INFANTS' ABILITY TO USE LANGUAGE AS A GUIDE TO INDUCTIVE INFERENCE

by

RENÉE NICOLE DESJARDINS

B.A. (Honours) Queen's University, 1985

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

Department of Psychology

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

November 1992

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature)

Department of Psychology

The University of British Columbia
Vancouver, Canada

Date December 10/92

DE-6 (2/88)

ABSTRACT

The present studies were designed to address whether infants use language to guide their inferences about novel objects. In Study 1, forty-eight infants from two age groups (16-18 and 20-22 months) participated in a task which used exploratory play as a window on inferential abilities. Infants explored novel toys presented in pairs: The first toy produced an interesting but non-obvious property while the second toy was invisibly altered such that it failed to produce the target property. In the Labeled condition both toys were moderately similar in appearance and given a common label. In the Unlabeled Control condition moderately similar toys were talked about in a general way as were dissimilar toys in the Baseline Control condition. The frequency of infants' attempts to elicit the target property was marginally higher in the Unlabeled Control condition compared to the Baseline Control condition. This suggests that infants may have, with difficulty, perceived the greater degree of similarity between moderately similar toys in the Unlabeled Control condition compared to the dissimilar toys in the Baseline Control condition. When toys were labeled, infants across the 16-22 month range attempted to reproduce the target property with the second toy more frequently than in the Unlabeled Control condition indicating that they used the object labels to guide their inferences about object properties. Study 2 compared the frequency of infants' attempts to elicit the target property from the second toy when two toys were given a common label (Same Label condition) as opposed to different labels (Different Label condition). Eighteen 20- to 22-month-olds showed a marginally higher frequency of target actions in the Same Label condition compared to the Different Label condition suggesting that they understood that two toys should share a <u>common</u> label before that label can be used to guide expectations about object properties. Together these studies provide evidence that infants

from 20 months of age understand that object labels can be used as a guide to inference, and some suggestion that this ability may be present from 16 months.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENT

I would like to thank the parents and infants who participated in these studies. Thanks to CheeSiah Chang, Judith Hole, and Siamac Saatchi for their help with coding. Also, to Brigitte Bill, Jane Irwin, Ruth Kirson, and Glynnis Tidball, thank-you for all your moral support and help during data collection. The biggest thank-you of all goes to Dare Baldwin for all her help along the way: intellectual guidance, emotional support and material help. And last, but not least, thank-you to Ross McKitrick for being there for me.

INTRODUCTION

Everyday object categories, such as dog and chair, tend to group together objects which have a variety of physical properties in common. Quite often, however, the properties shared by category members cannot be determined by visual inspection alone; they occur at a deeper, non-obvious level. For example, one important feature of the category "computer" is a CPU (Central Processing Unit) which is not readily observable. One way in which object labels function is to identify members of a category which otherwise might not be obvious from their surface appearance (e.g., a radio which looks like a Coke can). Given knowledge that labels function in this way, hearing two objects given a common label may lead one to search for the non-obvious similarities which are indicated by the common label. There is evidence that preschool-aged children already use language in this way to structure their thinking and expectations about the object world. However, at present little is known about the development of this ability to use language for conceptual purpose. The present research explores this developmental question.

Current evidence suggests that children of preschool age already appreciate that objects can share hidden properties and they are fairly sophisticated at using both perceptual and linguistic information to guide their conjectures about such non-obvious commonalties. For example, perceptual features such as wings suggest that a creature can fly even if that creature is currently immobile. Massey and Gelman (1988) have shown that children of 3 to 4 years of age can go well beyond such simple perceptually-based inferences. They showed children photographs of unfamiliar mammals, non-mammals and statues with animal-like forms and parts, asking them whether the photographed animal could move up and down a hill by itself. Children were remarkably accurate on this task even when the stimulus was an

unfamiliar animal; they predicted that mammals and non-mammals could move up a hill by themselves. At the same time they predicted that statues which shared physically similar properties (e.g., legs, etc.) could not, indicating that their predictions were based on a complicated set of perceptual cues and were not just a simple prediction from visibly perceptible parts.

In many cases the cues of surface similarity and object labels may converge to indicate that the objects in question share deeper, non-obvious properties. However, some categories exist for which perceptual information is either not as useful or even misleading as an indicator of membership (e.g., whales are mammals although they resemble fish). When these perceptual versus linguistic cues lead to opposite predictions about category membership, preschool children are capable of relying on object labels to guide their inferences about category membership. Gelman and Markman (1986) pitted surface similarity cues against object labels in order to determine whether preschool children would place more importance on the category label or perceptual similarity. Children were shown two familiar objects and taught a new fact about each one (e.g., "This squirrel eats bugs. This rabbit eats grass"). Then they were asked which property applied to a third object which looked like <u>one</u> of the objects (e.g., the rabbit) but was given the same category label as the <u>other</u> object (e.g.,"squirrel"). For example, children were asked, "Does this squirrel eat bugs like this squirrel or grass like this rabbit?" If children's inferences were based on perceptual cues they should choose the property belonging to the similar-looking object. On the other hand, if children understand that objects with common labels have similar non-obvious properties, then they should choose the property belonging to the object with the same label as the target object. Gelman and Markman found that children inferred that the target object had the same property as the object with the same

label. Hence, these children demonstrated an understanding that object labels can be informative about non-obvious object properties. In a simplified design, Gelman and Markman (1987) replicated this result with three-year-old children.

In a series of studies similar in design to the Gelman and Markman (1986; 1987) research, Davidson and Gelman (1990) also found that preschoolers tend to use novel labels to guide their inferences about category membership. However, in certain circumstances when the linguistic information was inconsistent with children's perceptually-based partitioning of the target and test objects, they relied on perceptual similarity to guide their inferences. If children saw an object which looked similar to the target object but was given a different label and saw another object which looked different from the target but was given the same label, children did not consistently choose the object with the same label as having the same property as the target object. In this case, children ignored object labels and relied on similarity since the experimenter's use of labels was inconsistent with respect to surface cues. Thus, although preschool children are capable of relying on labels to guide their inferences about non-obvious object properties, they will not do so if labels are entirely inconsistent with a perceptually-based nonlinguistic partitioning of the objects in question.

As previous research demonstrated that preschoolers are able to use object labels to guide their expectations about object properties (when labels are not entirely inconsistent with respect to surface cues), Gelman and Coley (1990) investigated whether even younger children's inferences showed this pattern. They found that at 2 1/2 years of age, children accurately inferred an object's property more often when the test and target objects were given the same label than when they were not. For example, children were told about the target picture that it was a bird and lived in a nest. If they were then shown the

test picture, children were more likely to agree that it lived in a nest if they were told the picture was of a bird than if neither target or test pictures were labeled. This study was replicated using temporary state adjectives (e.g., "This is sleepy") instead of object labels in order to determine whether the effect was due merely to a bias to attend to the word associated with each picture. In this case, children's accuracy did not differ from chance indicating that they only use object labels to guide their inferences and not temporary state adjectives. These studies indicate that by 2 1/2 years of age children are already using object labels to guide their expectations about object properties. Furthermore, they understand that object labels per se, not temporary state adjectives, can be used in this way.

To date however, there is no evidence to clarify whether children under the age of 2 1/2 use labels to guide their inferences about non-obvious object properties. Existing methodologies for testing inferential abilities have a strong verbal component and thus cannot be used for children under 2 years of age. Yet, there is reason to suspect that they may be able to do so. Evidence for two prerequisite abilities for this competency can be found in younger children: comprehension of object labels and the ability to draw simple inductive inferences.

As early as 9 months of age, infants begin to show signs of comprehending words (Oviatt, 1980; 1985). In the next 6 to 12 months they develop an even larger comprehension vocabulary and a productive vocabulary of 30 or more words (Benedict, 1979; Goldin-Meadow, Seligman & Gelman, 1976; Huttenlocher, 1974, Nelson, 1988). At approximately 18 months, some infants show dramatic increases in their vocabularies--the"naming explosion"--adding in some cases even as many as 10-20 new words to their productive vocabulary in a single day (Gopnik & Meltzoff, 1986;

Nelson, 1977; 1991). Thus, it seems that before the age of two children may have at least the requisite linguistic competency for using language to guide inferences. Furthermore, research suggests that they also have the ability to draw simple inductive inferences.

One method for determining infants' inferential abilities is by examining their exploratory play. If after exploring an object with a certain non-obvious property, infants expect a second similar-looking object to have the same property, this should be manifested in their attempts to elicit that property. Using such an exploratory play methodology, Baldwin, Markman and Melartin (in press) have demonstrated that infants as young as 9 months are able to draw inferences about non-obvious properties after only a brief exposure to a novel object. In one condition (violated expectation), infants explored a novel object with a non-obvious and interesting property for 30 seconds, then they explored a second toy similar in appearance but invisibly altered such that it failed to produce the interesting property. Because the second toy failed to produce the property, infants' actions to elicit the property could be attributed to a prior expectation. A second condition (boredom) in which infants explored two similar novel objects neither of which produced an interesting property was included to distinguish actions which might result from boredom at having received two similar toys in a row from true inferences in the violated expectation condition. A third condition (different) was included in which infants explored a novel object with a non-obvious and interesting property followed by a second toy dissimilar in appearance which failed to produce an interesting property. In both the boredom and different conditions infants explored novel toys which failed to produce an interesting non-obvious property; these two conditions provided a base-rate of the frequency with which infants produced the target actions on a novel toy through miscellaneous exploration. Infants

made a first attempt more quickly and made more frequent subsequent attempts to elicit the target property from the second toy in the violated expectation condition than in either the boredom or different conditions suggesting that infants were able to use perceptual similarity to guide their expectations about the non-obvious object properties.

Altogether it appears that children under 2 years of age may well have the necessary prerequisites of language comprehension and inferential abilities in order to be able to use object labels to guide their inferences. However, it is possible that the ability to rely on object labels to guide inferences about non-obvious object properties is still beyond their capabilities. Perhaps they are not able to exploit their knowledge of object labels for conceptual purpose. For example, children's early words may be simple associations between a word and an object. Nelson (1985, 1991) suggests that associative word-to-object links may be characteristic in the early phases of word learning. She uses the term "referential" to describe these early meanings in which labels make little contact with children's conceptual knowledge. If infants' comprehension of labels is comprised of simple word-to-object links, we would not expect them to be able to use object labels to guide their inferences about non-obvious object properties. If on the other hand, their understanding of object labels is such that meanings make contact with conceptual knowledge--Nelson refers to this as denotational meaning and expects meanings to become increasingly denotational with age--then we might expect that infants may indeed be capable of using language for inferential purposes.

The following studies make use of the Baldwin, Markman & Melartin (in press) exploratory play methodology to examine whether children under 2 years of age are able to use object labels to guide their inferences about non-obvious object properties. Specifically, will hearing the same label applied to

somewhat different looking objects lead infants to search for shared non-obvious properties? If so, infants should demonstrate their expectations by their manner of exploring the second toy in a pair. If they expect the second toy to produce the same non-obvious property as the first toy, they should be more persistent in attempts to elicit the target property when the two toys are given a common label than when they are unlabeled.

Infants in two age groups, 16-18 months and 20-22 months, had the opportunity to explore a number of novel toys in pairs. Infants between 20 and 22 months were chosen as previous research indicates that by this age children have acquired a sizable comprehension vocabulary of object labels (Nelson, 1977). Furthermore, Gopnik and Meltzoff (1987; 1992) observed a relationship between the onset of the naming explosion and an interest in categorizing objects in infants of 18 months. For infants from 15 to 21 months of age, vocabulary size and performance on categorization tasks were obtained. A significant relationship first appeared between vocabulary size and certain categorization tasks (e.g., basic level sorting) at 18 months. Thus, we might expect that using language to guide inferences might appear also during this period and would thus be readily detectable in infants between 20-22 months. A group of younger infants (16- to 18-month-olds) was chosen to examine whether there is a relationship between age and/or vocabulary size and infants' inferential abilities. Perhaps younger infants might be less adept at using language to guide inferences.

STUDY 1

In Study 1, infants explored novel pairs of toys in each of 3 conditions: Labeled, Unlabeled Control and Baseline Control. In each pair of toys, the first toy produced a novel, non-obvious property while the second toy was invisibly altered such that it failed to produce the target property. As in Baldwin et al. (in

press), the second toy in each pair was disabled in order to show that infants' attempts to reproduce the target property are the result of inferences and not merely the result of a rewarding object property. Furthermore, novel toys were used in order that infants' expectations about the second toy would be based on one previous exposure (the first toy) rather than resulting from more general background knowledge.

Toys in a pair were chosen to be only moderately similar in appearance (although capable of the same non-obvious property). This is in contrast to the Baldwin et al. studies in which toys were quite similar in appearance. In the present study, if toys were too similar in appearance, object labels would not provide any additional cues above those given by perceptual similarity as to the underlying non-obvious property shared by the toys.

A comparison between the Labeled and Unlabeled Control conditions is of most interest in the present context. If infants understand that toys with a common label share the same non-obvious property they should expect the second toy in a pair to possess the same non-obvious property as the first toy. Thus, they should quickly attempt to reproduce the the target property and persist in their attempts. Such expectations should not necessarily arise when toys are not labeled. Furthermore, a Baseline Control condition was included which paired toys dissimilar in appearance. This provided a measure of the frequency with which infants produce the target action on the second novel toy through miscellaneous play. If infants' target actions on the second toy are more frequent in the Unlabeled Control condition than in the Baseline Control condition, this would indicate that infants' exploration was the result of inference and not merely the result of miscellaneous play.

The Baseline Control condition also allowed us to examine whether infants were likely merely to repeat on the second toy the target action which

they employed with the first toy even in the absence of perceptual similarity. For example, if the property of the first toy was a light activated by a pressing a button, would infants show pressing actions on a periscope regardless of the obvious lack of similarity? If instances of such transfer are rare, this would suggest that target actions in the Unlabeled Control condition were indeed the result of inferences based on perceptual similarity and not merely the perseveration of a successful action pattern.

Although exploratory play behaviour may be indicative of infants' expectations about toy properties, visual comparison of the toys may reveal their expectations as well. Mandler (1988) reviews a number of studies in which pairs of objects are available for visual inspection. She notes that from around 8 months of age infants look back and forth between a novel object and a familiar object in a comparison process which appeared to be different from a simpler process of recognition observed at younger ages. In the present study, infants had the opportunity to look at the first toy in a pair while exploring the second toy; the frequency of such looks may be indicative of infants' attempts to discern similarity between the two toys. If hearing two toys labeled increases infants' expectations that the two toys share the same non-obvious property this might be revealed in more frequent looks in the Labeled condition to the first toy in order to compare the toys and their properties than in either the Unlabeled Control or Baseline Control conditions.

Method

Subjects

Sixty-eight infants in two age groups (16- to 18-month-olds and 20- to 22-month-olds) participated. Twenty infants were eliminated, due to fussiness (9), prior familiarity with the stimuli[1] (5), design criteria (two or more trials eliminated due to fussiness and familiarity of stimuli) (3), and experimenter error (3),

yielding a final sample of 48 infants: 24 infants in each of the two age groups. All infants were from monolingual English-speaking families, were developing normally and had no history of serious ear infection. Infants were recruited through newspaper advertisements and postnatal information sessions at maternity hospitals in the area.

Materials

Stimuli. Two toy sets were used, comprised of four pairs of toys each. Toy pairs were chosen to be novel, manipulable, interesting and visually distinctive from other pairs. Each toy pair consisted of two toys only moderately similar in appearance, although capable of the same non-obvious property. For example, one pair consisted of a janitor's retractable key-ring and a plastic retractable measuring tape. Toy pairs were assigned to sets such that the pairs in a set would be maximally distinctive. See Table 1 for a list of the toy pairs that were used, their target properties and the actions required to elicit the properties.

For all three conditions, the second toy in each pair was disabled such that it failed to produce the target property but showed no surface evidence of this alteration. For example the measuring tape was glued so that pulling on the metal tab failed to release the tape from the housing.

Equipment. A time-stamped video record provided continuous information about the flow of events.

Design

Each infant explored pairs of toys in three conditions: Labeled, Unlabeled Control and Baseline Control.

Labeled condition. Infants explored the first novel toy in a pair which was capable of producing a non-obvious property. Then infants explored a

Table 1

Toys, Their Properties, and Target Actions for Study 1

First Toy	Second Toy[a]	Target Property	Target Action
Set 1:			
toothbrush case	corkscrew	hidden stickers	remove inner portion
little bobber (handleless)	big bobber (long handle)	hinge movement/clacking	lift & drop arm piece
round magnifier	square magnifier	centre emerges	hinge tab
collapsible cup	periscope	expands/contracts	pull apart ends
Set 2:			
suction cup toy	suction soap pad	adheres/popping sound	suction against table & lift
sharpener	telescope	contracts/expands automatically	depress lid
janitor's key-ring	measuring tape	hidden chain/ automatic retraction/ 'zippy' sound	pull tab to extend & release
square flashlight	cylindrical flashlight	light flashes	press on button

[a] The second toy is disabled such that it fails to produce the target property.

second moderately similar novel toy which was unable to produce the non-obvious property. Both toys were given the same novel label.

Unlabeled Control condition. As in the Labeled condition, infants explored pairs of novel toys which were moderately similar in appearance. The first toy produced a non-obvious property whereas the second toy failed to produce the non-obvious property. Toys were talked about in a general way but were not labeled.

Baseline Control condition. Infants explored a novel toy which was capable of producing a non-obvious property. Then they were given a second novel toy dissimilar in appearance to the first toy which failed to produce the non-obvious property. Toys were talked about in a general way but were not labeled. Dissimilar pairs were generated by taking the first toy from one pair and combining it with the second toy from another pair (e.g., a flashlight was paired with a measuring tape). Only four dissimilar pair combinations per set were used out of a possible twelve so that the second toy in a pair appeared equally often in each of the three conditions.

All possible orders of the three conditions were used. Order of conditions and toy pairs were counterbalanced such that each toy pair appeared equally often in each condition (i.e., Labeled, Unlabeled Control, Baseline Control). Infants each participated in two blocks of trials. Each block consisted of one trial of each condition. Toy sets were randomly assigned to the blocks such that for half of the subjects toy set 1 was assigned to the first block and for the other half of the subjects toy set 2 was assigned to the first block. Novel labels (i.e., peri and toma) were roughly counterbalanced in assignment by block, toy set and toy pair.

Each block was preceded by a pair of novel identical toys given the same novel label which were both capable of producing the same interesting and

non-obvious property. One pair consisted of identical magnifying glasses on a handle which bent in the middle; the other pair consisted of binoculars with several pieces which moved on hinges. No data were collected from infants' exploration of these toys. These "interim" pairs were included only in order to maintain infants' interest in the game and to prevent infants from forming a response set that might undercut their exploration of the toys (i.e., a notion that the second toy never works). The interim pair assignment was counterbalanced by block and toy set; the assignment of labels, nup and bik, to these interim toys was counterbalanced by interim toy pair and block.

Procedure

After a ten minute warm-up period, the infant was seated in an infant seat which was attached to a table. The experimenter knelt on the floor on the other side of the table facing the infant. The infant first explored an interim toy pair followed by three toy pairs, one in each of the three conditions: Labeled, Unlabeled Control, and Baseline Control. Then infants explored another interim pair before exploring three more toy pairs, one in each of the three conditions.

First, infants were given one of the toys from the interim pair. For the interim toy pairs and for toys in the Labeled condition, the toy was labeled six times in motherese (e.g., "I'm going to show you a toma! You want to see the toma? You can play with the toma." etc.) The experimenter also demonstrated the property of the toy before handing it over to the infant (e.g., how to make the handle of the magnifying glass bend). Labeling did not occur during the demonstration in order to help prevent infants from associating the novel label with the action rather than the object itself. During the 30s toy exploration period that followed, the experimenter did not talk about the objects or make eye-contact but let infants play as they wished without interference. When the

30s period was up, the experimenter retrieved the toy and placed it on the table top out of infants' reach where it remained for the duration of infants' second toy exploration. This technique was used to reduce possible memory demands, because infants were able to look at the first toy as an aid to remembering its property and the target action required.

For the second toy, the experimenter did not demonstrate the toy's property. For the interim pairs and toys in the Labeled condition the toy was labeled six times before it was handed over to the infant. During the 20s toy exploration period, the experimenter again refrained from talking about the objects or interfering with infants' play. When the 20s period was up, the experimenter retrieved the toy and also collected the first toy from the corner of the table. Both toys were put away and the whole procedure was repeated with a new pair of toys.

The only difference between the Unlabeled Control and Baseline Control conditions was the relation between the first and second toys in a pair; toys were moderately similar in the Unlabeled Control and quite different in the Baseline Control condition. In all other respects the procedures were the same. Toys in both the Unlabeled Control and Baseline Control conditions were presented in the same fashion as in the Labeled condition except the experimenter talked about them with six general phrases instead of six labeling phrases. For example, the experimenter said, "I'm going to show you a new one. You want to see it? You can play with it!", etc. Please see Table 2 for the exact list of phrases used in the different conditions.

Coding

Target Action Coding. Coders reported the frequency of infants' target actions with the first and second toys in the Labeled, Unlabeled Control and Baseline Control conditions. The coding categories for each target action

Table 2

Carrier Phrases Used in Study 1

Labeled	Unlabeled Control & Baseline Control
First toy:	
I'm going to show you a toma!	I'm going to show you a new one!
You want to see the toma?	You want to see it?
You can play with the toma.	You can play with it.
Look! a toma!	Look! yes!
You want to play with the toma?	You want to play with it?
Okay, here's the toma.	Okay, here it is.
Second toy:	
You want to see the next toma?	You want to see the next one?
I've got a toma!	I've got it!
I'm going to show you the next toma.	I'm going to show you the next one.
You want to see this toma?	You want to see this one?
Look at this toma.	Look at this one.
A toma!	Ah yes!

were different for each toy pair (see Table 1). The time at which each of these actions occurred was recorded from the time-stamped video record. In this way, latency to the first target action from the time the toy was presented could be calculated.

Coders were blind to experimental condition. This was accomplished by having a different individual record the time at which the toy was presented. Using these times, coders then covered all the television screen except for the time-line and advanced to the beginning of the play period before removing the cover. In this way, coders did not see which toy the infant had previously explored and thus did not know if it was similar to or quite different from the second toy. Furthermore, coding was completed with the sound off, so that coders could not hear whether the toy had been labeled or not.

The reliability of coders' judgments was determined by having another coder who was also blind to experimental condition recode all conditions from eight randomly-selected subjects, resulting in 48 20s trials altogether. Percent agreement was calculated by dividing the number of instances on which coders agreed by the total number of instances observed. Coders demonstrated 88% agreement on the frequency of target actions.

Transfer Coding. Instances of transfer of target action from the first toy to the second toy in the Baseline Control condition were coded by having the coder observe the infant's actions on the first toy. The number and nature of actions on the second toy that appeared similar to target actions performed on the first toy were noted. For example, if pressing a button elicited the target property on the first toy, a flashlight, an infant might also demonstrate a pressing action on the second (dissimilar) toy, a measuring tape. The coder also rated the degree of similarity (1=not very similar to 5=very similar) and confidence (1=very uncertain to 5=very certain) that a transfer of action was observed. A

relatively liberal criterion was used in judging instances of transfer: actions of only moderate similarity (3 and above) about which the coder was only moderately confident (3 and above) were counted as instances of transfer[2]. An alternative explanation for what might appear to be inferential behaviour is that infants merely repeat the last action which produced a successful result. In order to give such an explanation a fair test any plausible instance of perseverative action was included.

Looks Coding. Coders, blind to condition, recorded the number of times infants looked at the first toy resting on the corner of the table while the second toy was available for exploration. To ensure that coders remained blind to condition, the corner of the screen where the first toy could be seen was covered. As well, coding was completed with the sound off so that coders could not hear whether the toy had been labeled. Coders also judged whether the infant seemed to look at the first toy for the purpose of comparing the first and second toys.

The reliability of coders' judgements concerning looking patterns was determined by having another coder, blind to experimental condition, recode all conditions from eight randomly-selected subjects, resulting in 48 20s trials altogether. Percent agreement was calculated by dividing the number of instances on which coders agreed by the total number of instances observed. Coders demonstrated 82% agreement on frequency of looks and 82% agreement on judgements of "comparative" looks.

Results and Discussion

Two measures of inferential ability were obtained from infants' exploration of the second (disabled) toys: the number of times infants produced the target action required to elicit the non-obvious property and latency to the first target action. If infants perceive the similarity between the two toys in the

Unlabeled Control condition, they should produce more target actions and show a shorter latency to their first action in the Unlabeled Control condition than in the Baseline Control condition.

Furthermore, if infants are able to use object labels to guide their inferences about the non-obvious object properties, infants should show more target actions and show a shorter latency to their first target action in the Labeled condition relative to both the Unlabeled Control and Baseline Control conditions.

Only trials in which infants demonstrated the target action on the first toy were included in the analyses. Unless infants displayed at least one action on the first toy, we could not be sure that they had noticed the toy's non-obvious property. If infants had not noticed the property, then they would be unable to draw inferences about a second toy's non-obvious property. Hence, only those trials on which infants attempted the target action with the first toy of the pair were analyzed. This happened rarely: only six out of a total of 288 trials were eliminated because of failure to attempt the target action on the first toy. And an additional ten trials were eliminated because of fussiness or prior familiarity of toys. When trials from both blocks were available, the frequency and latency scores were averaged across the trials in separate blocks. When only one trial was available, one score was used in place of an average score for that infant.

The raw data included two extreme scores (one 4.5 standard deviations and the other 6 standard deviations from the mean), leading us to remove the outliers and submit the remaining data to a mixed design MANOVA with Age as a between-group factor and Condition as a within-subject factor. Furthermore, a natural log transformation of the raw data including extreme scores was carried out and the transformed data were submitted to a mixed design MANOVA. As the pattern of results remained the same using the transformed data, only

results from the analyses using raw data with extreme scores removed are reported here. Non-parametric tests with the outlying scores included were used as an additional strategy to deal with a distribution skewed by the inclusion of extreme scores.

Both frequency and latency measures were included in a mixed design MANOVA with one between-group factor, Age (16-18 months and 20-22 months) and one within-subject factor, Condition (Labeled, Unlabeled Control, & Baseline Control). This analysis revealed a significant main effect of Condition, F (4, 43) = 3.28, p < .02, but no significant main effect due to Age[3]. Univariate analyses of frequency, F (2, 45) = 6.26, p<.005, and latency, F(2, 45) = 4.03, p<.05, also revealed a significant main effect of Condition. No significant interactions were found between Condition and Age in either the multivariate or univariate analyses. The average frequency of target actions and the average latency to first attempt are displayed with respect to age and condition in Figures 1 and 2.

The significant multivariate and univariate condition effects lead to a number of questions as to the locus of the effect; these questions will be taken up in turn.

Effect of Similarity

The first question of interest is whether infants noticed the greater similarity between the moderately similar toys in the Unlabeled Control condition compared to the dissimilar toys in the Baseline Control condition. Multivariate planned contrasts with one within factor, Condition, revealed a marginally significant difference between the Unlabeled Control condition compared to the Baseline Control condition in frequency of target actions and latency to the first action, Hotelling's t (45) = 1.50, p<.06 (one-tailed). Univariate contrasts indicated that the frequency of infants' target actions was greater in

Figure 1. The frequency of target actions on the second toy as a function of condition. (Study 1)

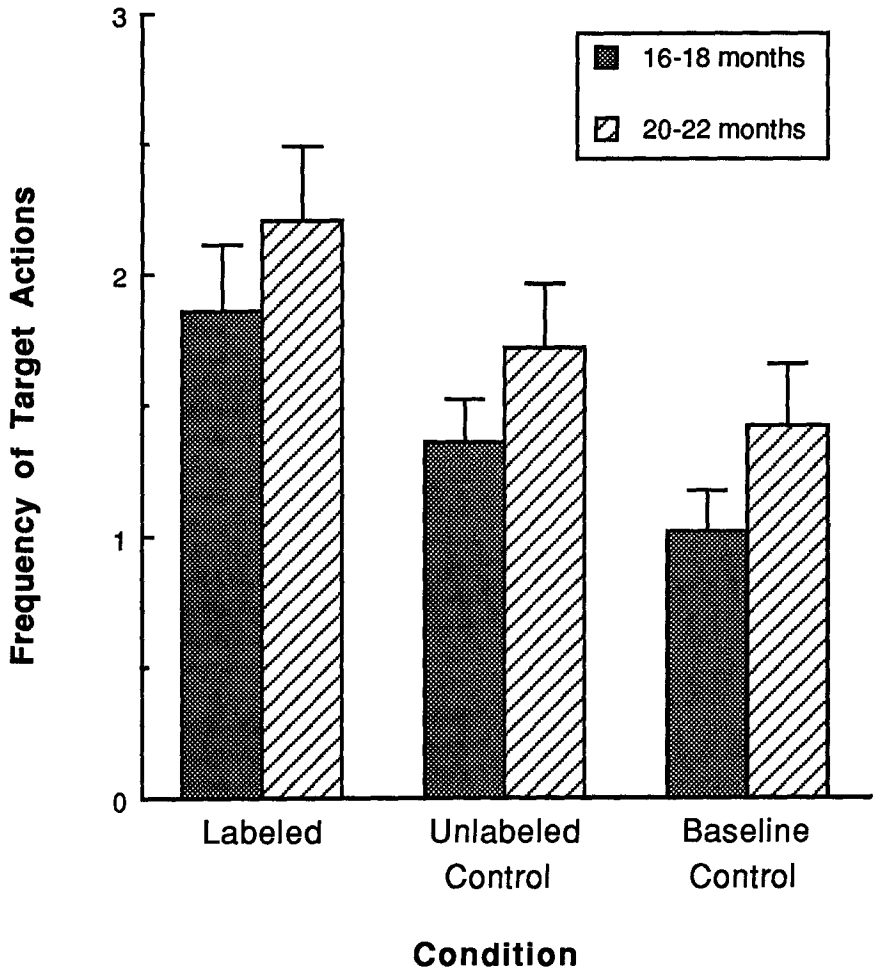
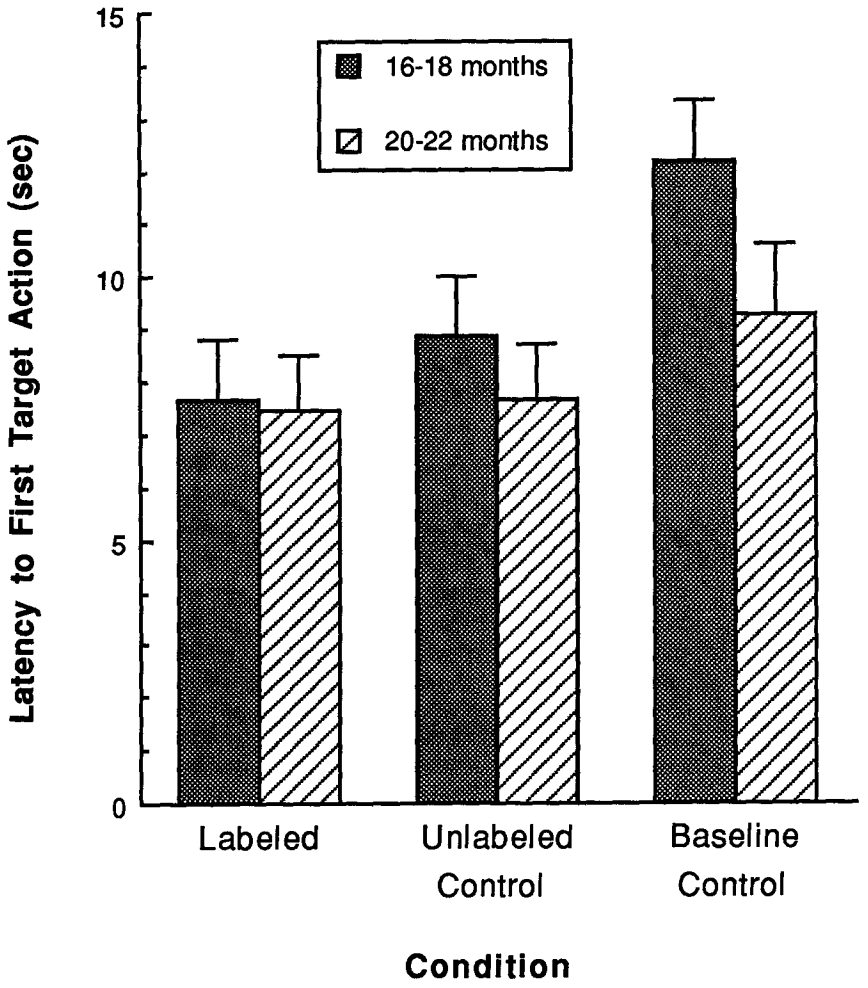


Figure 2. Latency to infants' first target action on the second toy as a function of condition (Study 1)



the Unlabeled Control condition compared to the Baseline Control condition, t (46) = 1.70, p<.05 (one-tailed). As well, the latency to infants' first target action was shorter in the Unlabeled Control condition than in the Baseline Control condition, t (46) = 2.13, p<.02 (one-tailed). Wilcoxon Matched-Pairs Signed-Ranks tests confirmed that the frequency of target actions was greater in the Unlabeled Control condition compared to the Baseline Control condition, z = -2.26, p<.02 (one-tailed), and that the latency to the first target action was shorter in the Unlabeled Control condition compared to the Baseline Control condition, z = -2.10, p<.02 (one-tailed). The greater frequency of target actions in the Unlabeled Control condition (M = 1.53 target actions, SD = 1.05) compared to the Baseline Control condition (M = 1.22 target actions, SD = 0.99) and the shorter latency to the first target action in the Unlabeled Control condition (M = 8.25 seconds, SD = 5.42) compared to the Baseline Control condition (M = 10.73 seconds, SD = 6.31) suggest that infants noted the greater similarity between the moderately similar toys in the Unlabeled Control condition and thus drew inferences about the non-obvious object properties based on the perceived similarity.

Because some trials had to be eliminated due to an absence of attempts to elicit the target property with the first toy, toy set by condition counterbalancing was not maintained. We were concerned that the observed condition effects might be confounded by toy effects. Thus, Condition differences were also examined across toys as opposed to across subjects (i.e., toy was treated as a random factor). Multivariate planned contrasts with one within factor, Condition, yielded no significant difference between the Unlabeled Control and Baseline Control conditions on frequency and latency. Univariate contrasts were also non-significant indicating no significant difference in either frequency or latency to the first target action in the Unlabeled Control condition

relative to the Baseline Control condition. This suggests that the obtained effects using subjects as a random factor must be interpreted with caution.

Effect of Labeling

The second question of interest is whether infants used the object labels in the Labeled condition to guide their expectations about the non-obvious object properties. To answer this question, the frequency of target actions and latency to the first action in the Labeled condition were compared to the Unlabeled Control condition. Multivariate planned contrasts with one within factor, Condition, revealed a significant difference in infants' exploratory action in the Labeled condition compared to the Unlabeled Control condition, Hotelling's t (45) = 1.82, p<.05 (one-tailed). Univariate analyses indicated that infants demonstrated more target actions in the Labeled condition than in the Unlabeled Control condition, t (46) = 2.22, p<.02 (one-tailed). However, latency to the first target action did not differ between the two conditions. Furthermore, Wilcoxon Matched-Pairs Signed-Ranks tests indicated a marginally significant difference between the Labeled and Unlabeled Control conditions in frequency of target actions, z = -1.52, p<.07 (one-tailed), and no significant difference in latency between the two conditions. The greater frequency of target actions in the Labeled condition (M = 2.03 target actions, SD = 1.31) relative to the Unlabeled Control condition (M = 1.53 target actions, SD = 1.05) suggests that infants used the object labels in the Labeled condition to guide their inferences about non-obvious object properties. Against expectations, however, no corresponding effect of labeling on latency to the first target actions was observed (M = 7.53 seconds, SD = 5.43 for the Labeled condition and M = 8.25 seconds, SD = 5.42 for the Unlabeled Control condition). Perhaps we failed to observe a difference in latency between the two conditions because a) infants' initial attempt to elicit the target property is based on perceptual similarity alone,

whereas b) hearing the two toys labeled leads to persistence in such attempts as reflected in the greater frequency of attempts in the Labeled condition relative to the Unlabeled condition.

Differences between the Labeled and Unlabeled conditions were also examined across toys (as opposed to subjects). Multivariate planned contrasts revealed that when toys were presented in the Labeled as opposed to the Unlabeled Control condition, the frequency of infants' target actions was greater and the latency to the first target action shorter, Hotelling's t (6) = 2.62, p<.02 (one-tailed). Univariate analyses indicated that infants demonstrated more frequent target actions on the toys when toys were shown in the Labeled condition compared to the Unlabeled Control condition, t (7) = 3.44, p<.01 (one-tailed); however, no significant effect for latency was obtained. Thus the effect of labeling generalized across toys suggesting that the effects reported above in which subjects were used as a random factor can be interpreted with confidence as indicating that labels were guiding infants' expectations about object properties.

Transfer

The coding of the rate of transfer of target actions in the Baseline Control condition provided some additional information that infants were indeed making inferences based on physical similarity in the Unlabeled Control condition rather than merely repeating the last action which yielded an interesting result. Instances of transfer of target action were observed on only 14 of the 93 trials (15%) in the Baseline Control condition (in which toys were dissimilar in appearance) compared to on 67 of the 92 trials (73%) in the Unlabeled Control condition (in which toys were moderately similar). For one pair of toys in the Baseline Control condition, the collapsible cup and the measuring tape, the transfer of target action had a clear perceptual basis although in fact the action

would not have successfully elicited the target property and the toys were, in a global sense, quite different in appearance. Excluding this toy pair resulted in instances of transfer being observed on only 6 of 81 trials (7%) in the Baseline Control condition compared to on 58 of 80 trials (73%) in the Unlabeled Control condition. The relatively small number of trials in the Baseline Control condition on which transfer was observed suggests that instances observed in the Unlabeled Control were indeed due to inferences based on perceptual similarity.

Frequency of Looking

Another question of interest was whether the number of looks to the first toy during second toy exploration was influenced by labeling. Although the total number of looks did not differ between the Labeled condition and Unlabeled Control condition or Baseline Control condition, the number of looks judged to be for the purpose of comparing the two toys did differ between conditions. The number of comparative looks was greater in the Labeled condition than in Unlabeled Control and Baseline Control conditions ($\chi^2(2, \underline{N} = 48) = 8.89$, $p<.02$). Although such comparative looks were not all that frequent, infants were almost twice as likely to make a comparative look to the first toy in the Labeled condition than in the Unlabeled Control condition or Baseline Control condition (33 comparative looks in 93 (35%) trials in the Labeled condition versus 16 in 96 (17%) trials in the Unlabeled Control and 16 in 93 (17%) trials in the Baseline Control conditions). Expectations about the common non-obvious properties shared by two similar looking toys given the same label may have led infants in the Labeled condition to check the first toy more often to see how to elicit this property than in the Unlabeled Control and Baseline Control conditions. This result lends support to the finding that hearing two toys given a

common label leads infants to actively pursue an expectation that the two toys share common non-obvious properties.

Summary

These analyses suggest that infants perceived the similarity between the moderately similar toys in the Unlabeled Control condition and hence inferred that the objects possessed similar non-obvious properties. However, this effect must be interpreted with caution as the analysis using toys as a random factor fails to support this conclusion. More importantly, the analyses provide clear evidence that when infants hear two moderately similar toys given the same name, they expect the two toys to share the same non-obvious property.

The present study leaves the source of this labeling effect unresolved. It is not clear whether infants were indeed paying attention to the fact that the two toys were given a common label. Perhaps the increased frequency of attempts observed in the Labeled condition relative to the Unlabeled Control condition was due merely to the presence of any object label. We do not know whether infants were sensitive to the common label, per se. A second study attempted to clarify this issue by comparing the frequency of infants' target actions and the latency to the first action when two toys are given the same label as opposed to different labels.

STUDY 2

In this study, infants explored pairs of moderately similar novel toys in three conditions: Same Label, Different Label and Unlabeled Control. In the Same Label condition, corresponding to the Labeled condition of Study 1, infants heard two moderately similar toys given a common label. In the Unlabeled Control condition, as in the first study, infants explored pairs of moderately similar toys which were not labeled but instead talked about in a general way. This condition was included to allow a replication of the labeling

effect of Study 1. In the Different Label condition, infants explored pairs of moderately similar toys given different labels (e.g., the first toy was referred to as a peri while the second toy was called a bik).

As in Study 1, we expected infants to show a higher frequency of target actions in the Same Label condition relative to the Unlabeled Control condition. The failure to find a latency difference in Study 1 between the Labeled and Unlabeled Control conditions renders predictions about latency unclear in Study 2. Of interest is whether the same pattern will be obtained with the new toy set.

If infants are sensitive to the common label shared by the toys in the Same Label condition, then we would expect them to show more frequent target actions in the Same Label condition relative to the Different Label condition. Again, given that no latency difference was found between the Labeled condition and Unlabeled Control condition in Study 1, it is unclear what to expect from the latency measure when toys are given the same versus different labels.

Further, no clear-cut prediction arose regarding possible differences between the Different Label and Unlabeled Control conditions. It may be that hearing two moderately similar toys given different labels (Different Label condition) would inhibit the frequency and lengthen the latency of infants' attempts to produce the target property relative to those in the Unlabeled Control condition. On the other hand, infants may not regard the different labels as informative in any way about shared versus differing non-obvious properties. If this is the case, then no differences in frequency of target actions and latency to the first attempt between the Different Label and Unlabeled Control conditions would be expected.

Method

Subjects

Fifty-five normally developing infants in each of two age groups (16-18 months and 20-22 months) were recruited through newspaper advertisements and postnatal information sessions at maternity hospitals in the area. Nineteen infants were eliminated, due to fussiness (12), prior familiarity with the stimuli (2), design criteria (two or more trials eliminated due to fussiness and familiarity of stimuli) (4), and experimenter error (1), yielding a final sample of 36 infants: 18 infants in each of the two age groups. All infants were from monolingual English-speaking families, were developing normally and had no history of serious ear infection.

Materials

Stimuli. Two toy sets were used, comprised of three pairs of toys each. As in Study 1, toy pairs were chosen to be novel, manipulable, interesting and visually distinctive from other pairs. Each toy pair consisted of two toys only moderately similar in appearance, although capable of the same non-obvious property. The toy sets were redesigned with two criteria in mind: first, replacement of toys from Study 1 which were found to be familiar to several of the infants thus resulting in eliminated trials, and second, selection of toys which would result in a high level of inference. See Table 3 for a list of the toy pairs which were used, their target properties and the actions required to elicit the properties.

For all three conditions (as in Study 1), the second toy in each pair was disabled such that it failed to produce the target property but showed no surface evidence of this alteration.

Equipment. The same equipment was used as in Study 1.

Table 3

Toys, Their Properties, and Target Actions for Study 2

First Toy	Second Toy[a]	Target Property	Target Action
Set 1:			
lint brush	corkscrew	hidden red 'velvety' surface	remove inner part
janitor's key-ring	measuring tape	hidden chain/ automatic retraction/ 'zippy' sound	pull tab to extend & release
accordian squeeker	telescope	honking sound contracts/expands	squeeze top & bottom together
Set 2:			
wheel-shaped suction toy	barrel-shaped suction toy	adheres/popping sound	press to suction & lift
collapsible cup	periscope	expands/contracts	pull apart sides
folding binoculars	folding brush	hidden parts/ clacking sound	hinge edges

[a] The second toy is disabled such that it fails to produce the target property.

Design

Each infant explored sets of toys in three conditions: Same Label, Unlabeled Control and Different Label.

Same Label condition. Infants explored the first novel toy in a pair which was capable of producing a non-obvious property. Then infants explored a second novel toy which was moderately similar to the first toy. The second toy was unable to produce the non-obvious property. Both toys were given the same novel label.

Unlabeled Control condition. As in the Same Label condition, infants explored the first novel toy in a pair which was capable of producing a non-obvious property. Then infants explored a second novel toy which was moderately similar to the first toy and unable to produce the non-obvious property. Toys were talked about in a general way but were not labeled.

Different Label condition. Infants explored the first novel toy in a pair which was capable of producing a non-obvious property. Then infants explored a second novel toy which was moderately similar to the first toy and unable to produce the non-obvious property. The first and second toys were given different novel labels.

All possible orders of the three conditions were used. Order of conditions and toy pairs were counterbalanced such that each toy pair appeared equally often in each condition (Same Label, Unlabeled Control, Different Label). Two blocks of trials were run. Each block consisted of one trial of each condition. Toy sets were randomly assigned to the blocks such that for half of the subjects toy set 1 was assigned to the first block and for the other half of the subjects toy set 2 was assigned to the first block. Three novel labels were assigned to toy set 1 (peri, totóma, bik) and three to toy set 2 (mido, nánaku, fup). Labels within a set were chosen to maximize auditory distinctiveness so that in the Different

Label condition especially, the labels would stand out as being clearly different. Each label appeared equally often assigned to the first as opposed to the second toy in a pair and equally often in each condition (Same Label versus Different Label).

Infants had the opportunity before the experimental trials to play with a pair of familiar toys (e.g., dog and ball). No data were be collected from infants' exploration of these toys. These pairs were included only as "warm-up" toys to help infants become comfortable with the experimental situation. These toys were also presented again between blocks of trials to give infants the opportunity to engage in a less demanding task (i.e, a familiar game with familiar objects and familiar labels).

Procedure

The procedure was essentially the same as in Study 1 with three exceptions: first, the experimenter played a hide-and-seek game with the infants using a dog, a ball and a basket before each block of trials; second, the experimenter continued to look at the infant during toy exploration, rather than look away, to minimize fussiness; and third, labeling phrases were shortened in order to reduce the amount of time the experimenter spent talking. Please see Table 4 for the list of carrier expressions used.

Coding

Target Action Coding. The coding procedure was the same as in Study 1. The reliability of coders' judgments was determined by having another coder who was also blind to experimental condition recode all conditions for eight randomly-selected subjects, resulting in 48 20s trials altogether. Percent agreement was calculated by dividing the number of instances on which coders agreed by the total number of instances observed. Coders demonstrated 90% agreement on the frequency of target actions.

Table 4

Carrier Phrases Used in Study 2

Same Label & Different Label	Unlabeled Control
First toy:	
I'm going to show you a peri.	I'm going to show you a new one.
Yes, a peri.	Yes, uh huh.
Want to see the peri?	Want to see it?
Look! a peri!	Look! uh huh!
You want the peri?	You want it?
Here's the peri.	Here it is.
Second toy:	
Guess what I'm going to show you next?	Guess what I'm going to show you next?
A bik![a]	Uh huh!
See the bik?	See this one?
Look a bik.	Look at this.
Yes, a bik.	Yes, uh huh.
You want the bik?	You want it?
Here's the bik!	Here it is!

[a] In the Same Label condition, this label would be the same as the one used for the first toy.

Looks Coding. The coding procedure was the same as in Study 1. The same coder coded looking behaviour in both Studies 1 and 2. Because the coding scheme for both studies was identical and the coder's reliability was sufficiently high, an additional measure of reliability for Study 2 looks coding was not considered necessary.

Results and Discussion

As in Study 1, two measures of inferential ability were obtained from infants' exploration of the second (disabled) toys: the number of times infants produced the target action required to elicit the non-obvious property and latency to the first target action. If infants are able to use object labels to guide their inferences about the non-obvious object properties, we would expect infants to show more target actions in the Same Label condition relative to the Unlabeled Control condition. However, given the results of Study 1, it is unclear whether latency to infants' first target action will differ between the Labeled condition and the Unlabeled Control condition.

Furthermore, if infants understand that two objects must have the same label before that label can be used to guide inferences about object properties, then we would expect infants to show more target actions on the second toy in the Same Label condition relative to the Different Label condition. As mentioned earlier, given that in Study 1, labeling toys did not decrease the latency to the first target action compared to when toys were unlabeled, it is unclear whether a difference in latency in the Same Label compared to the Different Label condition will be observed.

As in Study 1, only trials in which infants demonstrated the target action on the first toy were included in the analyses. Unless infants displayed at least one action on the first toy, we could not be sure that they had noticed the toy's non-obvious property. If infants had not noticed the property, then they would

be unable to draw inferences about a second toy's non-obvious property. Hence, only those trials on which infants attempted the target action with the first toy of the pair were analyzed. This happened rarely: only two out of a total of 216 trials were eliminated because of failure to attempt the target action on the first toy. And an additional thirteen trials were eliminated because of fussiness or prior familiarity of toys. When trials from both blocks were available, the frequency and latency scores were averaged across the trials in separate blocks. When only one trial was available, one score was used in place of an average score for that infant.

Because Study 1 revealed differences between conditions on the frequency measure but not on the latency measure, univariate analyses were carried out rather than multivariate analyses. The frequency data were submitted to a mixed design ANOVA with one between-group factor, Age (16-18 months & 20-22 months) and one within-subject factor, Condition (Same Label, Different Label, Unlabeled Control). Univariate analyses of frequency revealed a significant effect of Age on frequency of target actions, $F(1,34) = 9.84$, $p<.004$. Twenty- to 22-month-old infants produced on average significantly more target actions during the 20s play period than did 16- to 18-month-old infants (M = 1.87 target actions, SD = 1.35, and M = 1.13 target actions, SD = 0.98, respectively). No significant main effect for Condition was noted. Furthermore, no significant interactions were found between Condition and Age on frequency of target actions. A natural log transformation of the raw data was carried out to correct for skewed data and the transformed data submitted to a mixed design ANOVA. As the pattern of results remained the same using the transformed data, only results from the analyses using raw data are reported here.

When latency data were submitted to a mixed design ANOVA with one between-group factor, Age (16-18 months & 20-22 months) and one within-

subject factor, Condition (Same Label, Different Label, Unlabeled Control), no significant main effects or interactions were observed. Furthermore the pattern of results remained the same when natural log transformed data were used.

As the significant overall Age difference indicated that the frequency of 16- to 18-month-olds' target actions was depressed relative to that of the 20- to 22-months-olds, condition differences may have been obscured. For this reason, separate analyses were conducted for each age group in addition to analyses collapsed across age group.

A repeated measures ANOVA with one within-factor, Condition (Same Label, Different Label & Unlabeled Control) was carried out for the 16- to 18-month-old group revealing no significant main effect for Condition on frequency of target actions. Furthermore, the same analysis examining the latency data revealed no significant main effect for Condition.

Likewise, repeated measures ANOVAs with one within-factor, Condition (Same Label, Different Label & Unlabeled Control) carried out for the 20- to 22-month-old group revealed no significant main effect for Condition for either frequency of target actions or latency to the first action. The average frequency of target actions and the average latency to first attempt are displayed with respect to age and condition in Figures 3 and 4.

Because a priori predictions were made about paired comparisons, comparisons between conditions were made even though there was no main effect of Condition on either the frequency or latency measures.

Effect of Labeling

To determine whether the effect of labeling on infants' inferential behaviour found in Study 1 was replicated in Study 2, a univariate planned contrast with one within-factor, Condition, was carried out. A significant effect of labeling was noted; infants demonstrated more target actions in the Same

Figure 3. The frequency of target actions on the second toy as a function of condition (Study 2).

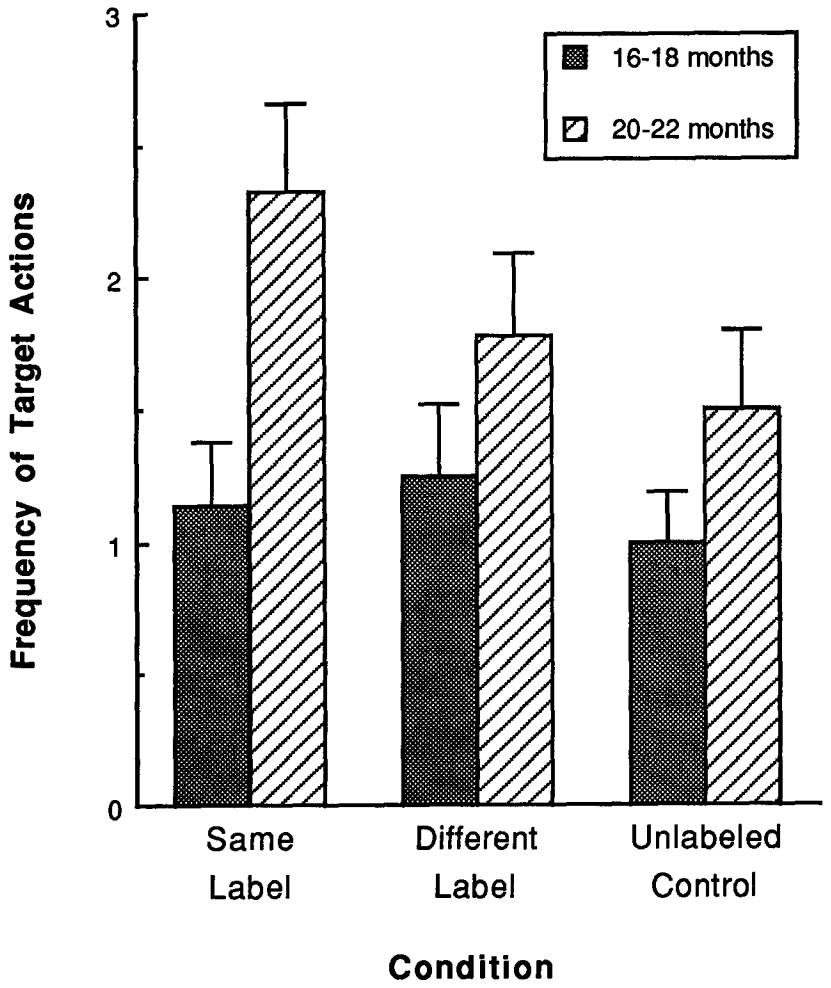
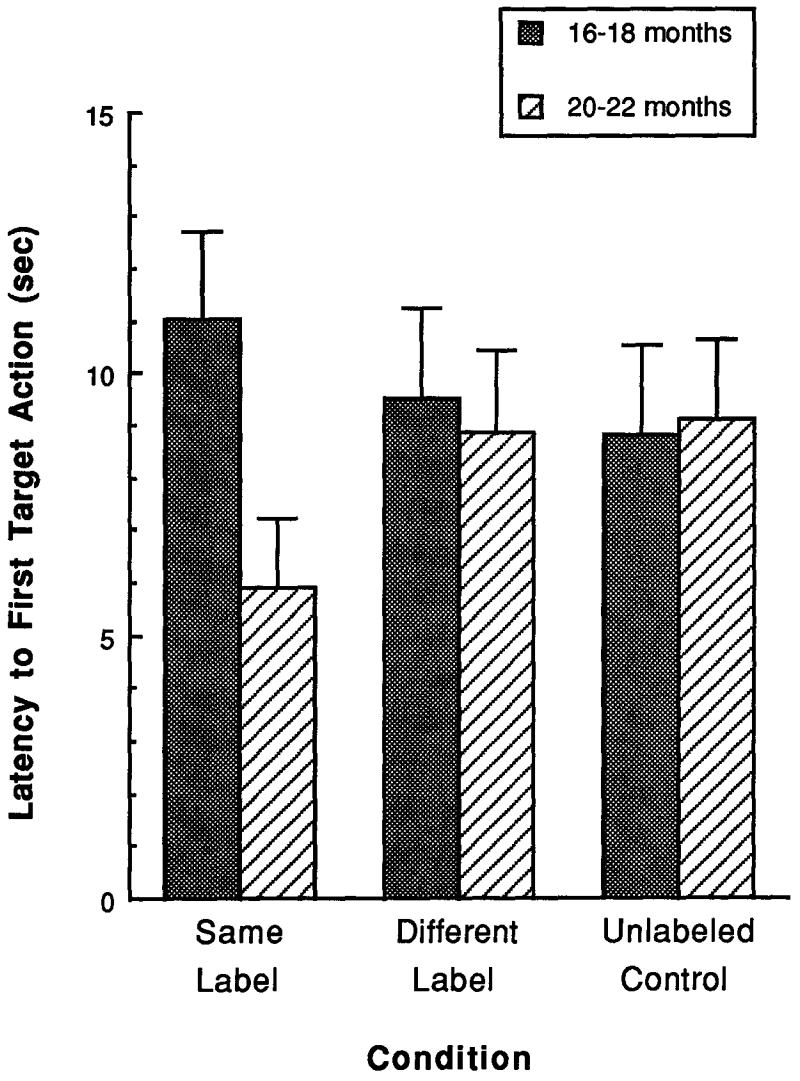


Figure 4. Latency to infants' first target action on the second toy as a function of condition (Study 2)



Label condition than in the Unlabeled Control condition, $t(34) = 1.66$, $p=.05$ (one-tailed) (M = 1.74 target actions, SD = 1.36 in the Same Label condition, M = 1.25 target actions, SD = 1.08 in the Unlabeled Control condition). As in Study 1, Wilcoxon Matched-Pairs Signed-Ranks tests were carried out as well, revealing a significant difference between the Same Label and Unlabeled Control conditions in frequency of target actions, $z = -1.84$, $p<.04$ (one-tailed). These results replicate the labeling effect found in Study 1: in the Same Label condition, labeling leads infants to persist in attempts to replicate the target property as reflected in the greater frequency of target actions in this condition relative to in the Unlabeled Control condition.

Using the latency data, a univariate planned contrast with one within-factor, Condition, was carried out. Latency to the first target action did not differ between the two conditions (M = 8.46 seconds, SD = 6.80 in the Same Label condition, M = 8.95 seconds, SD = 6.94 in the Unlabeled Control condition). A Wilcoxon Matched-Pairs Signed-Ranks test also failed to show a significant difference in latency between the two conditions. Thus, as in Study 1, no difference in the latency to the first target action was observed between the two conditions.

Analyses Across Toys. As in Study 1, Condition differences were also examined across toys as opposed to across subjects (i.e., toy was treated as a random factor). No reliable differences in either frequency of target actions or latency to the first target action were observed between the Same Label condition and Unlabeled Control condition. A Wilcoxon Matched-Pairs Signed-Ranks test using frequency data approached significance: the frequency of target actions was somewhat greater in the Same Label condition relative to the Unlabeled Control conditions, $z = -1.36$, $p<.09$. The failure to find significance

using toys as a random factor suggests that the obtained effects using subjects must be interpreted with caution.

16- to 18-Month-Olds. To examine whether the labeling effect on frequency of target actions held for each age group, univariate planned contrasts were carried out separately for each age group. For the 16- to 18-month-old age group, the univariate contrast did not yield a significant difference between the Same Label and Unlabeled Control conditions in frequency of target actions (M = 1.14 target actions, SD = 1.03 in the Same Label condition, M = 1.00 target actions, SD = 0.80 in the Unlabeled Control condition). Furthermore, a univariate planned contrast showed no significant difference in latency between the two conditions (M = 11.04 seconds, SD = 7.05 in the Same Label condition, M = 8.80 seconds, SD = 7.42 in the Unlabeled Control condition). Wilcoxon Matched-Pairs Signed Ranks tests were also not significant for either the frequency or latency measures.

20- to 22-Month-Olds. When parallel analyses to those performed for the younger age group were carried out for the older age group, an effect of labeling on frequency of target actions was found as well as signs of an effect on latency not revealed by the analyses collapsed across age group. The univariate contrast indicated that infants showed a marginally higher frequency of target actions in the Same Label condition than in the Unlabeled Control condition, $t(17) = 1.66$, $p<.06$ (one-tailed) (M = 2.33 target actions, SD = 1.40 in the Same Label condition, M = 1.50 target actions, SD = 1.27 in the Unlabeled Control condition). Furthermore, a marginally shorter latency to the first target action in the Same Label condition relative to the Unlabeled Control condition was observed, $t(17) = 1.46$, $p=.08$ (one-tailed) (M = 5.88 seconds, SD = 5.59 in the Same Label condition, M = 9.09 seconds, SD = 6.63 in the Unlabeled Control condition). These effects were confirmed by the non-parametric

analyses. Wilcoxon Matched-Pairs Signed Ranks tests indicated that the frequency of target actions was indeed significantly higher in the Same Label condition than in the Unlabeled Control condition, $z = -1.71$, $p<.05$ (one-tailed) and that the latency to the first target action was marginally shorter in the Same Label condition relative to the Unlabeled Control condition, $z = -1.46$, $p=.07$ (one-tailed). Thus, the 20- to 22-month-old infants showed the same pattern with respect to frequency as the contrasts collapsed across age group showed: infants produced more target actions in the Same Label condition than in the Unlabeled Control condition. Furthermore, the 20- to 22-month-old infants showed a somewhat shorter latency to the first target action in the Same Label condition relative to the Unlabeled Control condition--an effect which was not revealed by the contrasts collapsed across age group.

Effect of Same vs. Different Labels

The major question of interest in Study 2 was whether infants demonstrated a sensitivity to the fact that two toys must share a common label before that label can be used to guide inferences about shared non-obvious properties. To answer this question, univariate planned contrasts with one within-factor, Condition, were carried out to determine if the frequency of target actions or the latency to the first target action differed between the Same Label condition and the Different Label condition. The univariate contrast using frequency data indicated no significant difference in frequency between the two conditions although the means are in the predicted direction (M = 1.74 target actions, SD = 1.36 in the Same Label condition, M = 1.51 target actions, SD = 1.23 in the Different Label condition). Moreover, a univariate contrast using latency data also failed to show significant condition differences (M = 8.46 seconds, SD = 6.80 in the Same Label condition, M = 9.17 seconds, SD = 7.00 in the Unlabeled Control condition). Wilcoxon Matched-Pairs Signed-Ranks

tests also failed to reveal significant differences in frequency and latency between the Same Label and Different Label conditions. Furthermore, when condition differences were examined across toys, neither frequency nor latency differed between the Same Label and Different Label conditions. Previously reported analyses have demonstrated that collapsing across age groups obscures effects, therefore, condition differences will again be examined for each age group separately.

16- to 18-Month-Olds. The same pattern of results was found for the younger age group. The univariate contrast indicated no significant difference in frequency between the Same Label and Different Label conditions (M = 1.14 target actions, SD = 1.03 in the Same Label condition, M = 1.25 target actions, SD = 1.13 in the Different Label condition). Furthermore, latency did not differ between the conditions either (M = 11.04 seconds, SD = 7.05 in the Same Label condition, M = 9.50 seconds, SD = 7.33 in the Unlabeled Control condition).

20- to 22-Month-Olds. In contrast to the analysis performed on the data collapsed across age group, differences between the Same Label and Different Label conditions were observed in both frequency of target actions and latency to the first action for the older age group. The univariate contrast using frequency data revealed that infants showed a marginal increase in target actions in the Same Label condition relative to the Different Label condition, $t(17) = 1.56$, $p<.07$ (one-tailed) (M = 2.33 target actions, SD = 1.40 in the Same Label condition, M = 1.78 target actions, SD = 1.31 in the Different Label condition). Furthermore, a univariate contrast revealed a shorter latency to the first action in the Same Label condition than in the Different Label condition, $t(17) = 1.77$, $p<.05$ (one-tailed) (M = 5.88 seconds, SD = 5.59 in the Same Label condition, M = 8.83 seconds, SD = 6.85 in the Different Label condition). These

findings were confirmed by Wilcoxon Matched-Pairs Signed-Ranks tests, $z = -1.62$, $p=.05$ (one-tailed) for frequency, and $z = -1.61$, $p=.05$ (one-tailed) for latency. Twenty- to 22-month-olds understand that two toys should have the same label before that label can be used to guide inferences about non-obvious object properties as indicated by the higher frequency of target actions and shorter latency to the first target action in the Same Label condition relative to the Different Label condition.

Effect of Different Label vs. No Label

An additional contrast of interest involved the Different Label and Unlabeled Control conditions. Recall that we had no a priori expectations about the frequency of target actions or latency to first action in the Different Label condition relative to the Unlabeled Control condition. No significant differences between the two conditions in either frequency or latency were noted by the univariate contrasts or the Wilcoxon Matched-Pairs Signed-Ranks tests. The analyses across toys also failed to yield significant differences for frequency and for latency between the Different Label condition and Unlabeled Control condition. Contrasts carried out for each age group separately also failed to yield significant differences between the Different Label and Unlabeled Control conditions on both frequency and latency measures.

Frequency of Looking

As in Study 1, we examined whether the number of looks to the first toy during second toy exploration was influenced by labeling. The total number of looks did not differ between the Labeled condition and Unlabeled Control condition or Different Label condition. And, in contrast to Study 1, the number of looks judged to be for the purpose of comparing the two toys did not differ between the Same Label and Unlabeled Control conditions. Moreover, no other condition differences in number of comparative looks were revealed.

Failure to observe a difference in the number of comparative looks between the Same Label condition and Unlabeled Control condition in this study may have been due to a methodological difference between Studies 1 and 2. In Study 1, the experimenter averted her gaze during toy exploration whereas in Study 2 she looked at the infant. Perhaps having the experimenter's face to look at may have competed with the first toy as a possible source of information for the infant and thus inhibited the number of comparative looks towards the first toy.

Summary

These analyses partially replicate the labeling effect found in Study 1: in the Same Label condition, labeling led infants to persist in attempts to replicate the target property as reflected in the greater frequency of target actions in this condition relative to in the Unlabeled Control condition. Furthermore, for infants of 20 to 22 months, an increased level of inferences occurred only when toys were given a common label. No such effects of labeling or common labels were observed for the younger age group, perhaps because of the overall lower frequency of target actions in Study 2 relative to Study 1. Hence we are unable to clarify whether infants of 16 to 18 months of age are sensitive to the fact that two toys which share a common label share deeper, non-obvious properties.

GENERAL DISCUSSION

The results of Studies 1 and 2 provide new and interesting information about infants' inferential abilities. In Study 1 a comparison of the frequency of target actions and latency to the first action in the Unlabeled Control condition relative to Baseline Control condition provided evidence that infants were able to perceive the similarity between the only moderately similar toys and, based on that similarity, were able to draw inferences about non-obvious object properties. Furthermore, additional evidence provided by the Baseline Control condition that transfer of target actions occurred infrequently indicates that

infants' attempts to elicit the target property in the Unlabeled Control condition were not simply perseveration of a previously successful action. This finding replicates and extends the findings of Baldwin, et al. (in press) by demonstrating that infants are able draw inferences based on only moderate perceptual similarity, at least by 16-22 months.

Furthermore, Studies 1 and 2 together suggest that infants use object labels in addition to perceptual similarity to guide their inferences about non-obvious object properties. However, in Study 2, when the data for each age group was examined separately, the labeling effect was not obtained for the 16- to 18-month-olds. In contrast, a labeling effect was noted for the 20- to 22-month-olds. Moreover, in Study 2, weak support--as indicated by marginal differences in frequency and in latency-- was obtained for the suggestion that infants, at least by 20 months of age, understand that two toys must share a common label before that label can be used to guide expectations about object properties. Stronger support awaits a further replication with a toy set which elicits a higher level of target actions, thus giving more room for condition differences to occur.

There were some minor methodological differences between Studies 1 and 2 which may underlie the failure to replicate the labeling effect for the 16- to 18-month-old group. In Study 1 before each block of toys, infants played with a pair of novel identical toys given the same label which were both capable of producing the same interesting and non-obvious property. These "interim" pairs were included in order to maintain infants' interest in the game and to prevent infants from forming a response set (e.g., the second toy never works). We chose to replace these pairs in Study 2 with two familiar toys (a dog and a ball) and a basket in a hide-and-seek game. We were concerned that the additional labels required for this second study to achieve the different label condition

might be overwhelming for such young children. For this reason, we chose a less demanding but still interesting activity to maintain infants' cooperation in the game. Perhaps, however, the interim pairs of toys in Study 1 served an important purpose by helping infants to realize that a familiar piece of real-world knowledge was applicable in the current game--the notion that two identical toys with the same name share the same non-obvious property. Furthermore, omitting these pairs may have increased the chances that infants formed a response set--the second toy never works--simply because the proportion of novel toys that did not work was now half whereas in Study 1 it was only slightly more than one-third.

One issue which awaits future resolution is the degree of similarity necessary between two objects before infants will rely on object labels to guide their inferences. In these studies toys were chosen to be only moderately similar in appearance; if toys were very similar in appearance object labels would not provide any additional cues as to the underlying non-obvious property shared by the toys above those given by perceptual similarity. However, using toys only moderately similar in appearance may have partly masked the effects of same and different labels. If the objects had been more similar in appearance, even the younger infants (16- to 18-month-olds) might have demonstrated an understanding that two toys sharing the same label share the same non-obvious properties. Perhaps, additionally, the degree of perceptual similarity necessary before infants will rely on object labels to guide their inferences changes with age. It was not a goal of this study, however, to manipulate similarity in any fine-grained way in order to determine the relationship between degree of similarity and use of object labels as a guide to inference. Further research is necessary to determine whether infants of different ages will use object labels to guide their expectations about non-

obvious object properties when perceptual similarity is even minimal or nonexistent.

It might be argued that young infants would be more likely to use language as a guide to inference when familiar, rather than novel, labels and categories are used. However, novel exemplars would still be required as test toys. In particular, a novel exemplar dissimilar in appearance would be necessary so that infants failed to recognize it as a member of the familiar category based on perceptual features alone. If infants immediately recognized the test toy as belonging to a familiar category, no additional information would be provided by the label as to category membership over and above that already indicated by perceptual features. However, with a novel exemplar dissimilar in appearance, the situation becomes one in which infants must rely on the object label, in the absence of perceptual similarity, as a guide to inference. If infants are unable to use novel labels to guide expectations when two novel toys are moderately similar in appearance, it seems doubtful that they would be willing to extend a familiar label to a novel exemplar in the absence of perceptual support. Moreover, as described earlier, Davidson and Gelman (1990) found that even preschoolers are reluctant to rely on object labels if labels are entirely inconsistent with a perceptually-based nonlinguistic partitioning of the test objects.

The youngest infants to participate in these studies were 16 months of age; since younger infants have not been tested, there is no definite reason to believe that using object labels as a guide to inference is beyond their capability. From previous research it is evident that infants show basic inferential skills as early as 9 months of age (Baldwin, et al., in press). However, the earliest age at which there is reliable evidence of language comprehension is around 10 to 12 months of age (Oviatt, 1980;1985).

Therefore, it is possible that infants are capable of using object labels to guide inferences from the start of language comprehension. Evidence from other domains suggests that a task which does not require the infant to actively manipulate toys might be a more sensitive window on infants' abilities. For example, Baillargeon and colleagues (e.g., Baillargeon, Spelke & Wasserman, 1985; Baillargeon, 1987) demonstrated that infants as young as 3 1/2 months understand that an object continues to exist when hidden--several months before they show such understanding using Piagetian search tasks which require infants to actively search a spatial array. Furthermore, Baillargeon (Baillargeon, DeVos & Graber, 1989) showed that even for 9-month-old infants, a task which requires the infant to perform an action, such as the AB search task, may not be the best indicator of their understanding of object permanence. Baillargeon, DeVos & Graber (1989) attribute 9-month-olds' failure on the AB search task to an inability to integrate memory information into the planning and execution of action. They believe there may be a developmental lag between cognition and action in which understanding precedes the ability to act on understanding. If there is a lag between cognition and action, then using a non-action methodology (e.g., an adaptation of the preferential-looking methodology) to investigate infants' use of object labels as a guide to inference might reveal an early emerging understanding and provide a more sensitive window on developmental change.

These studies leave unanswered whether the observed effects of labeling on inferential abilities are indeed language-specific. Perhaps these effects are merely the result of an association between any sound (in this case an object label) or, for that matter, any other associate, and the non-obvious object properties. If labels are functioning in a simple associative manner, then other non-linguistic stimuli, such as a whistle sound or a coloured star, which

covary with non-obvious properties should also function to guide inference. For example, would exploring a novel toy on which the experimenter stuck a blue star lead infants to expect a similar-looking toy which also received a blue star to share the same non-obvious property? Such controls would help clarify whether the effects observed in this study are the result of a specific property of object labels or merely the result of general associative mechanisms. Given Waxman and Balaban's (1992) evidence which suggests that 9-month-olds use words, but not tones, to categorize objects, we might expect that 16-month-olds would use object labels, rather than any other associate to guide inferences about non-obvious object properties.

Even if infants restrict inferences to words rather than any associate, it is not clear whether they would regard any common words as candidates for guiding inferences or whether they would require common object labels, per se. For example, hearing the adjective "blueish" applied to two objects might lead infants to expect that both objects shared a common non-obvious property, such as a retractable chain. There is evidence that by 2 1/2 years of age children will use only object labels, not temporary state adjectives (e.g., "sleepy"), to guide their inferences about category membership (Gelman & Coley, 1990), but no such evidence yet exists for infants. Based on their research, Markow and Waxman (1992) suggest that 12-month-olds will use both nouns and adjectives for categorization purposes. However, it may well be that by 16 months of age, infants understand that only object labels (nouns) can be used for categorization. Implicit in this issue are two assumptions: (1) that infants can recognize when they have heard an object label as opposed to another kind of word; and (2) that infants understand that the label applies to the whole object rather than to one of its properties or its substance, for example. These assumptions will be taken up in turn.

For the assumption that infants can distinguish labels from other types of words, there is as yet no direct evidence. However, several researchers have recently provided evidence that from 17 months of age, infants have some knowledge of syntactic structure which may allow them to differentiate labels from other words (Gleitman, 1990; Hirsh-Pasek, Golinkoff, Fletcher, DeGaspe Beaubien, & Cauley, 1985; Naigles, 1990). In a preferential looking task that Hirsh-Pasek et al. devised, infants demonstrated syntactic knowledge by looking longer at the video display matching the sentence they heard; for example, if they heard "Big Bird is tickling Cookie Monster," infants looked longer at the screen depicting that event than at the one depicting Cookie Monster tickling Big Bird. Using a similar procedure, Naigles demonstrated that two-year-olds use syntactic knowledge to map a novel action term onto the correct event. Based on these and other similar findings, Gleitman argues that very young children use syntactic knowledge to acquire new action terms. Moreover, Katz, Baker and MacNamara (1974) demonstrated that from 18 months of age, female infants use syntactic cues to distinguish between common nouns and proper nouns. Gelman and Taylor (1984) replicated this finding with two-year-olds of both sexes using a more elaborate design. Perhaps, then, children before the age of two are sensitive to the syntactic cues which differentiate category labels from other terms such as attribute descriptions.

With regard to the second assumption, there is some evidence which suggests infants understand that object labels refer to the whole object rather than to its substance, for example. Some researchers suggest that when faced with a novel object for which they do not know the name, infants' default assumption is to treat a novel label as applying to the whole object rather than to its parts, substance or other properties (e.g., Markman, 1989; Markman &

Wachtel, 1988; Woodward, 1992). In the studies reported here, infants were faced with exactly the situation in which this default assumption is said to be used: a novel label is applied to a novel object. Therefore, it is possible that infants treated the novel labels as object labels rather than attribute or substance terms, for example. Although this issue warrants further research before it can be stated definitively that infants use object labels qua object labels to guide their inferences, the current methodology provides a framework within which such issues can be addressed.

What would it mean to the course of development that the ability to use object labels as a guide to inference is already available in infancy? Firstly, it would mean that infants could rapidly acquire new concepts and thus quickly expand their knowledge about the object world. Without such recourse to object labels, infants would have to explore each and every object individually to learn about its hidden properties unless perceptual similarity led them to recognize it as a familiar item. Thus, using object labels to guide inferences speeds conceptual development. Secondly, the ability to use language to guide inference would likely speed language development, as well as conceptual development, for the infant would be motivated to acquire new labels as a means to gaining information about the world. If object labels were, for the infant, merely a means of requesting a desired object, vocabulary acquisition might proceed more slowly since there are many objects an infant might be uninterested in possessing. An interesting possibility is that the ability to use object labels as a guide to inference may be related developmentally to the rapid increase in vocabulary--for example, the "naming explosion"--that begins in the later half of the second year of life (Benedict, 1979; Goldin-Meadow, Seligman & Gelman, 1976; Gopnik & Meltzoff, 1986, 1987, 1992; Nelson, 1977, 1985, 1991; Oviatt, 1980,1985). All in all, the early emergence of

contact between language and conceptual knowledge provides infants with excellent tools for quickly increasing their linguistic repertoire while simultaneously gaining knowledge and understanding of the world.

NOTES

[1]Parents were asked to indicate whether their infant had played before with any of the toys or with any toys similar in appearance and non-obvious property. Infants for whom toys from two or more trials were familiar were eliminated since novelty of toys was important in order to ensure that infants were drawing inferences from the first toy to the second toy and not drawing on previous experiences with similar toys.

[2]A measure of the coder's reliability was not made because it would not be possible to derive a stable reliability estimate. There were two reasons for this: a) the infrequency of actual instances of transfer, and b) the liberal criterion for actions considered as transfer.

[3]Although the multivariate analysis did not yield a significant main effect of Age, the univariate analysis of frequency yielded a marginally significant effect for Age, $\underline{F}(1, 46) = 3.03$, $p<.09$. However, no significant interactions with Condition were observed.

REFERENCES

Baldwin, D.A., Markman, M.E. & Melartin, R. (in press). Infants' Ability to Draw Inferences about Nonobvious Object Properties: Evidence from Exploratory Play. Child Development.

Benedict, H. (1979). Early lexical development: Comprehension and production. Journal of Child Language, 6, 183-200.

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

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

Gelman, S.A. & Markman, E.M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. Child Development, 55, 1535-1540.

Gelman, S.A. & Taylor, M. (1984). How two-year-old children interpret proper and common names for unfamiliar objects. Child Development, 55, 1535-1540.

Gleitman, L. (1990). The structural sources of verb meanings. Language Acquisition, 1, 3-55.

Goldin-Meadow, S., Seligman, M.E.P. & Gelman, R. (1976). Language in the two year old. Cognition, 4, 189-202.

Gopnik, A. & Meltzoff, A.N. (1986). Words, plans, things, and locations: Interactions between semantic and cognitive development in the one-word stage. In S.A. Kuczaj, III & M.D. Barrett (Eds.), The Development of Word Meaning (pp. 199-223). NY: Springer-Verlag.

Gopnik, A. & Meltzoff, A.N. (1987). The development of categorization in the second year and its relation to other cognitive and linguistic developments. Child Development, 58, 1523-1531.

Gopnik, A. & Meltzoff, A.N. (1992). Categorization and naming: Basic-level sorting in eighteen-month-olds and its relation to language. Child Development, 63, 1091-1103.

Hirsh-Pasek, K., Golinkoff, R., Fletcher, A., DeGaspe Beaubien, F. & Cauley, K. (1985, October). In the beginning: One word speakers comprehend word order. Paper presented at Boston Language Conference, Boston.

Huttenlocher, J. (1974). The origins of language comprehension. In R.L. Solso (Ed.), Theories in Cognitive Psychology. New York: Erlbaum.

Katz, N., Baker, E. & Macnamara, J. (1974). What's in a name? The case of object names. Cognitive Psychology, 19, 63-89.

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

Markman, E.M. (1989). Categorization and Naming in Children: Problems of Induction. Cambridge, MA: MIT Press.

Markman, E.M. & Watchel, G.F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. Cognitive Psychology, 20, 121-157.

Markow, D.B. & Waxman, S.R. (1992, May). The Influence of Labels on 12-Month-Olds' Object Category Formation. Poster presented at the International Conference on Infant Studies, Miami.

Massey, C.M. & Gelman, R. (1988). Preschoolers' ability to decide whether a photographed unfamiliar object can move itself. Developmental Psychology, 24, 307-317.

Naigles, L. (1990). Children use syntax to learn verb meanings. Journal of Child Language, 17, 357-374.

Nelson, K. (1985). Making Sense: The Acquisition of Shared Meaning. Orlando, Florida: Academic Press, Inc.

Nelson, K. (1977). The conceptual basis for naming. In J. Macnamara (Ed.), Language Learning and Thought (pp. 117-136). New York: Academic Press.

Nelson, K. (1985). Making Sense: The Acquisition of Shared Meaning. Orlando, Florida: Academic Press, Inc.

Nelson, K. (1988). Constraints on word learning? Cognitive Development, 3, 221-246.

Nelson, K. (1991). Concepts and Meaning in Language Development. In N.A. Krasnegor, D.M. Rumbaugh, R.L. Schiefelbusch & M. Studdert-Kennedy (Eds.), Biological and Behavioral Determinants of Language Development (1991). Hillsdale, NJ: Lawrence Erlbaum.

Oviatt, S.L. (1980). The emerging ability to comprehend language: An experimental approach. Child Development, 51, 97-106.

Oviatt, S.L. (1985). Tracing developmental change in language comprehension before twelve months of age. Papers and Reports on Child Language Development. No. 24. Stanford, CA: Stanford University.

Waxman, S.R. & Balaban, M.T. (1992, May). The Influence of Words vs. Tones on 9-Month-Old Infants' Object Categorization. Poster presented at the International Conference on Infant Studies, Miami.

Wellman, H.M. & Gelman, S.A. (1988). Children's understanding of the nonobvious. In R.J. Sternberg (Ed.), Advances in the Psychology of Human Intelligence, Vol. 4 (pp. 99-135). Hillsdale, NJ: Erlbaum.

Woodward, A.L. (1992). The Role of the Whole Object Assumption in Early Word Learning. Unpublished doctoral dissertation, Standford University.