The Influence of Shape Similarity and Shared Labels on Infants' Inductive Inferences about Nonobvious Object Properties

Andrea N. Welder and Susan A. Graham

This study examined the influence of object labels and shape similarity on 16- to 21-month-old infants' inductive inferences. In three experiments, a total of 144 infants were presented with novel target objects with or without a nonobvious property, followed by test objects that varied in shape similarity to the target. When objects were not labeled, infants generalized the nonobvious property to test objects that were highly similar in shape (Experiment 1). When objects were labeled with novel nouns, infants relied both on shape similarity and shared labels to generalize properties (Experiment 2). Finally, when objects were labeled with familiar nouns, infants generalized the properties to those objects that shared the same label, regardless of shape similarity (Experiment 3). The results of these experiments delineate the role of perceptual similarity and conceptual information in guiding infants' inductive inferences.

INTRODUCTION

Young children's ability to reason inductively is a fundamental component of early conceptual development. Put simply, inductive reasoning involves invoking the premise that things that are true for one exemplar of a category will hold true for other members of the same category (Moore & Parker, 1989). Thus, inductive reasoning allows for the generalization of knowledge to new instances and new situations, resulting in increased cognitive efficiency and the opportunity to benefit from past experiences. In recent years, a great deal of empirical attention has been devoted to examining young children's inductive abilities, with particular focus on the nature of the categories that guide their inferences (e.g., Baldwin, Markman, & Melartin, 1993; Gelman, 1988; Gelman & Coley, 1990; Gelman & Markman, 1986, 1987; Mandler & McDonough, 1996, 1998). In this article, we report three experiments that investigated the influence of shape similarity and shared object labels on infants' inductive inferences about nonobvious object properties.

Early investigations of young children's inductive abilities suggested that preschoolers were limited to reasoning on the basis of perceptually based categories and were unable to appreciate the inductive power of complex categories (for reviews, see Mandler, 1998; Wellman & Gelman, 1988). More recent research using inductive generalization tasks, however, has demonstrated that by 2½ years of age, preschoolers possess remarkably sophisticated inductive abilities. In a typical inductive generalization task, preschoolers are taught a fact about a target object and then are asked whether that fact can be generalized to other test objects. Using this methodology, studies have demonstrated that preschoolers can classify objects in multiple ways and use these classifications to generalize different types of properties, if they have sufficient knowledge about and experience with unobservable properties (e.g., Kalish & Gelman, 1992). Furthermore, preschoolers are able to limit the inferences they make based on factors such as previous knowledge about a particular category, property generalizability, and category homogeneity (Gelman, 1988). Finally, studies have demonstrated that preschoolers will rely on perceptual similarity to guide their inferences only when they do not have any information about the conceptual relation (or lack thereof) between the target object and the test objects. When preschoolers have conceptual knowledge about the underlying kind of target and test objects (e.g., when they are told that both objects are birds), they will overlook perceptual similarity and use this knowledge about kind to guide their inductive inferences (e.g., Gelman & Markman, 1986, 1987; Kalish & Gelman, 1992). For example, Gelman and Coley (1990) found that children as young as 2½ years of age will disregard perceptual similarity and generalize properties on the basis of shared underlying kind when the target and test objects are given the same count noun label.

Although preschoolers' impressive inductive abilities have received a great deal of empirical attention, only recently have researchers begun to examine the development of inductive capabilities during the infancy period. Yet it is clear that the study of inductive inferences in infancy is important for at least two rea-
Research on inductive reasoning during infancy has focused both on whether infants possess basic inductive reasoning abilities and on how infants will limit their inductive generalizations. In one of the first studies to examine this issue, Baldwin et al. (1993) examined whether infants would generalize nonobvious object properties to objects that were similar in appearance to a target object. In their first experiment, 9- to 16-month-old infants were allowed to explore a novel target object that possessed an interesting property (e.g., a can that made a “wailing” sound when tipped), using three within-subjects conditions. In the violated condition, infants were then presented with a similar-looking, but disabled, test object to explore (e.g., a highly similar can that could not “wail”). In the interest control condition, which served as a baseline condition, both the target and test objects were disabled. In the fulfilled expectation condition, both the target and the test objects possessed the nonobvious property (no data were collected from this condition). Baldwin et al. found that infants performed significantly more target actions on test objects in the violated condition than in the interest control condition, indicating that they expected the test objects to possess the same nonobvious properties as the target objects. In a second experiment, infants performed significantly more target actions on perceptually similar test objects (e.g., cans that differed only in color and pattern from the test object) than on perceptually different test objects, in the violated condition. The results of these two studies indicate that infants as young as 9 months of age can develop expectations about nonobvious object properties after a short exposure to a novel object. Furthermore, infants form specific expectations about nonobvious properties and will only generalize these properties to objects that are highly perceptually similar to the target object.

More recent research indicates that infants can constrain their inductive inferences based on their pre-existing knowledge about conceptual categories. In a series of studies, Mandler and McDonough (1996) used a generalized imitation paradigm to examine infants’ inductive inferences about broad classes of animals and vehicles. They presented 14-month-old infants with small objects and modeled actions that were either appropriate (e.g., a dog was shown drinking from a cup: Experiments 1 and 2) or inappropriate (e.g., a car was shown drinking from a cup: Experiment 3) for a given animal or vehicle. Infants were then presented with another exemplar from the same domain (e.g., a cat) and a distracter from the other
domain (e.g., an airplane) to assess their willingness to generalize these properties. Mandler and McDonough found that infants generalized the appropriate, but not the inappropriate actions, indicating that infants do not merely imitate any action they have seen modeled. Furthermore, infants limited their generalizations to exemplars in the same domain as the target (e.g., infants would imitate a cat drinking but not an airplane). Finally, perceptual similarity did not appear to influence infants' generalization of properties. Infants were as likely to generalize an “animal” property from a dog to a bird or fish as from a dog to a cat.

In a second series of studies, Mandler and McDonough (1998) examined whether 14-month-old infants would distinguish between actions that are appropriate to generalize within one domain only (domain-specific actions), and actions that are potentially appropriate to generalize to both domains (domain-neutral actions). In Experiment 1, Mandler and McDonough found that infants limited their generalization of domain-specific actions to domain-appropriate objects (e.g., they would put a horse, but not a car, to sleep), but generalized imitations of neutral actions to objects from both domains (e.g., they would wash both a horse and a car). In subsequent experiments, Mandler and McDonough found that infants limited their generalizations to categories at the global level (e.g., animals versus vehicles), rather than the basic level (e.g., cats versus dogs), although by 20 months of age, infants did make finer level basic distinctions.

Thus, research indicates that infants will draw inferences about nonobvious object properties based on knowledge gained during the experimental session (Baldwin et al., 1993). Furthermore, infants will limit their inductive generalizations when appropriate, indicating that their imitative actions are purposeful (Baldwin et al., 1993; Mandler & McDonough, 1996, 1998). Finally, research suggests that both perceptual similarity and conceptual knowledge may play a role in guiding infants’ inferences, under differing task demands. That is, Baldwin et al.’s (1993) studies indicate that when reasoning about novel objects, infants will limit their generalization of nonobvious object properties to other novel objects that are highly perceptually similar to the target object. It is important to note, however, that Baldwin et al. did not provide infants with any information about the objects, other than perceptual similarity. Mandler and McDonough’s (1996, 1998) findings suggest that when reasoning about properties that fall within the familiar domains of animals and vehicles, infants’ inductive inferences are guided by their conceptual knowledge about object kind and not by perceptual similarity per se. It has been suggested, however, that in tasks such as those used by Mandler and McDonough, infants may not necessarily be responding on the basis of conceptual knowledge (Oakes & Madole, 1999). Instead, infants may have formed associations between properties (e.g., drinking) and a type of object (e.g., things with eyes) and may have used this accumulated knowledge to guide their inferences (for a discussion countering this argument, see Mandler, 2000a). Although these studies have provided important insights into early inductive reasoning, they leave unanswered the question of the contribution of perceptual and conceptual information in guiding infants’ reasoning about novel object categories and novel object properties, an issue germane to the debate surrounding the nature of infant categorization.

In the present studies, we pursued the examination of inductive reasoning abilities during infancy with particular focus on two issues. First, we investigated the relative influence of perceptual similarity (in the form of shared object shape) and conceptual information (in the form of shared object labels) on infants’ inductive inferences when reasoning about novel categories and novel properties. Second, and in a related vein, we examined whether infants recognized the conceptual information conveyed by shared object labels. There is ample evidence that novel words can promote infants’ formation of object categories, even when the perceptual similarity among objects is low (e.g., Balaban & Waxman, 1997; Waxman, 1999a, 1999b; Waxman & Hall, 1993; Waxman & Markow, 1995). In our studies, we investigated whether infants would treat labels for novel objects as an index of object kind and thus appreciate that these labels could promote inferences about nonobvious properties. The use of novel objects in our studies allowed us to be confident that infants did not have any pre-existing knowledge to guide their inferences (for a discussion countering this argument, see Mandler, 2000a).

The first specific goal of our studies was to examine the role of shape similarity in guiding infants’ generalization of nonobvious object properties when they were presented with novel object categories. As noted earlier, Baldwin et al.’s (1993) results indicated that infants will not generalize nonobvious properties to test objects that differ radically in appearance from the target object (i.e., shared no properties at all with the target). In their studies, however, perceptual similarity was not varied systematically along a continuum. In the present studies, we examined the degree of shape similarity necessary for infants to perceive
objects as members of the same category, and thus, to generalize a nonobvious object property. We chose to focus on shape similarity because research indicates that infants and preschool-age children attend to object shape versus other perceptual properties (such as color and size) when categorizing objects (e.g., Baldwin, 1989; Graham & Poulin-Dubois, 1999; Landau, Smith, & Jones, 1988, 1992; Smith, Jones, & Landau, 1992).

A second goal of the present research was to examine whether infants would rely on shared object labels to direct their inductive inferences. In particular, we examined whether infants would extend a nonobvious property on the basis of a shared object label, even if the objects differed in shape. As noted earlier, when conceptual information about category membership is available, preschoolers will base their inferences on this type of information, disregarding perceptual similarity (e.g., Gelman, 1988; Gelman & Markman, 1986, 1987; Kalish & Gelman, 1992). Similarly, Mandler and McDonough’s (1996, 1998) studies indicate that infants will use their pre-existing conceptual knowledge about animal and vehicle kinds to guide their inductive inferences. However, no published studies have examined whether infants can make use of conceptual information provided on-line to guide their inductive inferences about novel object categories. That is, it is unclear whether infants can rapidly and efficiently abstract information about the object kind of a novel object and use that information to guide their inductive inferences. In the present studies, we examined whether infants would treat labels for novel objects, which varied in shape similarity, as a conceptual marker of object kind and expect those objects that shared the same label to possess the same nonobvious property.

To summarize, the present studies were designed to examine the relative influence of shape similarity and shared object labels on infants’ inductive inferences about nonobvious object properties. In three experiments, a generalized imitation paradigm was employed to examine these issues in infants ranging in age from 16 to 21 months. Infants were presented with novel target objects that possessed nonobvious properties (e.g., a cloth-covered object that squeaked when squeezed). The experimenter demonstrated the nonobvious property using a specific target action. Infants were then presented with a series of test objects that varied systematically in their degree of shape similarity to the target. We reasoned that if infants considered test objects to be members of the same category as the target, they would expect the test objects to share the same nonobvious property as the target. That is, infants’ imitation of a target action on test objects would provide evidence of inductive reasoning. For one group of infants in each experiment, the experimenter labeled the target and test objects using either a novel label (Experiments 1 and 2: “Look at this blint.”) or a familiar label (Experiment 3: “Look at this spoon.”). For the second group, the experimenter simply directed the infants’ attention to the objects (Experiments 1 and 2: “Look at this one.”).

EXPERIMENT 1

In Experiment 1, infants were presented with object sets consisting of a target object followed by a high-similarity match, a medium-similarity match, a low-similarity match, and a dissimilar object in three within-subjects conditions and in either a novel-label or no-label group. The high-, medium-, and low-similarity matches within each set varied in shape and color but shared similar textures. We chose to vary color within object sets as well as shape because previous research has indicated that infants do not view color as predictive of category membership (e.g., Graham & Poulin-Dubois, 1999). The dissimilar objects in each set, however, differed from the target object in texture, shape, and color. The dissimilar objects were included to ensure that infants’ inductive generalizations were specific to objects that they perceived as belonging to the same category (i.e., the high-, medium-, and low-similarity objects), and to assess the possibility that participants were merely imitating the experimenter’s actions on any object, regardless of whether an expectation was generated.

Infants were presented with target and test objects in three within-subjects expectation conditions (similar to those used by Baldwin et al., 1993). In the surprised condition, the target object possessed an interesting sound property (e.g., squeaked when squeezed), but the test objects were disabled so that they could not exhibit the property (e.g., could not squeak when squeezed). This condition was of particular interest, because infants’ performance in this condition would indicate whether they expected test objects to possess the same nonobvious property as the target. In the interest control condition, neither the target nor the test objects possessed the interesting property (e.g., neither produced a squeak sound). This condition provided a baseline measure of infants’ exploratory actions. A comparison of infants’ performance in the interest control condition to the surprised condition would indicate whether the target property of the target object was, in fact, nonobvious on visual inspection. That is, if infants attempted the target actions on the test objects in the surprised condition (in which they knew about the property of the target) but not the interest control condition (in which they were
not shown the property), then we could be confident that the test objects, in themselves, did not suggest the property through their appearances. In the predicted condition, both the target and test object possessed the interesting property (e.g., both could squeak). This condition was included to avoid infants developing the expectation that all test objects were disabled (because all test objects in both the surprised and interest control condition were), and becoming quickly bored or frustrated with the stimuli.

We expected that infants would use degree of shape similarity and the presence or absence of a label to guide their generalizations of the nonobvious property from the target object to a particular test object. First, we hypothesized that the greater the degree of shape similarity between a test object and a target object, the higher the frequency of target actions performed on that test object. Second, we expected that infants would treat the novel labels as identity information, indicating that the test objects belonged to the same category as the target object. Thus, we hypothesized that infants in the novel-label group would perform more target actions on test objects than infants in the no-label group. Third, we predicted that infants’ reliance on shape for inductive inferences would vary as a function of the presence of object labels. That is, we expected that infants in the novel-label group would perform a higher frequency of target actions on medium- and low-similarity matches than infants in the no-label group, because they would have the conceptual information necessary to perceive the medium-similarity and low-similarity matches as belonging to the same conceptual category as the target.

Method

Participants

Data from 57 infants (31 girls, 26 boys; M = 18.77 months, SD = 1.47; range = 15.79–21.46 months) were included in the final sample. Twenty-two additional infants were tested but were excluded from the final sample for the following reasons: excessive fussiness (n = 7), statistical outliers (n = 7; see Coding and Data Reduction section), and lack of performance of target actions on any of the test objects in the surprised condition (n = 8; see Coding and Data Reduction section). Participants were randomly assigned to one of two groups: novel label and no label. The novel-label group consisted of 16 girls and 14 boys (M = 18.76 months, SD = 1.56, range = 15.79–21.46 months), and the no-label group consisted of 15 girls and 12 boys (M = 18.77 months, SD = 1.39, range = 16.26–21.10 months).

Materials

Stimuli. Three objects were used for the warm-up trials: a garlic press, a roller ball, and a clicking clock. Three object sets (a “squeaking” set, a “ringing” set, and a “rattling” set), each novel and visually distinctive from one another, were constructed for use in the imitation task (see Figures 1–3). Each set consisted of five objects: a target object, a high-similarity match, a medium-similarity match, a low-similarity match, and a dissimilar object. The dissimilar objects included a plastic orange file (squeaking set), a small white strainer (ringing set), and a plastic green hosesplitter (rattling set). There were two versions of each target and match object in each set: a functional version that could produce the target sound and a nonfunctional version that was disabled and thus unable to produce the target sound. All stimuli within a given object set (except for the dissimilar object) were crafted from the same type of fabric and varied in only color. For the squeaking object set, hollow rubber balls (7 cm in diameter) were covered with pleated silky rayon and shaped in various ways with string and/or sponge; the functional objects in this set could produce a squeaking noise when squeezed. For the ringing set, metal bells (the kind on store counters that one taps for service) were placed inside a styrofoam shape and were covered with faux-fur material. The bells had a metal button on the top that was connected to a ringer inside; thus, these functional objects could produce a ringing noise when tapped. In the rattling set, 7 cm (rattle portion) × 4 cm (handle portion) rattles were covered with felt and shaped in various ways with sponge, or by removing portions of the rattle frame. The functional objects in this set could produce a rattling noise when shaken. Parents were asked whether their infant had ever seen objects identical to or similar to any of the objects used in this experiment to verify that the objects were indeed novel. All parents verified that the stimuli were novel.

Equipment. A Sony video camera, hidden behind a one-way mirror, was used to provide a visual and audio record of infants’ exploration of the stimuli. The experimenter used a small stopwatch to time the 10- and 20-s exploration intervals.

Adult Ratings

To establish whether test objects could reliably be categorized as high, medium, or low in shape similarity relative to the target object, 15 adult judges (9 females and 6 males, all over the age of 18) rated the similarity of each test object to its target. The adults were presented with the objects one pair at a time in
random order. They were instructed to focus solely on shape similarity and ignore color differences and texture similarities. The rating scale ranged from 1 (not at all similar) to 7 (very similar). Participants did not physically handle any of the test stimuli.

The adult ratings followed the expected pattern. That is, the high-, medium-, and low-similarity test objects in each object set were perceived as significantly different in shape from one another, all t tests: \( p < .05 \), in the direction intended (see Appendix). That is, the high-similarity test objects were rated significantly higher in shape similarity to their targets than the medium-similarity objects, which were rated significantly higher in shape similarity to their targets than the low-similarity objects.

**Design**

Infants were randomly assigned to one of two groups: a no-label group \( (n = 30) \) or a novel-label group \( (n = 27) \). In each group, infants were presented with all target and test objects from the squeaking, ringing, and rattling object sets. For each infant, one of the three object sets (e.g., the rattling set) was presented in the surprised condition (i.e., the target object possessed the nonobvious property but the test objects did not), the second object set (e.g., the squeaking set) was presented in the interest control condition (i.e., neither the target nor the test objects possessed the nonobvious property), and the third object set (e.g., the ringing set) was presented in the predicted condition (i.e., both the target and the test objects possessed the nonobvious property). The specific object set assigned to the surprised, interest control, and predicted conditions was counterbalanced across infants.

The imitation task was comprised of four blocks of three trials each. One of the high, medium, low, and dissimilar matches for each of the object sets was randomly assigned to one of the four trial blocks. The order of presentation of test objects within a trial block was completely randomized, with the stipulation that each of the high-similarity, medium-similarity, low-similarity, and dissimilar matches for each of the squeaking, ringing, and rattling object sets was to be presented once only across the four blocks. Order of presentation of expectation condition within each trial block was counterbalanced across infants. Once the order of presentation of expectation condition was determined, this order was fixed for each trial block.

**Figure 1** Squeaking set.

**Figure 2** Ringing set.
Procedure

Infants were tested one at a time. They were brought into the laboratory and seated in their parents’ lap. The experimenter sat across a table from the infants. Before testing began, the experimenter instructed parents to interact with their infants as little as possible, and not to mention, point to, or give objects to their infants during the testing session (although parents were told they could repeat their infants’ name, if they started to become fussy). Parents were instructed to silently place objects back on the table within the infants’ reach if objects were dropped on the floor near the parents, or if the infants handed objects to the parents.

During the warm-up phase, the experimenter demonstrated a target property of the warm-up objects to the infants and asked the parents to do the same. After demonstrating the target property, parents silently handed the object to their infants so that they could imitate the actions observed. Regardless of whether infants imitated the actions, all three warm-up trials were presented and the experimenter proceeded to the test phase.

The experimenter began each trial of the test phase by presenting infants with one of the target objects from a given object set. The experimenter introduced the object either with a novel label (e.g., “Look at this blint!”) or without a label (e.g., “Look at this one!”). The experimenter repeated these phrases three times while demonstrating the nonobvious property of the target object five times (e.g., shaking the rattle). Only the properties of target objects in the surprised and predicted conditions were demonstrated (because the target objects in the interest control condition did not possess the property). Once the experimenter completed the demonstration, parents demonstrated the property of the target object twice, and passed the object to their infants. Infants were allowed to explore the target object for 10 s. After this time elapsed, the experimenter retrieved the object and placed it within the infants’ view but out of their reach. The experimenter then presented the infants with a test object that was either introduced with a novel label or no label (using the same phrases as described earlier). Infants were allowed to explore the test object for 20 s. This same procedure was repeated for each of the other 11 target–test object pairs. Although the target object from each object set was reintroduced to infants on each trial, parents only demonstrated the property the first time a target object from each object set was introduced. The experimenter, however, continued to demonstrate the target object’s property on each trial. If an object was dropped off the table or passed/thrown out of the infants’ reach during the session, the experimenter quickly placed the object back within their reach. Time lost due to these actions was not compensated for, because they were considered to be intentional actions of frustration or disinterest (see Oakes, Madole, & Cohen, 1991).

Parents were asked to fill out the MacArthur Communicative Development Inventory: Toddlers Version (MCDI; Fenson et al., 1991), a checklist of infants’ productive vocabulary, to ensure that the novel-label and no-label groups were matched for productive vocabulary size. Parents filled out the MCDI either before or after testing or at home. All MCDIs that were completed at home were mailed back within a 1-week period, except for two that were not returned. Infants’ productive vocabulary size, as measured by the MCDI, ranged from 3 to 498 words ($M = 106.55$ words, $SD = 127.67$). In the novel-label group, vocabulary size ranged from 7 to 355 words ($M = 93.33$ words, $SD = 104.57$), and in the no-label group, vocabulary size ranged from 1 to 498 words ($M = 122.40$ words, $SD = 151.59$). The groups did not differ significantly in productive vocabulary size, $t(41.42) = .81$, $p > .10$, $t$ test corrected for unequal variances.
Coding and Data Reduction

Coders, blind to the hypotheses of the experiment, recorded the frequency of actions performed on the target and test objects. To ensure that the coders could not distinguish the surprised, interest control, and predicted conditions from one another, only the experimenter’s back was visible on the videotapes (thus, coders could not see whether the experimenter performed an action with the target objects). In addition, all sessions were coded with the volume on the monitor turned off. Thus, the coders could not see whether the experimenter had demonstrated a target action on an object and could not determine whether objects actually made sounds when actions were performed by either the experimenter or the infants.

A detailed coding scheme for each target action was developed for each object set. In addition, coders recorded the frequency of object transfer actions on the test objects. Object transfer actions were defined as the performance of a target action from one object set (e.g., the patting action from the ringing set) on a test object from another set (e.g., the squeaking set). The target action for the squeezing set was defined by a squeezing motion, that is, the infants gripped and then compressed their fingers together on the object (not tapping the object, hitting the object on the table, shaking the object, or gripping it to look at it, or passing or throwing it to the experimenter or parent). A release of the fingers after squeezing the object was not counted as a second target action. If the object was squeezed with two hands together, one action was counted, unless the squeezing occurred at two separate points in time, which was then counted separately. The target action for the ringing set was defined by a tapping, hitting, or patting motion (not squeezing the object, hitting it on the table, shaking it, or gripping it to look at it; or passing or throwing it to the experimenter or parent). As noted earlier, the bells had a metal button on the top that was connected to a ringer inside; thus tapping, hitting, or patting the top of the object produced the ringing noise. A downward motion making contact with the object was considered one action, but an upward motion to bring the hand or finger back from the object was not counted as a second action. Touching and stroking the object gently with the hand to feel its texture or using a finger to poke it was not counted unless it was a swift “tapping” action. If the object was tapped with two hands, a target action was only counted once unless these taps occurred at two separate points in time, which was then counted separately. Finally, the target action for the rattling set was defined by a shaking motion with the wrist, the whole arm, or both in a back and forth or up and down motion (not tapping the object, squeezing it, hitting the table or a body part with it, or gripping it to look at it; or passing or throwing it to the experimenter or parent). If the infant performed a fluid shaking movement, then only one target action was counted. An “upward” shaking motion was counted as one action and a “downward” shaking motion was counted as another action only if a pause occurred between them (i.e., the motion was not a continuous or rebound effect of moving the wrist or arm one way, but a true separate attempt to shake the object in another direction). If the object was shaken with two hands together, the same criteria as outlined above applied.

To establish interrater reliability, 20% of the data (n = 15 participants) was coded twice. Intraclass correlations (ICCs) were used to establish the level of agreement between the two coders. ICCs provide a more conservative measure of assessing interrater reliability than a traditional Pearson correlation, as both the pattern of agreement and the level of agreement of raters are considered (Sattler, 1992). ICC coefficients for target- and test-object frequency ratings were both significant, ICC(360) = 1.00, p < .001, and ICC(360) = .99, p < .001, respectively. Thus, the two raters were in almost perfect agreement in their coding.

Data were screened to assess for outliers, normality, and skewness. Infants with frequency of target-action standard scores greater than 3 SDs above or below the mean in the surprised or interest control condition were eliminated from the data analyses (n = 7). There was no pattern to these outliers (i.e., they were not specific to any particular object set or level of object similarity): 4 infants were from the novel label group and 3 were from the no-label group. All analyses were performed with and without excluding the outliers; the same pattern was found in both cases. Thus, only the analyses with the outliers excluded is reported, to ensure that sampling was from a normal distribution. Infants who did not perform target actions on any of the test objects in the surprised condition were also eliminated from the final analyses (n = 8). These infants were excluded because they did not have any scores in the surprised condition, our condition of interest. Again, these findings were not specific to any particular object set: Two infants were from the novel-label group, and 6 were from the no-label group. Because it could not be established with certainty why infants did not perform any imitation in the surprised condition, their data were excluded from all subsequent analyses. In the warm-up trials, only 3 infants failed to imitate target actions on at least one of the three objects. These infants did perform target actions on both target and test objects in
the experimental task, however; therefore their data were included in the following analyses.

## Results

### Frequency of Target Action Data Analyses

The mean frequency of target actions performed on the different test objects in each label and expectation condition at each level of shape similarity are presented in Table 1. In the first set of analyses, we examined whether the target properties of the object stimuli were indeed nonobvious to infants. The number of target actions infants performed on test objects after having first seen a functional target object (the surprised condition) versus a nonfunctional target object (the interest control condition) were compared.

Dependent *t* tests were used to compare the frequency of target actions in the surprised condition at each level of shape similarity, collapsed across the novel-label and no-label groups. As can be seen in Table 1, the frequency of target actions performed in the

<table>
<thead>
<tr>
<th>Shape Similarity to Target</th>
<th>Novel label</th>
<th>No label</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>Surprised</td>
<td>3.00</td>
<td>1.67</td>
</tr>
<tr>
<td>Interest control</td>
<td>.00</td>
<td>.33</td>
</tr>
<tr>
<td>Predicted</td>
<td>5.17</td>
<td>3.57</td>
</tr>
<tr>
<td>Surprised</td>
<td>4.41</td>
<td>1.93</td>
</tr>
<tr>
<td>Interest control</td>
<td>.04</td>
<td>.19</td>
</tr>
<tr>
<td>Predicted</td>
<td>5.00</td>
<td>3.41</td>
</tr>
</tbody>
</table>

Note: Values in parentheses are standard deviations.

Interest control condition was low, and in some cases, no actions at all were performed. Thus, we collapsed across the novel-label and no-label groups because some cells had zero variance and other analyses indicated no significant influence of labeling group. As expected, infants performed significantly more target actions on the high-similarity object in the surprised condition (*M* = 3.67, *SD* = 3.64) than in the interest control condition (*M* = .02, *SD* = .13), *t*(56) = 7.53, *p* < .001. Similarly, infants performed significantly more target actions on the medium- and low-similarity objects in the surprised condition (*M* = 1.79, *SD* = 2.02, and *M* = 1.32, *SD* = 1.92, respectively) than in the interest control condition (*M* = .10, *SD* = .40, and *M* = .07, *SD* = .26, respectively), *t*(56) = 6.35, *p* < .001 and *t*(56) = 4.76, *p* < .001, respectively. In contrast, infants did not differ significantly in their performance of target actions on the dissimilar object in the surprised condition (*M* = .24, *SD* = 1.31) versus the interest control condition (*M* = .05, *SD* = .40), *t*(56) = 1.05, *p* > .29. The analyses indicated that the appearances of the objects did not suggest that the objects possessed the nonobvious properties. Instead, infants performed target actions on test objects only after they had been exposed to the properties of particular functional target objects during the testing session.

In the second set of analyses, the influence of shape similarity and novel labels on infants’ generalization of nonobvious properties within only the surprised condition was examined. It is important to note that the data from the interest control condition was not included in these analyses because this condition was not relevant to our hypotheses and some cells within the interest control condition had zero variance. However, as can be seen in Table 1, the frequency of
target actions performed on the test objects in the interest control condition did not vary by shape similarity. A 2 (group: novel label versus no label) × 4 (shape similarity: high, medium, low, dissimilar) mixed-model analysis of variance (ANOVA) was used to compare how often infants performed the target action on test objects as a function of shape similarity, presence or absence of a label, and expectation condition. Group was a between-subjects factor and shape similarity was a within-subjects factor. Because Mauchley’s test of sphericity was significant for the interaction previously seen on other object sets, the number of target actions was compared to the number of transfer actions performed on test objects in the surprised condition. We first summed the number of appropriate target actions each child performed on the high-, medium-, and low-similarity objects and compared that number with the total number of transfer actions performed on these same objects, using directional t tests. We collapsed across the novel-label and no-label groups because previous analyses indicated no significant effects for group. As expected, infants performed significantly more target actions than transfer actions on each object set: squeaking set: M = 5.43 target actions (SD = 4.41), M = 1.11 transfer actions (SD = 2.26), t(23) = 4.18, p < .01; ringing set: M = 6.29 target actions (SD = 4.04), M = .47 transfer actions (SD = .20), t(13) = 5.60, p < .01; rattling set: M = 6.19 target actions (SD = 6.65), M = 1.24 transfer actions (SD = 2.36), t(18) = 2.85, p < .01. These results indicate that infants were more likely to perform target actions than transfer actions on the test objects, rather than simply trying any type of action previously seen. Furthermore, the mean total number of transfer actions performed on each object set (that is, summed across the high-, medium-, and low-similarity objects) was approximately 1 or less, indicating that infants rarely performed object transfer actions.

Although incidences of object transfer actions were rare, we examined whether labeling the objects or the degree of shape similarity between the target and test objects influenced infants’ tendency to perform transfer actions, using a 2 (group: novel label versus no label) × 3 (shape similarity: high, medium, low) mixed-model ANOVA. Neither the main effects for label and shape nor their interaction were significant. Thus, there were no significant differences in transfer performance as a function of object shape similarity or object labels.

Object Transfer Data Analyses

In the fourth set of analyses, instances of object transfer within the surprised condition were examined. Recall that object transfer was defined as the performance of a target action from one object set (e.g., the patting action from the ringing set) on a test object from another set (e.g., the squeaking set). To examine whether infants were restricting target actions to appropriate test objects or simply trying any type of target action previously seen on other object sets, the number of target actions was compared to the number of transfer actions performed on test objects in the surprised condition. We first summed the number of appropriate target actions each child performed on the high-, medium-, and low-similarity objects and compared that number with the total number of transfer actions performed on these same objects, using directional t tests. We collapsed across the novel-label and no-label groups because previous analyses indicated no significant effects for group. As expected, infants performed significantly more target actions than transfer actions on each object set: squeaking set: M = 5.43 target actions (SD = 4.41), M = 1.11 transfer actions (SD = 2.26), t(23) = 4.18, p < .01; ringing set: M = 6.29 target actions (SD = 4.04), M = .47 transfer actions (SD = .20), t(13) = 5.60, p < .01; rattling set: M = 6.19 target actions (SD = 6.65), M = 1.24 transfer actions (SD = 2.36), t(18) = 2.85, p < .01. These results indicate that infants were more likely to perform target actions than transfer actions on the test objects, rather than simply trying any type of action previously seen. Furthermore, the mean total number of transfer actions performed on each object set (that is, summed across the high-, medium-, and low-similarity objects) was approximately 1 or less, indicating that infants rarely performed object transfer actions.

Although incidences of object transfer actions were rare, we examined whether labeling the objects or the degree of shape similarity between the target and test objects influenced infants’ tendency to perform transfer actions, using a 2 (group: novel label versus no label) × 3 (shape similarity: high, medium, low) mixed-model ANOVA. Neither the main effects for label and shape nor their interaction were significant. Thus, there were no significant differences in transfer performance as a function of object shape similarity or object labels.
Discussion

The results of this experiment indicate that infants will generalize a specific nonobvious property from a target to those test objects perceived as members of the same category. Infants performed significantly more target actions on test objects in the surprised condition than in the interest control condition, indicating that the appearance of the objects did not suggest the nonobvious properties. This finding also indicates that infants formed specific expectations about the nonobvious properties of test objects from knowledge gained during the testing session. Further evidence for the specificity of infants’ expectations was obtained from the dissimilar object trials and object transfer data. Infants rarely performed target actions on dissimilar objects, indicating that they did not see these objects as members of the same category as the target object. Similarly, incidences of object transfer were rare, indicating that infants only performed target actions specifically intended for a particular object set. Thus, infants were not simply making associations between imitations and objects in general—they only extended the nonobvious property when they considered it appropriate.

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

Unexpectedly, infants’ generalization of the nonobvious properties did not vary as a function of the presence or absence of a novel count noun label. This finding suggests that providing infants with conceptual information in the form of shared labels did not influence their inductive inferences. It is possible, however, that the manner in which the novel labels were presented in this experiment may have diluted their effect. Recall that the experimenter presented the novel labels only three times and uttered these labels at the same time as she demonstrated the novel sound property of the object. Thus, it is possible that infants did not attend to the object labels because they did not have enough exposure to the label, because the demonstration of the objects’ sound properties distracted them, or both. To address this possibility, a second experiment was conducted in which infants were provided with more exposure to the object labels and the presentation of the labels was separated from the presentation of the sound property.

EXPERIMENT 2

The design of Experiment 2 was similar to that of Experiment 1 with two notable exceptions. First, the dissimilar object trials were not included. The results of Experiment 1 clearly indicated that infants were not indiscriminately imitating target actions on any object. Second, the infants were provided with more exposure to the novel labels: as the experimenter uttered the novel label three times before the sound property was demonstrated, and three times while the sound property was demonstrated. Thus, the presentation of the label information was isolated from the presentation of the sound property. We predicted that infants would generalize the nonobvious property to objects that shared the same label, even if they shared little shape similarity with the target object.

Method

Participants

Data from 58 infants (31 girls, 27 boys; M = 19.21 months, SD = 1.17, range = 16.52–21.16 months) were included in the final sample. Eighteen additional infants were tested but were excluded from the final sample for the following reasons: excessive fussiness (n = 6), parental cueing (n = 2), experimenter error (n = 1), statistical outliers (n = 5; see Coding and Data Reduction section), and lack of performance of target actions on any of the test objects in the surprised condition (n = 9; see Coding and Data Reduction section). Infants were randomly assigned to one of two groups: a novel-label group (16 girls, 13 boys; M = 19.43 months, SD = 1.01, range = 17.49–20.85 months), and a no-label group (15 girls, 14 boys;
Materials

Stimuli. The target and test objects were identical to those used in Experiment 1 (as noted earlier, the dissimilar objects were not included).

Equipment. The equipment was also identical to that used in Experiment 1.

Design

The design was identical to that of Experiment 1, with one exception: because the dissimilar object was not included, there were a total of nine trials per participant (3 sets of 3 trials).

Procedure

The procedure was similar to that of Experiment 1, with two exceptions: (1) the dissimilar object was not presented, and (2) infants were directed toward target and test objects with phrases before the sound property demonstration, and then accompanying the sound property demonstration (in the surprised and predicted conditions). As stated previously, the experimenter labeled the target objects a total of six times, three times before the sound property was demonstrated, and three times while the sound property was demonstrated. The experimenter labeled test objects a total of four times. Thus, the novel-object labels in Experiment 2 were repeated twice as often as in Experiment 1. The target and test objects in the no-label group were introduced in the identical manner and the identical number of times as in the novel-label group except the experimenter used the word “one” in place of the novel label (e.g., “blint”). The target’s property was demonstrated the same number of times as in Experiment 1 (i.e., five times).

As in Experiment 1, parents completed the MCDI either after testing or at home. All MCDIs that were completed at home were mailed back within a 1-week period, except for one that was not returned. Infants’ productive vocabulary size, as measured by the MCDI, ranged from 4 to 669 words (M = 146 words, SD = 146.41). In the novel-label group, vocabulary size ranged from 4 to 441 words (M = 136 words, SD = 126.74), and in the no-label group, vocabulary size ranged from 5 to 669 words (M = 156 words, SD = 164.37). The groups were not significantly different in terms of productive vocabulary size, t(55) = .52, p > .10.

Coding and Data Reduction

Coders, blind to the experimental hypotheses, coded the frequency of target actions performed on test objects using the same criteria as in Experiment 1. Approximately 20% of the sessions (n = 16 participants) were coded a second time to assess interrater reliability. As in Experiment 1, ICC coefficients for target and test object frequency ratings were both significant, ICC(288) = .98, p < .001 and ICC(288) = .97, p < .001, respectively. Thus, the two raters were in almost perfect agreement in their coding.

As in Experiment 1, infants with standard scores greater than 3 SDs above or below the mean in the surprised or interest control condition were eliminated from the data analyses (n = 5). Again, there was no pattern to these outliers (i.e., they were not specific to any particular object set or level of object similarity); 4 infants were from the novel-label group, and 1 was from the no-label group. Analyses were performed with and without the exclusion of the outliers, and findings followed the same pattern. Infants who did not perform target actions on any of the test objects in the surprised condition were also eliminated from the final analyses (n = 9). Again, these findings were not specific to any particular object set; 6 infants were from the novel-label group, and 3 were from the no-label group. In the warm-up trials, only 1 participant failed to imitate target actions on at least one of the three objects. This participant did, however, perform target actions on target and test objects in the experimental task; these data were not eliminated from the following analyses.

Results

Frequency of Target Action Data Analyses

As in Experiment 1, the number of target actions infants performed on test objects after having first seen a functional target object (the surprised condition) versus a nonfunctional target object (the interest control condition) were compared first. The mean frequency of target actions performed on the test objects in each group and expectation condition at each level of shape are presented in Table 2. Dependent t tests were used to compare the frequency of target actions in the surprised condition at each level of shape similarity for each group separately (because other analyses indicated a significant influence of group; see following paragraph). In the novel-label group, infants performed significantly more target actions on the high-, medium-, and low-similarity objects in the surprised condition than in the interest control condition, t(28) = 8.93, p < .001; t(28) = 5.40, p < .001, and
Table 2  Experiment 2: Frequency of Target Actions Performed on Test Objects in the Novel-Label and No-Label Groups at Each Level of Shape Similarity within each Expectation Condition

<table>
<thead>
<tr>
<th>Shape Similarity to Target</th>
<th>Novel label Surprised (M, SD)</th>
<th>Novel label Interest control (M, SD)</th>
<th>Novel label Predicted (M, SD)</th>
<th>No label Surprised (M, SD)</th>
<th>No label Interest control (M, SD)</th>
<th>No label Predicted (M, SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>7.48 (4.11)</td>
<td>.28 (.65)</td>
<td>8.34 (9.15)</td>
<td>3.79 (3.04)</td>
<td>.00 (.19)</td>
<td>6.93 (6.60)</td>
</tr>
<tr>
<td>Medium</td>
<td>3.62 (3.26)</td>
<td>.21 (.62)</td>
<td>6.38 (6.93)</td>
<td>1.86 (2.70)</td>
<td>.10 (.31)</td>
<td>4.55 (5.36)</td>
</tr>
<tr>
<td>Low</td>
<td>2.86 (3.40)</td>
<td>.07 (.26)</td>
<td>4.76 (6.89)</td>
<td>1.14 (1.85)</td>
<td>.00 (.00)</td>
<td>2.48 (3.15)</td>
</tr>
</tbody>
</table>

Note: Values in parentheses are standard deviations.

$t(28) = 4.38, p < .001$, respectively. Similarly in the no-label group, infants performed significantly more target actions on the high- and medium-similarity objects in the surprised than in the interest control condition, $t(28) = 6.72, p < .001$ and $t(28) = 3.40, p < .01$, respectively. No statistic was computed for the low-similarity objects because infants did not perform any actions on the low-similarity objects in the interest control condition (thus, there was no variance in that cell). Thus, consistent with the results of Experiment 1, the appearance of the objects did not suggest the properties of the objects.

In the second set of analyses, the effect of shape and novel labels on infants’ generalizations of the object properties within the surprised condition were examined. A 2 (group: novel label versus no label) × 3 (shape similarity: high, medium, low) mixed-model ANOVA was performed on the data. The ANOVA yielded a significant main effect for group, $F(1, 56) = 20.18, MSE = 12.32, p < .001$. Infants performed significantly more target actions in the novel-label group ($M = 4.65, SD = 3.79$) than in the no-label group ($M = 2.26, SD = 2.53$), collapsed across all levels of shape similarity. The main effect for shape similarity was also significant, $F(1, 56) = 24.97, MSE = 8.58, p < .001$, but the Group × Shape Similarity interaction did not reach significance, $F(1, 56) = 2.14, MSE = 8.58, p = .12$.

To follow-up the significant main effect for shape, pairs of test objects in the surprised condition were compared by collapsing over the novel-label and no-label groups, using paired $t$ tests with a Bonferroni correction of $0.05 / 3 = .0167$. Consistent with the results of Experiment 1, infants performed significantly more target actions on high-similarity objects than on the medium-similarity and low-similarity objects, $t(57) = 4.88, p < .001$, and $t(57) = 6.70, p < .001$, respectively. There were no significant differences between the frequency of target actions performed on medium- and low-similarity objects $t(57) = 1.46, p > .10$. Thus, infants relied on a high level of shape similarity to generalize nonobvious properties more often than the other levels of shape similarity within the surprised condition.

In the third set of analyses, the effect of providing a novel label on infants’ generalization of nonobvious object properties was explored using planned comparisons. As in Experiment 1, we predicted that novel labels would promote inductive generalizations at all levels of shape similarity. These comparisons indicated that infants performed significantly more target actions in the novel-label group ($M = 7.48, SD = 4.11$) than in the no-label group ($M = 3.79, SD = 3.04$) on the high-similarity objects, $t(56) = 3.89, p < .001$. Similarly, infants also performed significantly more target actions in the novel-label group ($M = 3.62, SD = 3.26$) than in the no-label group on the medium-similarity object, ($M = 1.86, SD = 2.70$), $t(56) = 2.24, p < .05$. Finally, when presented with the low-similarity objects, infants performed significantly more target actions in the novel-label group ($M = 2.86, SD = 3.40$), than the no-label group ($M = 1.14, SD = 1.85$), $t(43.20) = 2.40, p < .05$ ($t$ test corrected for unequal variances). These results suggest that novel labels did guide infants’ inductive inferences across each level of object shape.

Object Transfer Data Analyses

In the final set of analyses, instances of object transfer in the surprised condition were examined. As in Experiment 1, the number of target actions versus the number of transfer actions performed on test objects in the surprised condition were compared first, separately for the novel-label and no-label groups (because previous analyses had indicated a significant effect for group), using directional $t$ tests. In the novel-label group, infants performed significantly more target actions than transfer actions on each object set: squeaking set: $M = 12.83$ target actions ($SD = 5.00$), $M = 2.63$ transfer actions ($SD = 3.79$), $t(9) = 4.73, p < .01$; ringing set: $M = 13.50$ target actions ($SD = 8.19$), $M = 1.23$ transfer actions ($SD = 1.85$), $t(9) = 3.98, p < .01$; rattling set: $M = 9.59$ target actions ($SD = 2.87$), $M = .26$ transfer actions ($SD = 1.43$), $t(8) = 9.59, p < .01$. Similarly, infants in the no-label group performed significantly more target actions than transfer actions on each object set: squeaking set: $M = 4.39$ target actions ($SD = 2.53$), $M = 1.09$ transfer actions ($SD = 2.17$), $t(10) = 2.89, p < .05$; ringing set: $M = 9.67$ target actions ($SD = 5.35$), $M = .56$ transfer actions ($SD = 1.67$), $t(8) = 4.64, p < .01$; rattling set: $M = 4.41$.
target actions \((SD = 2.72), M = 2.04\) transfer actions \((SD = .68), t(8) = 1.90, p < .05\).

A 2 (group: novel label versus no label) \(\times 3\) (shape similarity: high, medium, low) mixed-model ANOVA was performed on the transfer data. Neither the main effects for group and shape nor their interaction were significant. Thus, consistent with the results of Experiment 1, infants were more likely to perform target actions than transfer actions on the test objects, and incidences of object transfer did not vary as a function of object shape similarity or presence or absence of a label.

**Discussion**

Consistent with Experiment 1, infants in the present experiment performed significantly more target actions on test objects in the surprised condition than in the interest control condition, indicating that they expected the test objects to share the same nonobvious property as the target object. As in Experiment 1, infants performed significantly more target actions on objects with a high level of shape similarity in both the novel- and no-label groups, suggesting that they expected objects that were highly similar in shape to share nonobvious properties.

In contrast to Experiment 1, the presence of a novel label did influence infants’ tendency to generalize nonobvious object properties. At each level of shape similarity, infants performed more target actions on test objects in the novel-label group than in the no-label group. Thus, when the labels were clearly isolated from the sound demonstrations, infants were able to attend to the label information and use it to guide their inferences, particularly when objects were highly dissimilar in shape. The presence of labels, however, did not lead infants to completely overlook shape information. Infants still performed significantly more target actions on the high-similarity objects than on either the medium- or the low-similarity objects in the novel-label group. Thus, object shape similarity continued to play a role in infants’ inductive strategies.

In Experiment 3, the possibility that familiar labels, rather than novel labels, might lead infants to disregard object similarity when generalizing nonobvious properties was explored. When presented with a novel label, infants have to attend to both the task of learning a novel word for a novel object and learning about a novel nonobvious property of that object. In contrast, when presented with familiar labels, infants have to integrate the nonobvious property into their existing category and decide whether another object will also share this newly learned property. Thus, we reasoned that the use of familiar labels may place fewer cognitive demands on infants, perhaps allowing them to attend more to the conceptual information provided by the label. Research with infants has indicated that by 12 months of age, familiar labels promote the formation of conceptually based categories (Waxman & Markow, 1995). Similarly, studies have demonstrated that 2½-year-olds will use familiar labels as an inductive base when generalizing properties of pictured natural kind objects, even when a target and test object differ in perceptual appearance (e.g., Gelman & Coley, 1990). Thus, we hypothesized that infants would perform as many target actions on medium- and low-similarity objects as on high-similarity objects when these objects were labeled with a familiar count noun.

**EXPERIMENT 3**

**Method**

**Participants**

Data from 29 infants (13 girls, 16 boys; \(M = 18.95\) months, \(SD = 1.29\), range = 16.56–21.39 months) were included in the final sample. Nine additional infants were tested but were excluded from the final sample for the following reasons: excessive fussiness \((n = 4)\), parental cueing \((n = 1)\), statistical outliers \((n = 3)\; \text{see Coding and Data Reduction section}\), and lack of performance on any of the test objects in the surprised condition \((n = 4)\; \text{see Coding and Data Reduction section}\). None of the infants had participated in Experiments 1 or 2.

**Materials**

*Stimuli.* The target and test objects were identical to those used in Experiment 2.

*Equipment.* This equipment was identical to that used in Experiments 1 and 2.

**Design**

This design was identical to that of Experiment 2, except that all infants participated in the familiar label group.

**Procedure**

The procedure was identical to that of the novel-label group of Experiment 2, except that the target and test objects were labeled with familiar count noun labels. The labels *ball* (squeaking set), *block* (ringing set), and *spoon* (rattling set) were used to introduce target and test objects (e.g., “Look at this
spoon!"). These familiar labels were chosen for two reasons: First, because the objects could reasonably be characterized by these labels (see Figures 1, 2, and 3) and second, because comprehension and production norms on the MCDI (Fenson et al., 1994) indicated that these words are understood and likely produced by infants in the age range of our study. That is, data from the normative study showed that the words we used are understood by 50% of the normative sample by the following ages: ball by 10 months, block by 15 months, and spoon by 13 months. It is important to note that none of the familiar labels we used labeled categories with the same nonobvious properties as our objects.

As in Experiments 1 and 2, all MCDIs were completed either after testing or at home, and those that were completed at home were mailed back within a 1-week period (with a 100% return rate). Productive vocabulary size, as measured by the MCDI, ranged from 4 to 459 words (M = 123 words, SD = 125.71). Participants in Experiment 3 and participants in the no-label condition of Experiment 2 (the comparison group in portions of the Results section to follow) were not significantly different in terms of productive vocabulary size, t(56) = .87, p > .10. Examination of completed MCDIs revealed that 86% of infants in the familiar-label group produced at least one of the three familiar labels, 48% produced at least two of the three familiar labels, and 34% produced all three of the familiar labels. Although not all of the children were producing all three words, we remain confident that these labels were familiar to children for the following reasons. First, the data from the normative MCDI study indicates that the three words used were all understood by 50% of the normative sample by 15 months of age, which is 1½ months younger than the youngest child tested in our studies (Fenson et al., 1994). Second, the familiar label group was divided into two subgroups (infants who were producing two or three of the familiar labels and infants who were producing none or one of the familiar labels) and these two subgroups were then compared on the frequency of target actions performed on the objects in the surprised condition at each level of shape. These analyses revealed no significant differences in performance. Thus, children who were producing at least two of the three words did not differ in their performance from children who were producing none or one of the three words. Finally, infants who produced none of the familiar words (n = 4) were compared with infants who produced all three of the familiar labels (n = 10) and no significant differences were found in the frequency of target actions performed on the objects in the surprised condition at each level of shape.

Codings and Data Reduction

Coders, blind to the experimental hypotheses, coded the frequency of target actions performed on target and test objects using the same criteria as in Experiments 1 and 2. Approximately 20% of the sessions (n = 8 participants) were coded a second time to assess interrater reliability. ICC coefficients for target and test object frequency ratings were both significant, ICC(144) = .99, p < .001 and ICC(144) = .99, p < .001, respectively. Thus, as in the two previous experiments, the two raters were in almost perfect agreement in their coding.

Infants with frequency of target action scores greater than 3 SDs above or below the mean were eliminated from the final analyses (n = 3). There was no pattern to these outliers (i.e., they were not specific to any particular object set or level of object similarity). Analyses were performed with and without the exclusion of the outliers, and findings followed the same pattern. Infants who did not perform target actions on any of the test objects in the surprised condition were also eliminated from the final analyses (n = 4). Again, these findings were not specific to any particular object set. In the warm-up trials, 100% of participants imitated target actions on at least one of the three objects.

Results

Frequency of Target Action Data Analyses

The mean frequencies of target actions performed on the different test objects in each expectation condition at each level of shape similarity are presented in Table 3. Infants did not perform any target actions on the objects in the interest control condition. Thus, we were not able to compute any statistical comparisons between the surprised and interest control condition. Nonetheless, the means indicated that the object properties were nonobvious, consistent with the results of Experiments 1 and 2.

The effect of shape on infants’ generalizations of the nonobvious object properties when objects were labeled with familiar count nouns was then examined. A repeated-measures ANOVA indicated no significant effect for shape, F(2, 56) = 1.07, MSE = 4.32, p > .35. As can be seen in Table 3, the number of target actions performed by infants did not vary as a function of the test objects’ degree of shape similarity to their target, in contrast to the results of Experiments 1 and 2. That is, infants generalized the nonobvious property to the medium and low similarity objects as frequently as they did to the high-similarity objects.
Object Transfer Data Analyses

As in Experiments 1 and 2, instances of object transfer in the surprised condition were examined, using directional t tests. As expected, infants performed significantly more target actions than transfer actions on each object set: squeaking set: M = 5.67 target actions (SD = 3.96), M = 1.77 transfer actions (SD = 2.63), t(9) = 2.32, p < .05; ringing set: M = 6.40 target actions (SD = 2.87), M = .40 transfer actions (SD = .97), t(9) = 5.22, p < .01; rattling set: M = 6.11 target actions (SD = 1.80), M = 2.48 transfer actions (SD = 3.85), t(8) = 2.34, p < .05. A 2 (group: novel label versus no label) × 3 (shape similarity: high, medium, low) mixed-model ANOVA was performed on the transfer data. Neither the main effects for group and shape nor their interaction were significant. Thus, consistent with the results of the previous experiments, infants were more likely to perform target actions than transfer actions on the test objects and incidences of object transfer did not vary as a function of object shape similarity or presence or absence of a label.

Cross-Experiment Comparisons

Given that Experiments 2 and 3 shared the same general procedure, participant characteristics, and experimenter, the frequency of target actions performed when infants were provided with familiar labels for test objects (Experiment 3) was compared with the frequency of target actions performed when infants were provided with either no labels or novel labels for test objects (Experiment 2). The frequency of target actions performed on surprised condition test objects in the familiar-label and no-label groups was compared first. The α level remained at .05 for each test, because they were each separate, planned t test comparisons. Infants performed significantly more target actions in the familiar-label group (M = 2.35, SD = 2.02) than in the no-label group (M = 1.14, SD = 1.85) on the low-similarity objects, t(56) = 2.37, p < .025.

No significant differences were found between the familiar-label group and no-label group for either the high-similarity objects, M = 3.00 familiar label (SD = 1.67), M = 3.79 no label (SD = 3.04), t(56) = 1.23, p > .10; or the medium-similarity objects, M = 2.28 familiar label (SD = 2.34), M = 1.86 no label (SD = 2.70), t(56) = 0.62, p > .10. These results suggest that familiar labels guided infants’ inferences about nonobvious object properties, particularly when an object’s shape was highly dissimilar to the target object. Furthermore, these results may indicate that familiar labels may not be as useful for guiding infants’ inferences about the high- and medium-similarity objects.

The frequency of target actions performed on surprised condition test objects in the familiar-label group was then compared to those in the novel-label group at each level of shape similarity. These comparisons made it possible to assess whether novel labels and familiar labels differed in their influence on infants’ generalizations and any possible interaction between shape similarity and label type. Planned t tests indicated that infants performed significantly more target actions on the high-similarity objects in the novel-label group (M = 7.48, SD = 4.11) than in the familiar-label group (M = 3.00, SD = 1.67), t(56) = 5.44, p < .001. There were no significant differences between the novel-label group and familiar-label group for either medium-similarity objects, M = 3.62 novel label (SD = 3.26), M = 2.28 familiar label (SD = 2.34), t(56) = 1.81, p > .05; or low-similarity objects, M = 2.86 novel label (SD = 3.40), M = 2.34 familiar label (SD = 2.02), t(56) = .70, p > .10.

Discussion

The results of Experiment 3 were strikingly different from those of Experiments 1 and 2. When infants were provided with familiar labels for novel target and test objects, shape similarity did not influence their inductive generalizations. That is, infants performed as many target actions on the medium- and the low-similarity objects as on the high-similarity objects. Thus, when labels refer to categories that are familiar to infants, they are able to use these labels to make inductive generalizations, overlooking the shape information available. In contrast, the degree of shape similarity between the target object and the test objects continued to influence infants’ generalizations when they were provided with novel labels for these objects in Experiment 2. This difference in attention to shape information between the novel-label and the familiar-label groups is likely due to differences in the cognitive demands of the task. In the novel-label group, infants were faced with the challenges of both
learning a novel word for a novel object and learning about a novel nonobvious property of that object. In contrast, in the familiar-label group, the labels were already well known to infants and thus, their primary task was to learn about the nonobvious property and infer that another member of that same category would also share this newly learned property. Thus, infants in the familiar-label group were able to disregard shape information and generalize the nonobvious properties solely on the basis of shared labels.

Consistent with the results of Experiment 2, infants generalized the nonobvious property more frequently to the low-similarity test object when it was labeled with a familiar count noun than when it was not labeled. This result suggests that when an object is dissimilar in appearance to a target, both shared familiar labels and novel labels help infants to generalize the property. We should note that an alternative explanation for these findings is that infants simply increased their rates of target actions on hearing the labels, rather than recognizing that shared object labels provide information about shared conceptual category. If this were the case, however, we would not expect to find the differences between the novel-label group and the familiar-label group that we did find. That is, we would expect infants to show similar patterns of generalization in the familiar-label and novel-label groups, relative to the no-label group, because labels simply acted to increase infants’ attention to the objects. Thus, infants’ reliance on the shared labels to generalize nonobvious properties does appear to indicate that they are using these labels as indicators of shared conceptual category. Supporting evidence for our finding is provided by a recent conference presentation by Desjardins and Baldwin (1992), using a similar paradigm. They found that 20- to 22-month-old infants generalized nonobvious object properties when objects were labeled with the same novel count noun but not when objects were labeled with different count nouns (e.g., when one object was labeled a toma whereas the second object was labeled a peri).

GENERAL DISCUSSION

The present studies were designed to examine the role of object shape similarity and object labels in promoting inductive inferences in infancy. In three experiments, 16- to 21-month-old infants were presented with novel target objects that either possessed (surprised and predicted conditions) or did not possess (interest control condition) a nonobvious sound property in novel-label and no-label groups (Experiments 1 and 2), or a familiar-label group (Experiment 3). Infants were then presented with test objects of varying degrees of shape similarity to the target. These test objects either possessed the same property (predicted condition) as the target or were disabled so that they could not make the sound (surprised and interest control conditions). Infants’ imitations of target actions on the various test objects were measured to assess their expectations about the generalizability of the nonobvious property.

The results of our studies provide four major insights into the nature of inductive reasoning during late infancy. First, the results of all three experiments provide evidence that infants between 16 and 21 months of age will form expectations about shared properties of novel objects after only a 10-s experience with a functional target object. Furthermore, infants will extend a specific nonobvious property from a target exemplar to other objects perceived as members of the same category, consistent with the results of previous studies (e.g., Baldwin et al., 1993; Desjardins & Baldwin, 1992; Mandler & McDonough, 1996, 1998; McDonough & Mandler, 1998). Thus, the speed and efficiency with which infants formed expectations about the nonobvious properties of novel objects in our studies provides further evidence that infants possess well-developed inductive reasoning abilities.

Second, our findings indicate that infants will rely on shape similarity to generalize nonobvious object properties, in the absence of other information about object kind. Unlike previous research in this area, we examined the degree of shape similarity necessary for infants to perceive objects as members of the same category, and thus, to generalize a nonobvious object property. In Experiments 1 and 2, infants were more likely to generalize a nonobvious object property to objects that were highly similar in shape rather than to objects that were less similar in shape, when objects were presented without a label. Interestingly, infants did not expect objects that shared only the same texture to also share nonobvious properties. These findings are consistent with a large body of evidence indicating shape similarity is often privileged over other types of other perceptual properties in young children’s object categorization (e.g., Baldwin, 1989; Graham & Poulin-Dubois, 1999; Landau et al., 1988, 1992). Furthermore, this finding bears on a current debate in the literature regarding the role of shape in young children’s concepts. Some researchers suggest that children attend to shape similarity when categorizing objects, because their categories reflect primarily their attention to salient perceptual features (e.g., Jones & Smith, 1993; Landau et al., 1988; Smith, Jones, & Landau, 1996). In contrast, other researchers suggest that children attend to shape information when categorizing objects because it serves as a perceptu-
ally available cue to the underlying structure of a category (e.g., Bloom 2000; Gelman & Diesendruck 1999; Gentner & Namy, 1999). Our findings provide clear evidence that infants expect that objects that share a high degree of shape similarity will also share other “deeper” characteristics, suggesting that they appreciate that shape similarity is an index of object kind (for similar results with preschoolers, see Davidson & Gelman, 1990; Florian, 1994; Gentner & Imai, 1994).

Third, our results indicate that infants expect objects that are labeled with the same novel count noun to share the same nonobvious property, even if they are highly dissimilar in shape. In Experiment 2, when objects were given the same novel label, infants performed more target actions on the test objects at all levels of shape similarity (versus when they were not labeled). Thus, infants relied on the shared novel-object labels to guide their inferences about object properties (see also Desjardins & Baldwin, 1992). It is important to note that infants performed more target actions on the high-similarity objects than on the medium- and low-similarity objects, even in the novel-label group. This finding suggests that shape information continues to influence infants’ inferences even when other information about category membership, in the form of a shared novel label, is available. Infants’ continued attention to shared shape similarity in the novel-label group is consistent with the results of studies with preschool-age children (e.g., Davidson & Gelman, 1990; Gentner & Imai, 1994; Shipley, 1993). For example, Florian (1994) found that both perceptual similarity and shared conceptual information (novel labels or conceptual attributes) guided 4-year-old children’s inferences about novel and familiar animals. However, also consistent with the results of research with preschoolers (e.g., Gelman & Coley, 1990; Gelman & Markman, 1986, 1987), infants expected objects that were highly dissimilar in appearance to also share nonobvious properties if they shared the same label.

Finally, our findings demonstrate that when a novel object is labeled with a familiar count noun, infants will disregard shape information and rely solely on the label to generalize the nonobvious property. In Experiment 3, infants performed as many target actions on low- and medium-similarity objects as on high-similarity objects when objects were labeled with familiar nouns. These findings provide the first clear evidence that infants, like preschoolers, will overlook perceptual similarity and expect those objects that share the same label to possess the same nonobvious property. This finding is particularly compelling given that the nonobvious properties of our objects typically would not have been associated with the familiar category labels. That is, infants would not have expected objects labeled with the term spoon to rattle when shaken. Therefore, infants had to integrate this new information (i.e., the nonobvious sound property) into their existing category knowledge and reason that another member of that same category, as indexed by a shared label, will also share this newly learned property. These results suggest that when the objects were treated like members of a familiar kind, perceptual similarity did not influence infants’ expectations about shared nonobvious properties. This finding is consistent with the results of Mandler and McDonough (1996, 1998) who also found that infants were able to overlook perceptual similarities and make inferences on the basis of familiar underlying kind. For example, in Mandler and McDonough (1996), children extended an unobservable property of “being able to drink from a cup” from a bird to a cat (same underlying kind as the bird, different appearance), rather than to an airplane (same appearance as the bird, different underlying kind).

Taken together, our findings lead to the following conclusions regarding the contribution of perceptual similarity and conceptual information to infants’ inductive inferences. First, infants will rely on perceptual similarity, more specifically shared shape similarity, to guide their inferences about novel objects’ nonobvious properties, when no other information about category membership is available. Second and more importantly, when infants are provided with information about conceptual category membership in the form of shared object labels, perceptual information is either attenuated in importance (in the case of novel labels) or disregarded (in the case of familiar labels). Our findings thus provide clear evidence that infants can form novel categories and make inductive inferences about novel properties based on a conceptual notion of object kind (see also Mandler & McDonough, 1996, 1998; McDonough & Mandler, 1998). Furthermore, our findings indicate that infants as young as 16 months of age appreciate the conceptual information conveyed by object labels. That is, infants, like preschoolers, appear to recognize that count noun labels supply information about underlying object kind and, furthermore, that members of the same kind share nonobvious properties. These findings add to a growing body of literature indicating that naming can foster infants’ formation of object categories (e.g., Balaban & Waxman, 1997; Graham, Baker, & Poulin-Dubois, 1998; Waxman, 1999a, 1999b; Waxman & Hall, 1993; Waxman & Markow, 1995) and, moreover, provide evidence that labels enhance the inductive potential of categories for infants. Finally, our findings suggest that there is continuity in induc-
tive reasoning abilities across the infancy and preschool periods. That is, infants, like preschoolers, will overlook perceptual similarity and generalize nonobvious properties on the basis of object kind (e.g., Gelman & Coley, 1990; Gelman & Markman, 1986, 1987). In summary, the present experiments have advanced our understanding of young children’s inductive abilities, demonstrating that infants are able to use shape similarity and count noun label information to make inferences about nonobvious object properties. The results of these experiments suggest a number of important directions for future research into early inductive abilities. First, it remains to be established whether young infants who are just beginning to acquire words can rely on object labels to guide their inferences about object properties. Second, it remains to be seen whether count noun labels are privileged in guiding infants’ inferences about nonobvious object properties or whether words from other form classes (e.g., adjectives) and other nonlinguistic stimuli (e.g., gestures) also promote inductive inferences. We are currently exploring these issues and expect that delineating the conditions under which infants utilize various kinds of linguistic and object information in their inductive judgements will shed light on the developmental processes underlying inductive reasoning during infancy.

**ACKNOWLEDGMENTS**

This research was supported by operating grants awarded to the second author from the Natural Sciences and Engineering Research Council of Canada (NSERC) of Canada and the University of Calgary Research Grants Committee. The first author was supported by graduate fellowships from NSERC of Canada and the Alberta Heritage Foundation for Medical Research. The authors thank the parents, children, and adults who participated in the studies as well as the staff at the following Calgary Regional Health Authority clinics for their kind assistance: Thornhill, North Hill, Ranchlands, and Midnapore. They also thank Tamara Demke, Carlene Adamache, Jennifer Storms, and Joanna Card for their assistance with this research and Christopher Sears and Suzanne Hala for their comments on experimental design. D. Geoffrey Hall, Dare Baldwin, and an anonymous reviewer provided the authors with invaluable comments on a previous version of this manuscript. Experiments 1 and 2 were included in a thesis submitted by Andrea N. Welder to the University of Calgary in partial fulfillment of the requirements for the Master of Science degree. Some of the data from Experiment 1 were presented at the annual meeting of the Canadian Society for Brain, Behaviour, and Cognitive Science, Edmonton, Alberta, Canada, June 1999. Data from Experiments 2 and 3 were presented at the International Conference on Infant Studies, Brighton, U.K., July 2000.

**ADDRESSES AND AFFILIATIONS**

Corresponding author: Susan A. Graham, Department of Psychology, University of Calgary, 2500 University Drive NW, Calgary, AB, T2N 1N4, Canada; e-mail: grahamss@ucalgary.ca. Andrea Welder is also at the University of Calgary.

**APPENDIX**

**Adult Ratings for the Shape Similarity Judgments of High-, Medium-, and Low-Similarity Test Objects Compared with a Corresponding Target**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Mean Ratings</th>
<th>Paired t test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bell</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-/medium-similarity</td>
<td>H: 6.13 M: 3.87</td>
<td>t(14) = 6.86, p &lt; .001</td>
</tr>
<tr>
<td>High-/low-similarity</td>
<td>H: 6.13 L: 1.67</td>
<td>t(14) = 12.76, p &lt; .001</td>
</tr>
<tr>
<td>Medium-/low-similarity</td>
<td>M: 3.87 L: 1.67</td>
<td>t(14) = 5.98, p &lt; .001</td>
</tr>
<tr>
<td>Rattle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-/medium-similarity</td>
<td>H: 6.20 M: 3.45</td>
<td>t(14) = 4.77, p &lt; .001</td>
</tr>
<tr>
<td>High-/low-similarity</td>
<td>H: 6.20 L: 2.40</td>
<td>t(14) = 9.13, p &lt; .001</td>
</tr>
<tr>
<td>Medium-/low-similarity</td>
<td>M: 3.45 L: 2.40</td>
<td>t(14) = 2.26, p &lt; .05</td>
</tr>
<tr>
<td>Ball</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-/medium-similarity</td>
<td>H: 6.33 M: 3.93</td>
<td>t(14) = 6.87, p &lt; .001</td>
</tr>
<tr>
<td>High-/low-similarity</td>
<td>H: 6.33 L: 1.60</td>
<td>t(14) = 12.75, p &lt; .001</td>
</tr>
<tr>
<td>Medium-/low-similarity</td>
<td>M: 3.93 L: 1.60</td>
<td>t(14) = 6.47, p &lt; .001</td>
</tr>
</tbody>
</table>

Note: Ratings were on a 7-point scale (1 = not at all similar, 7 = very similar).

**REFERENCES**


Quinn, P. C., & Eimas, P. D. (2000). The emergence of repre-


