelman and Coley's study, children extended an unobservable property from a bluebird to dodo bird (same underlying kind as the bluebird, different appearance), rather than to a pterodactyl (same appearance as the bluebird, different underlying kind. Thus, children in Gelman and colleagues' studies were able to overlook perceptual appearances, whereas in the present studies, infants still relied heavily on object shape when making inductive inferences. Infants in Mandler and McDonough's research (e.g., Mandler & McDonough, 1996, 1998b) were also able to overlook perceptual similarities and make inferences on the basis of underlying conceptual similarity. For example, in Mandler and McDonough (1996), children extended an unobservable property of "being able to drink from a cup" from a bird to a cat (same underlying kind as the bird, different appearance), rather that to an airplane (same appearance as the bird, different underlying kind). However, infants in their studies were not provided with object labels to use in their inductions, and were provided with familiar objects in the induction tasks, rather than novel ones.

There are a number of factors that might account for the inconsistent results between the studies that have examined the effect of object appearances and object labels on preschoolers' inductions, and the present studies. Methodologically, it is possible that because three-dimensional objects were used in the present experiments (not pictures as in Gelman & Coley, 1990; Gelman & Markman, 1986, 1987), the shape of the objects was more salient than was their labels. Déak and Bauer (1996) found that labels were utilized more often as a basis for generalization when stimuli were presented in the form of line drawings rather than 3-dimensional object models. Perhaps because there was other salient physical information in the objects in the present experiments (e.g., color, texture, shape), infants were not as compelled to attend to the label information as they might have, had the objects been presented in picture form. As Déak and Bauer demonstrated, the manner in which stimuli are presented can determine the properties preschoolers will use as the basis for their categorization judgements. Thus, one can imagine that this would be even more pronounced for infants.

Another methodological difference between the present studies and those of Gelman and colleagues (e.g., Gelman & Coley, 1990; Gelman & Markman, 1986, 1987) is that conceptual and perceptual information were not directly pitted against each other in the present studies. A comparison of infants' inductive performance was made between when they were provided with perceptual and label information (Experiment 1, label condition; Experiment 2) and when they were only provided with perceptual information (Experiment 1, no label condition). In Gelman and colleagues' studies, when appearances and labels were not directly pitted against one another, infants had little difficulty using both shape and label information to make their inferences. For example, in Gelman and Coley (1990), 2 1/2-year-old children made significantly more correct property generalizations when perceptual appearances and labels were consistent (e.g., when they were asked whether a stegosaurus lived in a nest), than when they were inconsistent (e.g., when they were asked whether a pterodactyl lived in a nest). Thus, it is possible that if infants in the present experiments were asked to choose between performing target actions on test objects that were either: (a) similar in appearance but possessing a different label than the prototype, or (b) dissimilar in appearance but possessing the same label as the prototype, they would try target actions on those dissimilar-appearing, similar kind of objects first, and/or more often. Examining infants' inductive performance when labels and perceptual appearances are pitted against one another is an important direction for future research, as it is still unknown how infants will perform in this situation.

An additional explanation which may account for the difference in findings between previous preschool induction research and the present experiments in terms of the label effect is that the label may have connected target and test objects in a more plausible way in previous research. For example, in Gelman and Markman (1986), children extended an unobservable property from a shark to a tropical fish. These objects shared a number of similar shape features which may have enabled children to perceive them as plausibly connected in some way. It is possible that infants only made inferences on the basis of shape when appearances were not interpreted as contradicting the shared label information. This argument is consistent with Davidson and Gelman (1990), who examined 4-year-old children's inductions to pictured test objects provided with either novel versus familiar labels. They found that children only made inferences on the basis of a shared novel label when perceptual appearances of test objects were not extremely discrepant from a target (prototype) object. Thus, the plausibility of two objects sharing a label based on their level of perceptual similarity may also be an important factor for infants' inductions. In the present experiments, if infants were presented with a low or medium shape object, even if it was labelled with the same name as the prototype object, they may not have generalized the sound property if it did not seem likely, or plausible, for these objects to share a target nonobvious property with a prototype.

Finally, it is possible that the finding that object shape promoted inductive inferences to a greater extent than object labels in the present studies was due to infants' lack of experience and knowledge about labels and their conceptual informativeness. This account suggests that an understanding of the role of labels as indicative of object identity and as an inductive base may more fully develop between 21 months of age and 2 1/2 years of age. Importantly, however, this understanding is emerging even by 16 months of age, at the same time that infants have generally started to produce coherent language, and are able to understand a number of words (particularly nouns). As infants approach toddlerhood, it is likely that at the same time their productive and receptive language skills are becoming increasingly developed, they are also becoming much more aware and knowledgeable about how labels assist in a conceptual understanding about objects and their nonobvious properties. In addition, as infants mature and become more independent explorers, they also become more experienced with all kinds of objects in their environment. Thus, they may learn over time that perceptual appearances are not always an adequate basis for making inductive inferences, and that objects that share the same appearances may not necessarily be the same kind of thing, or perform in the same way.

Most of the previous inductive development research has focused on investigating whether young children are capable of more "mature" induction strategies (i.e., those based on conceptual similarity rather than perceptual similarity; see Gelman & Wellman, 1988 for a review). Another way of thinking about the level of sophistication in inductive abilities, however, is in terms of the versatility and flexibility of such strategies, rather than in the use of one form of induction over the other. As noted by Baldwin et al. (1993), the possession of a versatile reasoning mechanism enables infants to utilize new knowledge about a variety of objects and properties in the world, and to extend this understanding to other situations they encounter. Thus, the present experiments help to explain how young children are able to learn so much about the world so quickly, generating hypotheses about how the world works based on a small number of experiences, by using both perceptual and conceptual information for generating inductive inferences well before they reach preschool age. An ability to use a variety of perceptual and conceptual kinds of information, while considering the context, the task, the information available, and the potential efficiency and cognitive demands required to make certain inferences, may instead more fully characterize the sophisticated thinker. The results of the present studies suggest that aspects of this kind of "mature" thought are present by late infancy.

However, there are still many aspects of infants' inductive abilities that are not yet understood. Future research addressing when infants first show an ability to make inductive inferences needs to be more fully investigated. Baldwin et al. (1993) found evidence for inductive abilities by 9 months of age. However, the conditions under which infants use varying kinds of information (e.g., shape, labels, texture, dimensionality) remain to be delineated. In addition, in future work it will be important to determine which kinds of measures are most pertinent and useful for various types of inductive tasks (e.g., frequency of target actions, latency to first target action, object transfer). Although the present experiments have shed light on some of the features of objects that promote inductive inferences in infancy, another goal of future work is to examine infants' inductive performance when they are taught familiar, rather than novel labels for novel objects. It is likely that in such situations, familiar labels will promote more inferences than novel labels, as infants will orient themselves to the familiar label and attend to it more readily during the prototype demonstration (perhaps even during an accompanying sound demonstration). This prediction is consistent with other research which has shown that familiar labels promote inductive inferences, allowing young children to overlook perceptual appearances when extending a property (e.g., Davidson & Gelman, 1990; Gelman & Coley, 1990; Gelman & Markman, 1986).

In conclusion, the present experiments have advanced our understanding of young children's inductive abilities, demonstrating that by 16 months of age, infants are able to use shape and label information for making inferences about nonobvious object properties. Although infants were able to use many kinds of information about objects for inductive generalization (e.g., shape, texture, labels), shape was found to be an important property driving infants' inductive inferences. Future research delineating when and how infants utilize various kinds of object information in their inductive judgements will provide important insights into the mechanisms underlying the development of inductive strategies during late infancy.

References

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

Bloom, L. (1998). Language acquisition in its developmental context. In W. Damon, D. Kuhn, & R. S. Siegler (Eds.), Handbook of child psychology: Cognition, perception, and language (pp. 309-370). New York: Wiley.

Carey, S. (1985). Conceptual change in childhood. Cambridge: MIT Press.

Davidson, N. S., & Gelman, S. A. (1990). Inductions from novel categories: The role of language and conceptual structure. Cognitive Development, 5, 151-176.

Déak, G. O., & Bauer, P. J. (1996). The dynamics of preschoolers' categorization choices. Child Development, 67, 740-767.

Eimas, P. D. (1994). Categorization in early infancy and the continuity of development. Cognition, 50, 83-93.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, P. P., Pethick, S., & Reilly, J. S. (1991). MacArthur Communicative Development Inventories. San Diego: San Diego State University.

Flavell, J. H. (1985). Cognitive development (2nd ed.). Englewood Cliffs: Prentice-Hall.

Flavell, J. H., Miller, P. H., & Miller, S. A. (1993). Cognitive development (3rd ed.). Englewood Cliffs: Prentice-Hall.

Florian, J. E. (1994). Stripes do not a zebra make, or do they? Conceptual and perceptual information in inductive inferences. Developmental Psychology, 30, 88-101.

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

Gelman, S. A., & Coley, J. D. (1990). The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. Developmental Psychology, 26 (5), 796-804.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. Cognition, 23, 183-209.

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

Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious. Cognition, 38, 213-244.

Graham, S. A., & Poulin-Dubois, D. (in press). Infants' reliance on shape to generalise novel labels to animate and inanimate objects. Journal of Child Language.

Graham, S. A., Williams, L. D., & Huber, J. F. (in press). Preschoolers' and adults' reliance on object shape and object function for lexical extension. Journal of Experimental Child Psychology.

Howell, D. C. (1999). Fundamental statistics for the behavioral sciences (4th ed.). Pacific Grove: Brooks/Cole.

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

Keil, F. (1989). Concepts, kinds, and cognitive development. Cambridge: MIT Press.

Keil, F. (1991). The emergence of theoretical beliefs as constraints on concepts. In S. Carey & R. Gelman (Eds.), The epigenesis of mind (pp. 237-256). Hillsdale, NJ: Erlbaum.

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. Cognitive Development, 3, 299-321.

Landau, B., Smith, L. B., & Jones, S. S. (1992). Syntactic context and the shape bias in children's and adults' lexical learning. Journal of Memory and Language, 31, 807-825.

Landau, B., Smith, L. B., & Jones, S. S. (1998). Object shape, object function, and object name. Journal of Memory and Language, 38, 1-27.

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

Madole, K. L., & Oakes, L. M. (1999). Making sense of infant categorization: Stable processes and changing representations. Developmental Review, 19 (2), 263-296.

Mandler, J. M. (1988). How to build a baby: On the development of an accessible representational system. Cognitive Development, 3, 113-136.

Mandler, J. M. (1992). How to build a baby: 2. Conceptual primatives. Psychological Review, 99, 587-604.

Mandler, J. M. (1998). Representation. In W. Damon, D. Kuhn, & R. S. Siegler (Eds.), Handbook of child psychology: Cognition, perception, and language (pp. 25-308). New York: Wiley.

Mandler, J. M., & McDonough, L. (1993). Concept formation in infancy. Cognitive Development, 8, 291-318.

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

Mandler, J. M., & McDonough, L. (1998a). On developing a knowledge base in infancy. Developmental Psychology, 34 (6), 1274-1288.

Mandler, J. M., & McDonough, L. (1998b). Studies in inductive inference in infancy. Cognitive Psychology, 37, 60-96.

McCarrell, N. S., & Callanan, M. A. (1995). Form-function correspondences in children's inference. Child Development, 66, 532-546.

Moore, B., & Parker, R. (1989). Critical thinking: Evaluating claims and arguments in everyday life (2nd ed.). Mountain View, CA: Mayfield.

Oakes, L. M., Coppage, D. J., & Dingel, A. (1997). By land or by sea: the role of perceptual similarity in infants' categorization of animals. Developmental Psychology, 33 (3), 396-407.

Oakes, L. M., Madole, K. L., & Cohen, L. B. (1991). Infants' object examining: Habituation and categorization. Cognitive Development, 6, 377-392.

Poulin-Dubois, D., Frank, I., Graham, S. A., & Elkin, A. (1999). The role of shape similarity in toddlers' lexical extensions. British Journal of Developmental Psychology, 17, 21-36.

Poulin-Dubois, D., Graham, S. A., & Sippola, L. (1995). Early lexical development: The contribution of parental labelling and infants' categorization abilities. Journal of Child Language, 22, 325-343.

Quinn, P. C., & Eimas, P. D. (1986). On categorization in early infancy. Merrill-Palmer Quarterly, 32, 331-363.

Rips, L. J. (1975). Inductive judgements about natural categories. Journal of Verbal Learning and Verbal Behavior, 14, 665-681.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. <u>Cognitive Psychology, 8,</u> 382-439.

Sattler, J. M. (1992). <u>Assessment of children</u> (3rd ed.). San Diego: Author.

Shepard, R. (1987). Towards a universal law of generalization for psychological science. <u>Science, 23,</u> 1317-1323.

Shipley, E. F. (1993). Categories, hierarchies, and induction. <u>The Psychology of Learning and Motivation, 30,</u> 265-301.

Siegler, R. S. (1991). <u>Children's thinking</u> (2nd ed.). Englewood Cliffs: Prentice Hall.

Smith, L. B., Jones, S. S., & Landau, B. (1992). Count nouns, adjectives, and perceptual properties in children's novel word interpretations. <u>Developmental Psychology, 28,</u> 273-286.

Tabachnick, B. G., & Fidell, L. S. (1996). <u>Using multivariate statistics</u> (3rd ed.). New York: Harper Collins.

Tversky, B. (1985). The development of taxonomic organization of named and pictured categories. <u>Developmental Psychology, 21,</u> 1111-1119.

Wellman, H. M., & Gelman, S. A. (1988). Children's understanding of the nonobvious. In R. J. Sternberg (Ed.), <u>Advances in the psychology of human intelligence</u> (Vol. 4). Hillsdale, NJ: Erlbaum.

Appendix A

Adult Ratings for the Shape Similarity Judgements of High, Medium, and Low Shape Test Objects Compared to a Corresponding Prototype

Comparison	Mean Ratings	Paired t-test
High/Medium Shape Bell	H: 6.13. M: 3.87	t (14) = 6.86, p < .001
High/Low Shape Bell	H: 6.13 L: 1.67	t (14) = 12.76, p < .001
Medium/Low Shape Bell	M: 3.87 L: 1.67	t (14) = 5.98, p < .001
High/Medium Shape Rattle	H: 6.20 M: 3.45	t (14) = 4.77, p < .001
High/Low Shape Rattle	H: 6.20 L: 2.40	t (14) = 9.13, p < .001
Medium/Low Shape Rattle	M: 3.45 L: 2.40	t (14) = 2.26, p < .05
High/Medium Shape Ball	H: 6.33 M: 3.93	t (14) = 6.87, p < .001
High/Low Shape Ball	H: 6.33 L: 1.60	t (14) = 12.75, p < .001
Medium/Low Shape Ball	M: 3.93 L: 1.60	t (14) = 6.47, p < .001

Note. Ratings are out of 7 with 1 (not at all similar) to 7 (very similar).

Appendix B

Frequency and Latency Outliers in Experiments 1 and 2

Table B1

Frequency Outliers in Experiments 1 and 2.

ID	Labelling Condition	Object Shape	Expectation Condition	Object Set	Frequency Count
101	No Label	Low	Surprised	Bell	19
110	No Label	Low	Interest Control	Rattle	07
116	Label	Low	Surprised	Bell	21
		High	Interest Control	Rattle	02
121	Label	High	Interest Control	Bell	03
		Medium	Interest Control	Bell	09
145	Label	Low	Interest Control	Ball	04
165	No Label	High	Surprised	Bell	18
176	Label	High	Surprised	Bell	24
		Medium	Surprised	Bell	15
		Low	Surprised	Bell	16
		Low	Interest Control	Ball	06
209	Salient Label	High	Interest Control	Bell	04
222	Salient Label	High	Interest Control	Rattle	04
232	Salient Label	Medium	Interest Control	Rattle	04

Note. ID's starting with 1 in the first digit were in Experiment 1, and ID's starting with 2 in the first digit were in Experiment 2.

Table B2

Latency Outliers in Experiments 1 and 2.

ID	Labelling Condition	Object Shape	Expectation Condition	Object Set	Latency Time
131	No Label	High	Surprised	Bell	08.43
150	No Label	High	Surprised	Bell	08.38
159	Label	Low	Surprised	Rattle	16.34
165	No Label	Medium	Surprised	Bell	09.81
207	Salient Label	High	Surprised	Rattle	14.03
227	Salient Label	Medium	Surprised	Bell	09.67

Note. ID's starting with 1 in the first digit were in Experiment 1, and ID's starting with 2 in the first digit were in Experiment 2. Latency time was in seconds.