Received: 19 December 2020 | Accepted: 5 January 2022

DOI: 10.1111/infa.12459

RESEARCH ARTICLE

INFANCY THE OFFICIAL JOURNAL OF THE INTERNATIONAL CONGRESS OF INFANT STUDIES WILEY

Discovering category boundaries: The role of comparison in infants' novel category learning

Jelena Sučević[1] | **Nadja Althaus**[2] | **Kim Plunkett**[1]

[1]Department of Experimental Psychology, University of Oxford, Oxford, UK

[2]School of Psychology, University of East Anglia, Norwich, UK

Correspondence
Jelena Sučević, Department of Experimental Psychology, University of Oxford, Oxford, UK.
Email: jelena.sucevic@psy.ox.ac.uk

Abstract

A key question in categorization is how infants extract regularities from the exemplars they encounter. Detecting similarities and dissimilarities across items is vital in order to determine category-relevant features. Previous research found evidence that infants acquire a single category more easily with paired presentations in comparison with single presentations (Oakes & Ribar, 2005, *Infancy*, 7, 85; Oakes & Kovack-Lesh, 2007, *Cogniție, Creier, Comportament / Cognition, Brain, Behavior*, XI, 661). Here, we focus on infants' acquisition of a category *contrast*, that is, when they are exposed to two categories. In an eye-tracking study, we examined 10-month-old infants' ability to learn two novel visual categories when presented with one item at a time and with items in pairs. Infants were familiarized with pairs of items from the same category or with pairs of items from different categories (cross-category pairs). Using a linear model with a priori contrasts, we show that infants' learning is directly related to the opportunity for category comparison: There is no evidence of category learning in the single-item condition, improved performance when familiarized with same-category pairs, and finally robust category learning when familiarized with cross-category pairs. We conclude that comparison which involves items from different categories promotes category formation, by highlighting differences and promoting a discovery of category boundaries.

1 | INTRODUCTION

Categorization represents the ability to form groups of objects, which involves understanding which items are "the same kind of thing" despite perceivable differences, but also deciding which items are *not* the same kind, even though they may appear similar in some aspects (Mareschal & Quinn, 2001). In other words, to discover categories, infants must detect what is shared among members on the one hand, and what sets non-members apart on the other hand. For instance, even though a Labrador and a Chihuahua look notably different, both belong to the category of dogs. At the same time, although a Chihuahua and a Himalayan cat look similar, they do not belong to the same basic-level category. Such distinctions can be made by infants already around 4 months of age (e.g., Eimas & Quinn, 1994), even on the basis of exposure to visual images alone. However, such instances of success in early perceptual categorization are often obtained when infants are familiarized with pairs of items, whereas familiarization with the very same items one-by-one may fail (Oakes & Ribar, 2005). A substantial body of research has shown that comparison of category exemplars, which is only possible when at least two exemplars are presented simultaneously, promotes categorization in infants (Kovack-Lesh et al., 2008; Oakes & Kovack-Lesh, 2007; Oakes & Ribar, 2005).

1.1 | Why does presenting items in pairs benefit infant category learning?

One of the factors that might make category acquisition difficult for young infants is limited memory capacity. As young infants have limited short-term memory capacity (Kwon et al., 2014; Ross-Sheehy et al., 2003), difficulties in maintaining and updating information in short-term memory when presented with one item at a time may lead to a failure to extract category-relevant information. In contrast, presenting pairs of items provides the opportunity to examine two items side by side, thereby reducing short-term memory demands, and facilitating the identification of category-relevant information.

When only one item is presented at a time, infants must compare the current item with a mental representation of the previously presented item. Given infants' limited memory capacity, representations retained in working memory might lack a sufficient level of detail to conduct a suitable comparison. In contrast, when two items are available in the visual field, the infant has the opportunity to immediately compare items and discover similarities and differences. In particular, to perform a comparison in the single-item condition, the infant must maintain the representation of the previously presented item for much longer than in the paired condition: When one picture is replaced by another in the single-item condition, there will be a delay of seconds (or even minutes), as opposed to merely the duration of a saccade when shifting gaze from one picture to another when both pictures are visually available. When seconds pass between picture presentations, the visual representation of the first picture will inevitably deteriorate. The difficulty of single-item presentation is exacerbated by the fact that more trials are needed to present the same range of items as compared to paired presentation, further increasing the memory load. Consequently, a comparison between a sparse representation of the previously presented item with the currently viewed item in the single-item condition may fail to highlight relevant similarities and differences and therefore undermine discovery of the category-relevant information. In the paired presentation condition, infants can compare detailed representations of individual exemplars as they are both visually available and so are better able to form category representations.

Several studies have provided evidence of the benefits of paired presentation in infants' category learning (Oakes, 2012; Oakes & Kovack-Lesh, 2007; Oakes & Ribar, 2005). For instance, Oakes and Ribar (2005) presented 4-month-old infants with the same set of pictures of cats and dogs, either one at a time or in pairs. They found that 4-month-old infants failed to distinguish the categories after being familiarized with one item at a time, but succeeded after being familiarized with these items in pairs. Paired presentation was found to be beneficial for category learning in older infants, too. Although Oakes and Kovack-Lesh (2007) found that 6-month-old infants are able to distinguish categories of cats and dogs after being familiarized with a single item at a time, for categories that involve highly variable items, such as land and sea animals, 6-month-olds fail in a single-item condition, but succeed when presented with the very same items in pairs. The authors suggest that presenting items in pairs helps with distinguishing category contrasts at any age. It should be further noted that presenting category exemplars one item at a time inherently involves more trials and a consequent increase in memory load as compared to paired presentation.

To the best of our knowledge, all studies that compare successive and paired presentation in infants have tended to focus on categories that infants may have experienced in their everyday environments (Kovack-Lesh et al., 2014; Oakes et al., 2009; Oakes & Ribar, 2005; Younger & Furrer, 2003). It remains, therefore, unclear whether the facilitative effects of paired presentation on categorization are present only when infants need to *reactivate existing* category knowledge (e.g., cats) or whether the opportunity to compare can also influence formation of entirely novel categories. Paired presentation of familiar objects may serve to remind infants of the categories they have already formed, rather than facilitating the category formation process itself. The present study aims to fill this gap and determine how single and paired presentation conditions affect the process of acquiring novel categories. Furthermore, existing studies that explored the role of comparison in infant categorization involved initial exposure to a *single* category: Every participant was familiarized with exemplars from one category (e.g., pictures of dogs). A subsequent test trial evaluated novelty preference for an out-of-category item (e.g., preference for a cat over a dog). In real life, novel category learning probably rarely involves encountering many members of the same category in swift succession. Rather than incorporating successive items into the same category, infants need to track the co-occurrence of features over time, potentially forming multiple, contrasting category clusters. In the present study, we examine the role of exemplar comparison in infants' formation of such contrastive categories, by familiarizing 10-month-olds with a single item at a time or with pairs of items. To determine whether comparison emphasizing similarities or differences promotes categorization, we presented infants with pairs of items from the same category or with pairs of items from different categories.

The present study focused on infants' ability to learn correlation-based, that is, statistical categories. This class of categories is thought to be particularly important because it reflects the structure of natural categories in which objects have a high correlational structure and features do not occur independently of each other (Rosch et al., 1976). For instance, animals with feathers usually have wings, whereas furry animals rarely do. The ability to detect this relational information, that is, that certain features co-occur, is considered to play a central role in category formation (Younger & Cohen, 1986) and underlies categorization in adults (Rosch et al., 1976). Previous research has investigated infants' ability to learn correlation-based categories and has revealed that by 10 months of age, infants are able to encode these feature correlations and use them to form 2 distinct categories (Mather & Plunkett, 2011; Plunkett et al., 2008; Younger, 1993; Younger & Cohen, 1986). Given that individual features are not prognostic for these categories, but rather correlations between the features provide the basis for category formation, the process of comparison across category exemplars must play a particularly important role in promoting the discovery of these correlations.

1.2 | The role of same-category and cross-category comparisons in category learning

Comparing two items that are available in the visual field supports the discovery of commonalities and identification of differences among the presented items. Presenting two similar items may promote detection of similarities among the presented items, which can be particularly useful for organizing exemplars within a category and discovering an internal category structure. In contrast, presenting two dissimilar items can promote the detection of differences and aid the identification of key contrasts between categories. As a result, this can lead to the discovery of a boundary between different categories. For instance, presenting two cats simultaneously (e.g., a Persian cat and Siamese cat) can help discern that both have pointy ears and long tails. In contrast, presenting a Persian cat and a Labrador can promote identifying differences in eye separation or shape of claws.

Both the detection of similarities and differences are necessary for organizing items into categories. There are contrasting views as to which of the two processes is central to categorization. While some authors argue that similarity detection is at the core of category learning (Gentner & Namy, 1999; Namy & Gentner, 2002), some studies find that emphasizing differences promotes category learning (Mather & Plunkett, 2011; Posner & Keele, 1968).

According to the *structure-mapping theory*, comparison serves to highlight commonalities and leads to the process of structural alignment, whereby two representations are aligned and common structures are preferentially highlighted (Gentner, 1983). The process of comparison motivates children to attend to and search for commonalities between compared items.

At the same time, there is evidence that increased variability of category exemplars during familiarization promotes category learning even in young infants (Mather & Plunkett, 2011; Oakes & Ribar, 2005). These findings are in line with earlier classic studies on abstract categorization in adults revealing that higher variability among category exemplars leads to better categorization performance (Posner & Keele, 1968). Oakes and Ribar (2005) reported that infants show robust categorization effects when they are familiarized with pairs of different exemplars from the same category but not when familiarized with pairs of identical exemplars from the same category. Mather and Plunkett (2011) presented 10-month-old infants with the same stimuli taken from the Broad condition used by Younger (1985) in her study of infant categorization. They found that only infants familiarized with sequences that maximized variability between successive exemplars showed evidence of category learning, whereas infants presented with sequences that maximized similarity between consecutive exemplars failed to do so. The authors argued that maximizing differences between consecutive items results in highlighting and encoding contrasting regions of category space and leads to more robust category representations.

On the basis of these studies with infants and adults, we predict that presenting pairs of items which maximize dissimilarities between items should result in more robust category learning. Presenting pairs of items from different categories means that a larger region of category space is traversed in any given familiarization trial, as compared to the category space covered when presenting pairs of items from the same category. Furthermore, presenting items from different categories should highlight the differences between categories and facilitate identification of relevant categorical distinctions. Given that individual features in our study are not diagnostic, rather feature correlations define categories, we hypothesize that high variability between exemplars might be particularly relevant for discovering these categories. The present study tests this hypothesis by contrasting one condition in which infants are presented with pairs of items that come from the same category (maximizing similarities) and one in which infants are presented with pairs of items from different categories (maximizing dissimilarities). A control condition further establishes infants' performance in a single-item presentation paradigm. Differences in category formation under such conditions have the potential to provide a better understanding of how the level of similarity between the to-be-compared items impacts category learning.

1.3 | Overview of the study

To determine whether paired presentation facilitates novel visual category learning, we presented infants with one item at a time (*single condition*) or in pairs. To elucidate mechanisms underlying potential benefits of paired presentation, infants were presented either with pairs of items from the same category (*same-category pairs*) or with pairs of items from different categories (*cross-category pairs*).

To ensure that the effects were most likely to be a result of experimental manipulation, several methodological precautions were implemented. Infants in all three experimental conditions were familiarized with the same set of eight stimuli, therefore representing an identical category space. In addition, infants were tested using identical test items in all three conditions. It is noteworthy that there are limitations to interpreting any differences between single and paired presentation conditions, given that single-item presentation involves more trials than paired presentation. However, existing studies that explore the effect of comparison on infants' learning are subject to the same constraint due to the inherent characteristic of the single-presentation condition that involves additional memory decay (e.g., Oakes & Ribar, 2005). Critically, comparing the effects of the within-category and cross-category conditions has a potential to reveal deeper insights into the role of comparison in category learning.

Given the limitations in the memory abilities of young infants, we expected that the process of category learning when presented with pairs of items would be faster and lead to more robust category representation, as indexed by stronger looking preferences, in comparison with learning about the same items one item at a time. Based on the existing body of research demonstrating benefits of high variability in category learning, we expected that presenting pairs of items from different categories would promote category learning to a greater extent as compared to presenting pairs of items that belong to the same category. Existing studies show that the ability to form correlation-based categories is present at 10 months (e.g., Younger & Cohen, 1986), so the study was conducted with 10-month-old infants in order to enable comparison with the relevant body of research.

2 | METHOD

2.1 | Participants

Eighty-two participants took part in this study (M_{age} = 10.19 months, age range = 9.61–10.69 months, 39 girls).[1] Participants were recruited at the local maternity ward, and all were full-term babies with no known health conditions. All participants came from homes where English was the only language spoken. The present study was conducted according to guidelines laid down in the Declaration of Helsinki, with written informed consent obtained from a child's carer before any data collection. This study was approved by the Medical Sciences Ethics Committee at the University of Oxford. 29 infants were allocated to the *single-item* condition (M_{age} = 10.24 months, age range = 9.64–10.68 months, 14 girls), 27 infants to the *same-category pairs* condition (M_{age} = 10.20 months, age range = 9.51–10.69 months, 13 girls), and 26 infants to the *cross-category pairs* condition (M_{age} = 10.14 months, age range = 9.61–10.57 months, 12 girls). An additional five participants were tested, but not included in the analysis due to fussiness and refusal to look at the screen during the experimental session.

2.2 | Stimuli

A set of novel objects was designed for the purposes of this study. Colored and textured 3-dimensional-looking objects represented novel creatures (called Sukis). The set of Sukis was inspired by stimulus properties used in several categorization studies (Mather & Plunkett, 2011; Plunkett et al., 2008; Younger, 1993; Younger & Cohen, 1986). As illustrated in Figure 1, each Suki consisted of four features: body, antennae, hands, and legs. Each feature could vary systematically on a scale of seven levels (number of antennae, hand size, body shape, length of legs, for full specification of the set see Appendix S1). For example, stimulus 2222 (see Figure 1) has two antennae (level 2 of the antenna feature), second smallest hands (level 2 of the hands feature), smaller rounded body (level 2 of the body feature), and medium length of legs (level 2 of the legs feature).

The key novelty introduced in the Suki set is that it allowed the construction of test stimuli with feature levels not seen during familiarization. Extending the feature range from five levels

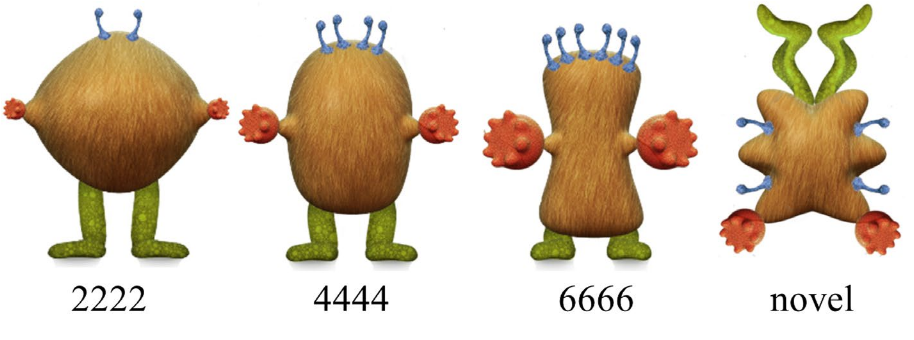

FIGURE 1 Examples from the Suki set: category prototypes (items 2222 and 6666), inter-category item (4444), and a novel out-of-category item

[1]Sample size was determined based on previous studies with the same age-group and using correlation-based categories.

of each feature used in the aforementioned studies, to seven feature levels, enabled us to create entirely novel test items.

For the familiarization set, eight Sukis were created such that features were correlated, that is, levels of one feature being predictive for levels on other features, thus inviting participants to form two categories. Our stimulus set thus loosely corresponds to the "narrow" condition in Younger (1985). To illustrate, Sukis with long legs always had small hands, whereas Sukis with a rounded body had fewer antennae (see Figure 2). In other words, all Sukis consisted of combinations of either feature values 1 and 3 *or* of combinations of values 5 and 7. Importantly, individual features were not prognostic, but correlations among the features provided the basis for category formation.[2]

Three additional Sukis were designed to be presented as test items: an inter-category item (consisting of overall mean levels on each feature, i.e., 4444) and two category prototypes (each consisting of average levels on each feature for its category, i.e., items with values 2222 and 6666 on 4 relevant features). In addition, a completely novel, out-of-category object which comprised the same features as other objects, but organized in a completely different manner, was presented in the final trial of the test phase in order to make sure that infants were in novelty preference mode, rather than familiarity preference mode (cf., Hunter & Ames, 1988). All objects were 450 × 450 pixels in size and depicted against a 5% gray background.

2.3 | Experimental design

2.3.1 | Familiarization

The study employed a standard familiarization-novelty preference procedure. First, infants were familiarized with a series of items from two novel visual categories. Infants were familiarized with either one item at a time or with pairs of items. Some studies present two identical items instead of a single item at a time (e.g., Oakes et al., 2009). We decided against this because that would result in doubling the familiarization time for the single condition, or presenting fewer exemplars to keep familiarization time equated. Presenting one item at a time enabled us to present an identical set of exemplars and keep the familiarization time equated across conditions.

In the *single-item condition*, there were 8 familiarization trials. Each trial started with a presentation of an animated star in the central location of the screen accompanied by a chiming sound for a duration of 2,000 ms. Following this, a single item was presented in a central location on the computer monitor for 6000 ms. As a previous study has shown that the order in which stimuli are presented may affect category formation (Mather & Plunkett, 2011), we calculated the mean Euclidean distance (as an average of seven distances between eight consecutive objects) for all possible stimuli sequences (40320 sequences) and selected sequences that fall between the 40th and 60th percentile (8112 sequences). Then, for each participant, a particular sequence from this pool was randomly selected.

[2]In the process of validating the stimulus set, we conducted a pilot study in which 10-month-old infants were presented with the same set of stimuli, but there was no correlation among the features. Despite all other parameters, apart from the lack of feature correlations being identical, infants did not show a preference for the inter-category item, suggesting that correlations among the features are used as the basis for category formation.

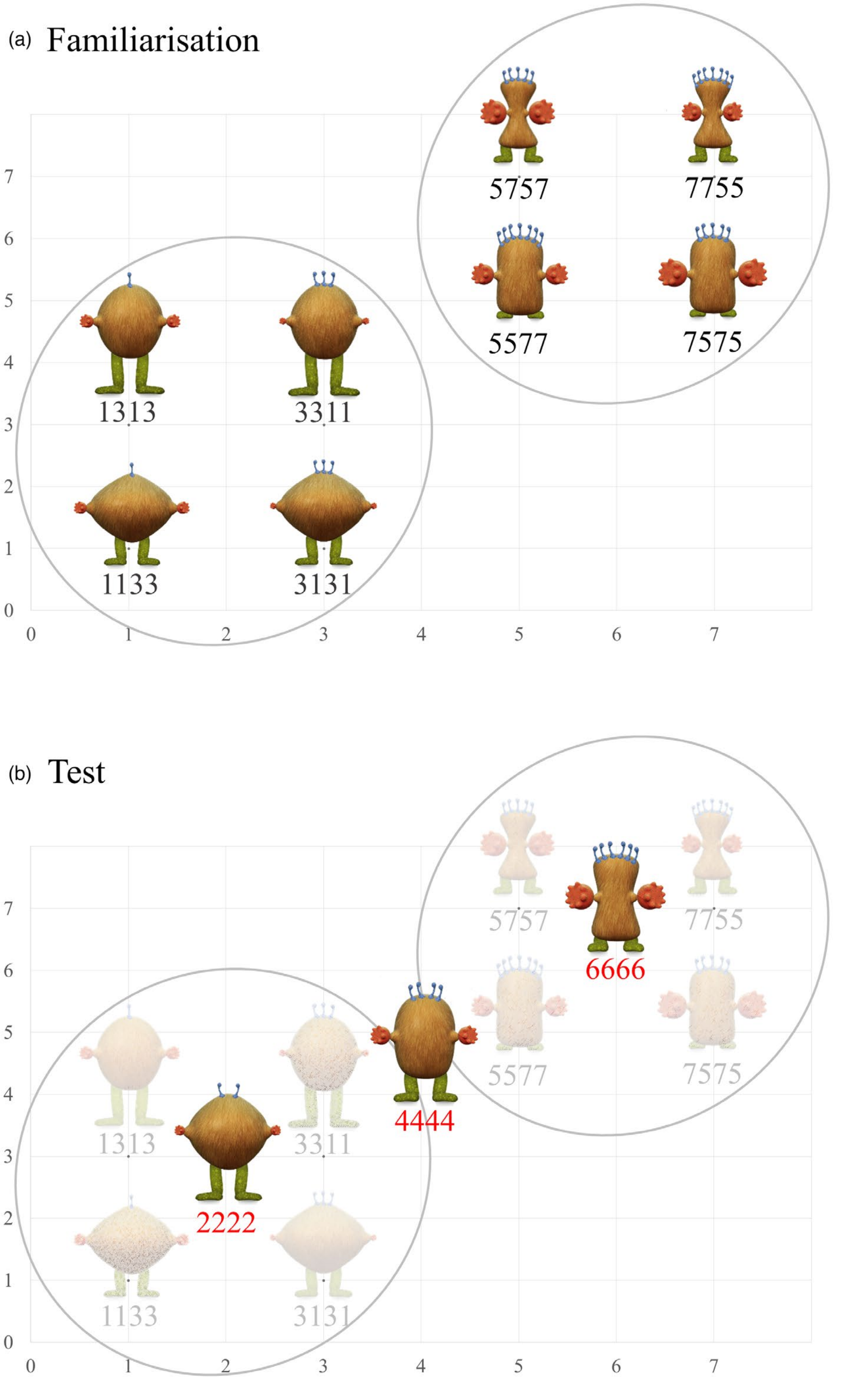

FIGURE 2 Schematic depiction of the category structure and stimuli examples: each stimulus can be described in a 4-dimensional perceptual space, where each dimension represents one feature, i.e., arms, hands, body, and legs); (a) familiarization phase and (b) test phase

In *the same-category and cross-category pairs conditions*, the same set of familiarization items was presented. A total of four trials were presented in each familiarization sequence, with items presented in pairs. Following the presentation of an attention getter, two objects were presented simultaneously for 12,000 ms. Familiarization time was increased from 6000 ms (*single-item condition*) to 12,000 ms in order to keep the total familiarization time (as well as the time available per object) identical across experimental conditions. Pairs of objects presented in familiarization trials were always either from the same category (*same-category pairs condition*) or from different categories (*cross-category pairs condition*). In the cross-category pairs condition, the positioning of items and categories was counterbalanced, that is, exemplars of both categories were equally likely to appear in each location, so that location could not be used as a cue for category membership.

The attention getter was presented for 2000 ms in all three experimental conditions. Although this led to longer overall presentation of the attention getter in the single-item condition as compared to the paired presentation, the duration of the attention getter during the inter-trial interval (ITI) was kept constant across all conditions, as the duration of the ITI has been found to have an important impact on infants' learning (Kovack-Lesh & Oakes, 2007). As a result of a fixed ITI, across all three conditions, the representation of familiarization items in working memory had to be retained for the same duration until they could be compared to the next familiarization trial. Shorter presentation of the attention getter in one condition would result in shorter ITIs, and consequently shorter retention periods and faster updating of the working memory representation. In their study of 10-month-old infants ability to perform the horses-dogs categorical distinction, Kovack-Lesh and Oakes (2007) found that reducing the duration of the inter-trial interval facilitated infants' performance. Therefore, although not ideal, we believe that the fixed ITI provided a more suitable test of the role of comparison in infant category learning. Nevertheless, presentation of category exemplars one at a time increased the number trials (and hence ITIs) needed for the infant to view all items.

It is important to note that despite the difference in the number of familiarization trials, the amount of time familiarization items were presented on the screen was identical across the three conditions. Therefore, any potential differences in the effect of the single and paired familiarization regimes cannot be due to different amounts of time infants were given to explore the objects. Yet, despite these efforts to keep the parameters identical across conditions, there are limitations when comparing single and paired presentation due to an additional memory decay stemming from the greater number of ITIs in in the single-item condition. However, this limitation stems from an inherent feature of presenting items one at a time that is difficult to avoid, also present in other studies that explored the impact of comparison on categorization outcomes.

2.3.2 | Test

After the familiarization trials, infants' category formation was evaluated in a test block. This set of test trials was identical across all three conditions.

During the first two test trials, dubbed "category formation test trials," infants were presented with two novel items: a prototype of one of the familiarized categories (consisting of category-specific average values in each feature, i.e., 2222 or 6666) and an item that consisted of average features of both categories (the inter-category item, with values 4444). Infants' looking preference was used as an index of category formation. The logic of these trials is this: If infants perceived the category contrast and formed two distinct categories, it is expected they will spend more time looking at the inter-category item (4444) because that item should appear, compared to the two category prototypes, less familiar and more surprising (cf., Younger, 1985). By contrast, if infants failed to perceive the category contrast and instead grouped all items together in a single large category, we would expect them to perceive this the item with the values 4444 as highly familiar (because it corresponds to the single category's centroid) whereas the items with values 2222 or 6666 would be less familiar by comparison (as they are further away from the category centroid).

Each trial began with an attention getter that was presented for 2000 ms. Then, two test objects were presented simultaneously for 10,000 ms (timing parameters were based on Plunkett et al., 2008). The positions of the two objects were counterbalanced across the two trials. Half the participants saw the item 4444 side by side with 2222, and half the participants saw 4444 with 6666.

The third test trial was always an out-of-category novelty preference test in which one of the training items from the familiarization phase was presented along with a novel, previously unseen out-of-category object. The purpose of this trial was to establish that infants were engaging in the task and expressing the expected novelty preference. A preference for the out-of-category item, which contains the same features but in a very different spatial configuration, also indicates that infants have encoded details of the spatial configuration across the familiarization items. Test trials were identical across all three experimental conditions, providing a legitimate evaluation of category learning across the three conditions.

2.4 | Experimental procedure and data pre-processing

After written consent was obtained from the carer, the infant was seated on their carer's lap approximately 60 cm from a 1920 × 1080 inch monitor screen in a sound-attenuated experimental booth. The carer was asked to keep their eyes closed for the duration of the experiment. Gaze data were recorded using a Tobii TX300 Eye Tracker with a 120 Hz sampling frequency. The study was run with PresentMate, a custom Matlab stimulus presentation software based on the Psychophysics Toolbox. The session started with a four point calibration procedure. Infants' behavior was monitored via a centrally located camera above the screen.

Fixations were calculated from raw gaze data using custom Matlab routines. A second-order Savitzky–Golay filter with a length of seven samples was used for data filtering (Nyström & Holmqvist, 2010). Minimum fixation duration was 100 ms. Eye-tracking data were considered valid if the eye-tracker validation flag indicated that at least one of the eyes was detected and the recorded gaze was within the screen area. Blinks were detected as sections of the data with an instantaneous rate of change of pupil diameter greater than 0.1 mm for both eyes and the corresponding samples were flagged as invalid. If the invalid data span was shorter than 100 ms, it was replaced with last valid value.

3 | RESULTS

3.1 | Visual exploration during familiarization

3.1.1 | Looking time during familiarization

Familiarization time was equated across the three conditions in order to guard against the possibility that the observed performance might be due to different levels of familiarization in the three conditions. As shown in Figure 3a, there was no systematic difference in the amount of looking time at the objects that infants accumulated during familiarization in the three experimental conditions. This was further confirmed in the statistical analysis which revealed a significant intercept term, showing that the total looking time was different from 0. Non-significant linear and quadratic terms show that there was no difference in the mean looking time across the three experimental conditions (the model estimates are reported in Table 1).

3.2 | Comparison propensity in the same- and cross-category pairs conditions

Familiarizing infants with pairs of items provides them with an opportunity to compare the two items, while both are visually accessible. We expect that the higher propensity for comparison might be reflected in more gaze shifts between the two simultaneously presented items. To explore the extent to which infants engage in comparison in the same-category and cross-category pairs conditions, we examine the rates of shifts between the two presented items during familiarization in the two experimental conditions.

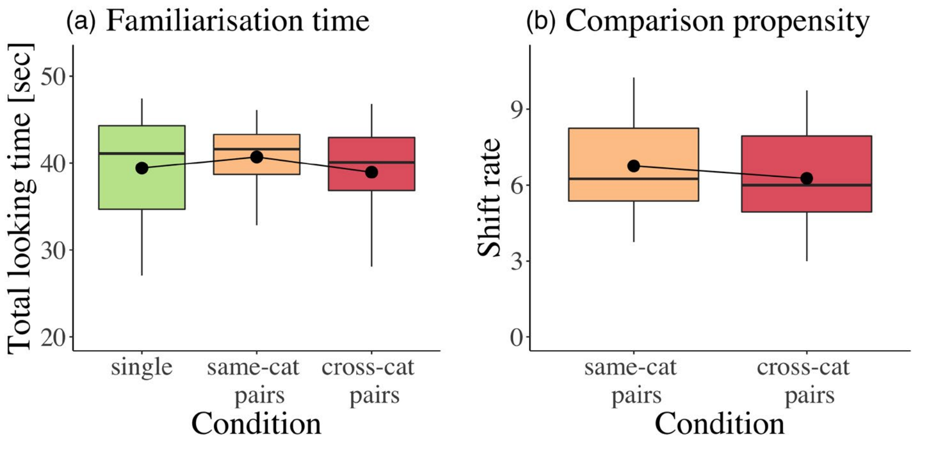

FIGURE 3 Visual exploration during familiarization: (a) total looking time accumulated during familiarization across three experimental conditions and (b) mean shift rates in same-category pairs and cross-category pairs conditions (black dots indicate means)

TABLE 1 Model estimates for the familiarization phase: looking time in familiarization was similar across the three experimental conditions

Familiarization looking time			
Predictors	**Estimates**	**CI**	***p***
Intercept	39.69	38.61 – 40.77	<.001
Linear	−0.34	−2.21 – 1.52	.721
Quadratic	−1.23	−3.11 – 0.64	.201
Observations	82		
R^2/R^2 adjusted	.022/−.003		

A shift was defined as a transition of looking from one familiarization object to another, that is, when looking to one object was followed by looking to another object rather than to another area of the screen or away from the screen. Each object was defined by a region of interest (ROI) consisting of a 450 × 450 pixels square surrounding the object. Mean number of shifts was calculated by averaging the number of shifts across the four familiarization trials for each infant.

As shown in Figure 3b, infants expressed similar shift rates, suggesting that infants engaged in comparison with a similar extent across the two conditions, $t(51) = 0.97$, $p = .33$, $d = .27$. This implies that any beneficial effects of cross-category comparison on category learning are not driven by different levels of shifting, but rather by the content of the to-be-compared pairs of items.

To further investigate familiarization looking patterns, we tested whether there was a drop in the amount of looking time in the initial and the final familiarization phase. Looking time in the first quarter of the familiarization was compared against the looking time in the last quarter of familiarization, that is, in the first and the last familiarization trial in the paired presentation conditions, and sum of looking times in the first two and the last two familiarization trials in the single-item condition. The mean looking time was submitted to an ANOVA with a within-subjects factor *Phase* and *Condition* as a between-subjects factor. The analysis revealed significant main effects of *Phase* ($F(2, 158) = 3.300$, $p = 0.039$, $\eta^2 = .039$) and *Condition* ($F(1, 158) = 4.529$, $p = .035$, $\eta^2 = .027$), while no *Phase x Condition* interaction was present (see Figure 4), suggesting that infants in all three experimental conditions show a familiarization effect.

To summarize, the analyses of the looking patterns during familiarization suggested that infants accumulate similar amounts of looking time across the experimental conditions and express a familiarization effect and that infants in the same-category and cross-category pairs conditions similarly engage in comparison with the two simultaneously presented items.

3.3 | Looking preferences at test

Our main hypothesis was that category learning will be most difficult for infants when familiarized with a single item at a time and that the optimal learning scenario involved presenting pairs of items from different categories, with same-category pairs representing an immediate level of difficulty. Therefore, we predicted that any novelty preference for the inter-category item would increase across the three experimental conditions. To directly test this hypothesis, we fit a linear model with *a priori* contrasts using *polynomial contrasts* (Schad et al., 2020) in R (R Core Team,

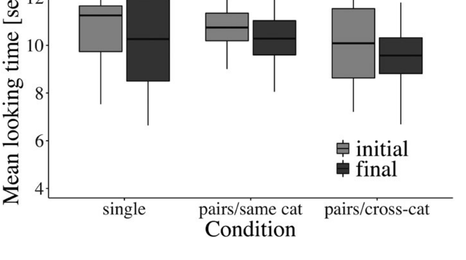

FIGURE 4 Mean looking time in the initial and the final familiarization phase across the three experimental conditions

2019), version 3.6.1. Polynomial contrasts are particularly suitable for testing the presence of a trend that spans across multiple conditions, especially when individual levels of a factor are not strong enough to be detected in the statistical model (Schad et al., 2020).

If the preference score increased by a similar magnitude from the *single-item condition* to the s*ame-category pairs condition* as it increased from the *same-category pairs condition* to the *cross-category pairs condition*, we expect to see a significant linear trend. Alternatively, we expect a significant quadratic trend if both paired conditions were equally effective, that is, there was no systematic difference between the same-category pairs condition and the cross-category pairs condition, but both paired conditions were better than the single-item condition. Finally, if there was no difference between the three conditions, we expect neither a linear nor quadratic term would be significant.

In addition, to determine whether there was any evidence of learning in each experimental condition, we ran a linear model with *cell means parametrization*. This enabled us to test whether the mean looking preference in each condition was different from chance, without testing for the differences between conditions.

3.4 | Category formation test

Preference scores were calculated by dividing looking at the inter-category item by the total looking time at both objects, averaged across the two test trials. The score was centered to 0 by subtracting 0.5 (performance if at chance). The analysis revealed a significant intercept and a significant linear trend across the 3 experimental conditions. The significant intercept confirmed that the overall preference score was different from 0, suggesting that infants showed evidence of category learning. The significant positive linear term revealed that there was a positive trend in the mean preference score, suggesting an improvement in the performance from the *single-item*, to *same-category pairs* to *cross-category pairs* conditions, see Figure 5a and Table 2 (upper section).

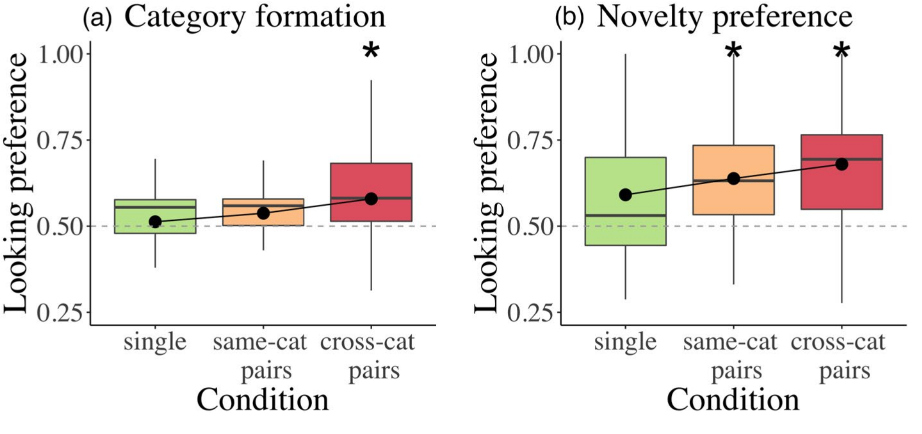

FIGURE 5 Looking preferences in (a) category formation test (higher preference indicates preference for the inter-category item, suggesting infants formed two categories); (b) novelty preference test (higher preference value indicates novelty preference, black asterisk depicts $p < .05$, and gray dashed line depicts expected preference score if performance was at chance (0.5)

TABLE 2 Model estimates for *category formation* test. Testing for a trend in performance across experimental conditions (upper section) and assessing whether performance differs from chance within each condition (lower section): Infants' preference differed from chance, and there was a positive linear trend across conditions

Category formation						
Predictors	**Estimates**	**CI**	***p***	**Estimates**	**CI**	***p***
Intercept	0.04	0.01 – 0.07	**.007**			
Linear	0.06	0.01 – 0.10	**.025**			
Quadratic	0.00	−0.05 – 0.05	.938			
Single				0.00	−0.05 – 0.05	.984
Same-cat pairs				0.04	−0.01 – 0.09	.130
Cross-cat pairs				0.08	0.03 – 0.13	.002

In the next step, we tested whether the mean looking preference in each condition is different from chance, without testing for the differences between conditions. The results revealed that the mean preference was significantly different from chance only in the *cross-category pairs* condition, whereas mean preference scores in the *single-item* and the *same-category pairs* conditions did not significantly deviate from chance, see Table 2 (lower section). This suggests that only infants in the *cross-category pairs* condition showed robust evidence of separating the stimulus set into two categories.

3.5 | Out-of-category novelty preference test

The final test served to validate whether infants' looking preferences in test trials were driven by novelty preference, but also to establish, especially in the absence of a preference in the category formation test, whether infants had encoded at least some information about the objects' features.

We fitted models the same way as reported above for the category formation test. The polynomial contrast testing for a trend in novelty preference across the three experimental conditions revealed a significant intercept term, while no effects for the linear and the quadratic terms, Figure 5b and Table 3 (upper section). The significant intercept reveals that the overall preference score was different from 0, that is, there was a significant novelty preference. The lack of linear and quadratic terms suggests there is no systematic difference in the out-of-category novelty preference test between conditions.

Testing the mean looking preference in each condition against chance, without testing for the differences between conditions, revealed that the mean novelty preference differed from chance in the *cross-category* pairs and same-category pairs conditions, while there was no novelty preference in the single-item condition, see Table 3 (lower section).

To summarize, the analyses of looking preferences in the category formation test showed that infants' performance improved across conditions: There was no evidence of category learning in the single-item condition, improved performance when familiarized with pairs of items from the same category, and finally robust category learning when familiarized with cross-category pairs. While there was no overall difference in preference for the out-of-category item across the experimental conditions, significant novelty preference was present in both the same-category

TABLE 3 Model estimates for *out-of-category novelty preference* test. Testing for a trend in performance across experimental conditions (upper section) and assessing whether performance differs from chance within each condition (lower section): Infants' preference differed from chance

Novelty preference						
Predictors	**Estimates**	**CI**	***p***	**Estimates**	**CI**	***p***
Intercept	0.10	0.05 – 0.15	<.001			
Linear	0.03	−0.05 – 0.12	.431			
Quadratic	−0.03	−0.11 – 0.06	.553			
Single				0.06	−0.02 – 0.14	.151
Same-cat pairs				0.12	0.03 – 0.20	.009
Cross-cat pairs				0.11	0.02 – 0.20	.016
Observations	82			82		
R^2/R^2 adjusted	.013/−.012			.163/.131		

and cross-category pairs conditions. This suggests that infants' looking at test was driven by the novelty preference. It also implies that infants in the paired conditions, even those who saw same-category pairs, represent some aspects of the spatial configuration of the stimuli.

3.6 | The impact of ITI on learning

In order to ensure that the representation of familiarization items in working memory had to be retained for the same duration until it could be compared to the next familiarization trial in all conditions, duration of the ITI was fixed across all three conditions. As a result, the total amount of time the attention was presented in the single-item condition was greater than in the paired item condition. To test whether the amount of looking at the attention getter interfered with learning, we examined whether there was a correlation between the looking time at the attention getter and looking preference at test. The analysis revealed no correlation between the mean looking time at the attention getter and looking preferences at test in any of the experimental conditions (*single-item*: $r = .11$, $p = .58$; *same-category pairs*: $r = .26$, $p = .18$; *cross-category pairs*: $r = .025$, $p = .9$). This suggests that the amount of time infants spent fixating the attention getter was not systematically related to their performance at test.

Lastly, to check for potential ceiling effects in looking times at test scores, we calculated mean looking times in the category formation and novelty preference test. As shown in Table 4, there is no evidence of ceiling effects in any of the conditions (trial duration was 10 s).

3.7 | Statistical approach: Polynomial contrasts vs. ANOVA

To test our main hypothesis that learning improves across the three learning scenarios, we adopted a statistical approach based on polynomial contrasts. Schad et al. (2020) highlight that this approach can yield more useful confirmatory tests of hypotheses than standard omnibus F-tests. Indeed, the advantages of this approach become apparent when we compare the outputs of the present analysis to the outputs of a traditional statistical approach using ANOVA, followed by planned comparisons (see Appendix S1). While a corresponding pattern of results

TABLE 4 Mean looking time (s) in category formation and novelty preference test trials

Condition	Category formation test	Novelty preference test
Single	7.03 (2.13)	7.76 (1.66)
Pairs: within-category	6.42 (1.46)	6.70 (1.95)
Pairs: cross-category	6.21 (1.77)	6.10 (2.07)

Standard deviation in parentheses.

was observed, the ANOVA failed to detect a significant effect of *Condition* ($p = .079$). This result would suggest that there is no systematic difference in performance across conditions, that is, that learning is similarly efficient in the single and pairs conditions. However, post hoc tests clearly show that learning does differ across conditions and that drawing a conclusion that the three learning scenarios are equally effective might not be justified. In addition, a Bayes factor of 0.81 for the main effect of *Condition* in the ANOVA does not provide sufficient support for accepting the null hypothesis, that is, for concluding that there is no difference across conditions.

Lastly, the observed effect size corresponds to the range typically observed in studies on novel category learning using the same experimental paradigm (e.g., d ranging from .47 to .61 in Plunkett et al., 2008). These effects are reliable, but small, and an ANOVA might not be sufficiently sensitive to capture these effects in the current dataset, while specifying a priori contrasts has more statistical power to detect these effects (Schad et al., 2020). Hence, in the present study, relying on a traditional statistical approach would have resulted in concluding that the three investigated learning scenarios are equally (in)efficient, whereas the present analysis clearly demonstrated that such a conclusion would not be justified.

4 | DISCUSSION

This study set out to explore the role of comparison in infant novel category learning. We familiarized infants with exemplars from two novel categories either one-by-one, with pairs of items from the same category, or with pairs of items from different categories. Infants' learning was directly related to the level of opportunity for cross-category comparison: There was no evidence of category learning when infants were familiarized with one item at a time. Performance improved when familiarized with pairs of items from the same category: While the preference score in the category formation test did not significantly deviate from chance, the fact that infants in the same-category pairs condition show a preference for the out-of-category item suggests that they had learnt something about the category exemplars during familiarization. They appear to have recognized the general spatial arrangement of object parts in the familiarization stimuli and recognize that the out-of-category stimulus is different. However, they demonstrated no evidence of extracting the correlation structure present within the familiarization set, which underpins the two category clusters.

Finally, when familiarized with pairs of items from different categories, infants show robust evidence of category learning at test, demonstrating a systematic preference for the inter-category item, as well as for the out-of-category object. Clearly, infants in this condition have represented the familiarization categories at a sufficient level of detail to recognize not only the overall spatial configuration of features (out-of-category test) but also the pattern of feature correlations across exemplars, that is, to recognize that there are two categories rather than just one (category formation test). The opportunity to compare items from different categories directly promotes infants' attention to details that are relevant for category discrimination.

It is important to note that the differences between single and paired presentation conditions were not driven by differing amounts of familiarization preceding test. Infants in all three experimental conditions were given the same opportunity to visually explore items. Therefore, the observed differences cannot be due to differing *amounts* of familiarization experience. Despite our efforts to minimize the differences between single and paired presentation, there are inherent differences between these conditions that might have contributed to poorer performance in the single-presentation condition. The single-item condition poses higher memory demands, as infants need to maintain a memory representation of an item until the next item is presented. In addition, single-item presentation results in more familiarization trials as compared to presenting items in pairs. As a result, there are more ITIs in the single-item condition than in the paired condition, which additionally increases memory demands.

In contrast to the existing body of research that showed that 10-month-old infants can form categories when presented with one item at a time (Mather & Plunkett, 2011; Plunkett et al., 2008; Younger, 1985; Younger & Cohen, 1986), in the present study we found no evidence of category learning in the single-item condition. As the experimental procedure used in this study was based on the one reported in Plunkett et al. (2008), the observed differences between these two studies are likely due to the increased complexity of the stimulus set used in the present study. In particular, the increased perceptual range of the stimulus features could be one of the factors leading to the observed results. The feature range was extended from five levels of each feature (the range used in previous studies) to seven feature levels. This increase in the perceptual range of stimulus features provided an opportunity to create test items that were entirely new as none of the feature levels at test were previously shown during familiarization. In addition, whereas the stimuli used in the previous studies were simple line drawings, the stimuli used in the present study were designed as more realistic, 3-dimensional-looking, textured items which further increased visual complexity. Thus, the combination of increasing feature variations and visual complexity may have hindered infants' ability to learn these categories when exposed to category exemplars one at a time.

A classic model of infant habituation postulates that the three key factors driving infants' visual preference are age, familiarity, and complexity (Houston-Price & Nakai, 2004; Hunter & Ames, 1988), where increased familiarity leads to novelty preference. Hence, in light of this model, the results of the present study suggest that the increased complexity of the categories might require further extended familiarization for the infant to succeed in forming categories when items are presented one at a time. One possibility consistent with this view is that infants struggle to encode each of these complex stimuli as they are encountered within a trial to a level of detail that is sufficient to allow the subsequent item to be mentally compared. As a consequence, it may be difficult for infants to extract regularities underlying the category distributions, in particular how feature values co-vary in this set. At test, the inter-category item therefore does not stand out as different, and infants do not exhibit a preference for either item.

Infants across all experimental conditions expressed a familiarization effect, as indexed by a significant drop in the amounts of looking time from the initial to the final stage of familiarization part of the experiment. At the same time, infants were in novelty preference mode across conditions at test. Therefore, after familiarization during which they became sufficiently familiar with the stimuli, infants expressed a novelty preference in the novelty preference test. Hence, the preference for the inter-category item in the paired presentation condition indicates that infants have formed two categories. In contrast, the lack of systematic preference for the inter-category item at test in the single-item condition, coupled with the familiarization effect and novelty preference effect, suggests that infants in the single-item condition acquired some knowledge and understanding of the novel categories, but not to a level sufficient to split items into two categories. Therefore, it seems that infants in the single-item condition were at an earlier stage of their learning trajectory when tested for category formation, as compared to infants presented with items in pairs.[3]

Presenting multiple items simultaneously can solve the problem of insufficient level of encoding details by allowing direct comparison and eliminating the need to rely on memory, thereby facilitating category formation. The present set of results supports this claim, as infants successfully learned these categories when familiarized with items in pairs.

The finding that infants failed to form categories when familiarized with one item at a time, but succeeded when the very same items were presented in pairs, is in line with previous studies demonstrating positive effects of paired presentation on learning categories in younger infants (Oakes et al., 2009; Oakes & Ribar, 2005). While these studies focused on existing categories, such as different animal classes, our results demonstrate that the opportunity to directly compare category instances facilitates category formation even when these categories are entirely novel. Therefore, we show that comparison plays a critical role not only in the reactivation of existing categories, but also in the formation of a new categories.

What makes presenting items in pairs beneficial for learning novel categories? Presenting items in pairs provides more opportunity to compare and contrast available information. Since visual contrasts are constantly available in the paired conditions, infants can immediately examine similarities and differences between the two items. Furthermore, infants are at liberty to decide for themselves how to allocate attention to the paired visual stimuli. In contrast, in the single-item familiarization regime, direct comparisons are not possible. Infants must compare one visually presented item with a working memory representation of the previously presented items. Inevitably, this leads to comparisons of an immediate perceptual representation with a more abstract memory representation. In particular, during stimulus encoding, relevant features must be selected. Identifying such diagnostic elements is possible immediately if contrasts are directly observable, for example, if a rounded "body" (e.g., item 3131) is observed next to a more rectangular one (e.g., item 7755), then "body shape" can be noted as a distinctive feature discriminating items, whereas the body's texture is not. For cross-category pairings, this should be particularly easy as there will be contrasting features, whereas same-category pairings by definition offer more similarities and therefore may not lend themselves to the identification of category-diagnostic features to the same degree. The results of this study thus suggest that direct lower-level perceptual comparisons are more useful for the processes of category formation in infants than the comparison of a perceived item to a representation in short-term memory. The results are therefore consistent with categorization models that assume comparison is one of the central factors in category and concept formation (Gentner & Namy, 1999; Kotovsky & Gentner, 1996; Kurtz et al., 2013). Although these models focused on categorization in children and adults, the present results speak for a similar effect even in early infancy.

[3]It is noteworthy that pilot work suggested that extended familiarisation led to novelty preference when infants were presented with one item at a time, suggesting that infants in the single-item condition were indeed at an earlier stage of learning as compared to infants in the pairs conditions.

More generally, the present set of results contributes to a more nuanced view of the role of comparison in infants' novel category learning. The results revealed that there is some advantage in learning from pairs of items from the same category as compared to the single-item condition. However, robust evidence of discovering a category boundary, that is, learning not just that there is a class of objects here, but also that it is subdivided into two clusters, is only obtained when infants are presented with pairs of items from different categories. The presentation of pairs of items crossing a category boundary serves to highlight contrasting information and promotes identification of those boundaries. Presenting pairs of items from different categories results in traversing a greater distance in category space within familiarization trials, with each gaze shift crossing the category boundary and covering longer distances in the category space as compared to same-category pairs. Emphasizing how items vary along different dimensions can thereby promote discovery of the category boundaries, which is critical for acquiring correlation-based categories, as they are defined by subtle contrasts as compared to other kinds of categories.

Given that the perceptual distance between pairs of items from different categories is greater than the perceptual distance between pairs of items from the same category, it is also possible that the cross-category advantage is primarily driven by perceptual distance, rather than category membership, per se. However, since most natural categories entail greater perceptual similarity between items from the same-category and smaller cross-category perceptual similarity, perceptual similarity and categorical effects are invariably intertwined. Category membership and perceptual similarity are certainly confounded in the current set of familiarization stimuli. The greatest same-category distance equates to the smallest cross-category distance.[4] Thus, a more careful interpretation of the present study is that increasing the perceptual distance enhances formation of novel categories. Cross-category pairing necessarily enhances perceptual distance between items in the current stimulus set. Further research is needed to disentangle the relative contributions of category membership and perceptual distance in paired familiarization regimes. Nor can we be sure that the impact of cross-category pairing holds for category structures that are not based on feature correlations. It is conceivable that feature correlations across category exemplars trigger the kind of comparison effects observed in this study. Non-correlational categories may trigger different patterns of comparison.

Our results also revealed that infants express similar shift rates in same-category and cross-category pairs condition, suggesting that they engaged in comparison with a similar extent. This suggests that it is not the overt process of comparison per se, but rather the content of comparison, that is, what is being compared, that can influence category learning.

Despite efforts to keep parameters in *single-item* and *paired presentation* conditions as similar as possible, there are certain limitations in evaluating the two familiarization regimes that need to be acknowledged. As a result of presenting items in pairs, infants saw more items earlier in the paired presentation condition as compared to infants in the single-presentation regime. They were exposed to a *broader range* of category exemplars earlier. Another difference between the conditions is that there was a change in the format between familiarization and test in the single-presentation regime (from presenting one item in familiarization to presenting two items in test). This was not the case in the paired presentation conditions, where the transition between familiarization and test was not apparent.

[4]Assuming equal salience between feature levels for all feature dimensions.

Keeping the duration of the ITI across all three conditions fixed meant that the representation of familiarization items in working memory had to be retained for the same duration until it could be compared to the next familiarization trial in all conditions. As there were twice as many trials in the single-item as compared to the paired presentation conditions, this meant that the total amount of time the attention getter was presented during familiarization was longer in the single-item as compared to the paired presentation conditions. This raises a possibility that looking at the attention getter during the ITI might have interfered with learning. However, our data provide no support for this possibility, as there was no correlation between the looking times at the attention getter during familiarization and performance at test. Nevertheless, as there were seven ITIs in the single-item condition, more than twice as many as compared to the paired presentation, it is not possible to completely rule out the possibility that this contributed to the poorer performance in the single-item condition: A *critical* memory decay may occur when the visual information about the categories is not available, that is, during the ITI. An infant must maintain the item representation in their working memory during this time. Keeping the duration of the ITI identical in all experimental conditions helped to ensure that the opportunity for a decay to occur between trials was identical across conditions. However, this control does not rule out the possibility of increased memory decay resulting from the increased number of attention getters in the single-item condition.

Infants' failure to learn categories in the single-presentation condition points to the importance of comparison in category learning. However, due to the different amount of ITI accumulated, this does not prove the importance of comparison. In addition, comparison occurs in the single-item condition as well. The critical evidence for the beneficial role of comparison in learning comes from comparing infants' performance in the same-category and cross-category conditions. Better performance in the cross-category conditions clearly demonstrates that it is the *type of comparison* that is important for infants' learning.

Despite these differences, by the time the test was presented, infants in all conditions had seen exactly the same items and were presented with the familiarization items for the same amount of time. Again, further research is needed to evaluate these other potential contributing factors (such as the duration of the ITI and duration of familiarization) to novel category learning in young infants.

The present study focused on 10-month-old infants' ability to form correlation-based, that is, statistical categories. In their study on the development of correlation-based categorization in infancy, Younger and Cohen's (1986) discovered that infants can detect isolated features at 4 months, they can detect of feature correlations of single items at 7 months, while feature-correlation-based categorization is present at 10 months (Younger & Cohen, 1986). Therefore, it would be interesting to explore whether the beneficial effects of comparison, and specific effects of different kinds of comparisons, are present in older infants as well. One possibility is that comparison facilitates learning because it reduces the requirements imposed on the limited working memory capacities of young infants. If this is the case, comparison might not be as beneficial for learning later in development. Alternatively, it is possible that this is a general learning principle and that similar effects would be observed with different age groups, although this effect is likely to interact with factors such as processing speed and working memory capacity. In line with this idea, some studies suggest that comparison plays an important role in category learning in children and adults (e.g., Hammer et al., 2009), indicating that comparison supports identification of category boundaries, above and beyond overcoming limitations in processing capacities in young infants.

In sum, the present study demonstrates that comparison promotes novel category learning in 10-month-old infants. What stands out is that a comparison which promotes contrasting information seems to be particularly helpful in the process of extracting category-relevant information. Comparing items from different categories results in highlighting the category boundaries during learning and in turn results in forming more robust category representations.

ACKNOWLEDGEMENTS

We would like to thank Irina Lepadatu for her help with data collection, and infants and their families for taking part in this study. The authors declare no conflicts of interest with regard to the funding source for this study.

ORCID

Jelena Sučević https://orcid.org/0000-0001-5091-5434
Kim Plunkett https://orcid.org/0000-0003-0216-7480

REFERENCES

Eimas, P. D., & Quinn, P. C. (1994). Studies on the formation of perceptually based basic-level categories in young infants based basic-level categories in young infants. *Child Development*, *65*(3), 903–917. https://doi.org/10.2307/1131427

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*(2), 155–170. https://doi.org/10.1016/S0364-0213(83)80009-3

Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development*, *14*(4), 487–513. https://doi.org/10.1016/S0885-2014(99)00016-7

Hammer, R., Diesendruck, G., Weinshall, D., & Hochstein, S. (2009). The development of category learning strategies: What makes the difference? *Cognition*, *112*, 105–119. https://doi.org/10.1016/j.cognition.2009.03.012

Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development*, *13*(4), 341–348. https://doi.org/10.1002/icd.364

Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research*, *5*, 69–95.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, *67*, 2797–2822. https://doi.org/10.2307/1131753

Kovack-Lesh, K. A., Horst, J. S., & Oakes, L. M. (2008). The cat is out of the bag: The joint influence of previous experience and looking behavior on infant categorization. *Infancy*, *13*(4), 285–307. https://doi.org/10.1080/15250000802189428

Kovack-Lesh, K. A., McMurray, B., & Oakes, L. M. (2014). Four-month-old infants' visual investigation of cats and dogs: Relations with pet experience and attentional strategy. *Developmental Psychology*, *50*(2), 402–413. https://doi.org/10.1037/a0033195

Kovack-Lesh, K. A., & Oakes, L. M. (2007). Hold your horses: How exposure to different items influences infant categorization. *Journal of Experimental Child Psychology*, *98*(2), 69–93. https://doi.org/10.1016/j.jecp.2007.05.001

Kurtz, K. J., Boukrina, O., & Gentner, D. (2013). Comparison promotes learning and transfer of relational categories. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *39*(4), 1303–1310. https://doi.org/10.1037/a0031847

Kwon, M. K., Luck, S. J., & Oakes, L. M. (2014). Visual short-term memory for complex objects in 6- and 8-month-old infants. *Child Development*, *85*(2), 564–577. https://doi.org/10.1111/cdev.12161

Mareschal, D., & Quinn, P. C. (2001). Categorization in infancy. *Trends in Cognitive Sciences*, *5*(10), 443–450. https://doi.org/10.1016/S1364-6613(00)01752-6

Mather, E., & Plunkett, K. (2011). Same items, different order: Effects of temporal variability on infant categorization. *Cognition*, *119*(3), 438–447. https://doi.org/10.1016/j.cognition.2011.02.008

Namy, L. L., & Gentner, D. (2002). Making a silk purse out of two sow's ears: Young children's use of comparison in category learning. *Journal of Experimental Psychology: General*, *131*(1), 5–15. https://doi.org/10.1037/0096-3445.131.1.5

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior Research Methods*, *42*(1), 188–204. https://doi.org/10.3758/BRM.42.1.188

Oakes, L. M. (2012). Advances in eye tracking in infancy research. *Infancy*, *17*(1), 1–8. https://doi.org/10.1111/j.1532-7078.2011.00101.x

Oakes, L. M., & Kovack-Lesh, K. A. (2007). Memory processes and categorization in infancy. *Cogniție, Creier, Comportament / Cognition, Brain, Behavior*, *XI*(4), 661–677.

Oakes, L. M., Kovack-Lesh, K. A., & Horst, J. S. (2009). Two are better than one: Comparison influences infants' visual recognition memory. *Journal of Experimental Child Psychology*, *104*(1), 124–131. https://doi.org/10.1016/j.jecp.2008.09.001

Oakes, L. M., & Ribar, R. J. (2005). A comparison of infants' categorization in paired and successive presentation familiarization tasks. *Infancy*, *7*(1), 85–98. https://doi.org/10.1207/s15327078in0701

Plunkett, K., Hu, J. F., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. *Cognition*, *106*(2), 665–681. https://doi.org/10.1016/j.cognition.2007.04.003

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*(3), 353–363. https://doi.org/10.1037/0097-7403.29.3.c2

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*, 382–439. https://doi.org/10.1016/0010-0285(76)90013-X

Ross-Sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The development of visual short-term memory capacity in infants. *Child Development*, *74*(6), 1807–1822. https://doi.org/10.1046/j.1467-8624.2003.00639.x

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, *110*, 104038. https://doi.org/10.1016/j.jml.2019.104038

Team, R. C. (2019). *R: A language and environment for statistical computing*. Team, R. C.

Younger, B. A. (1985). The segregation of items into categories by ten-month-old infants. *Child Development*, *56*(6), 1574–1583. https://doi.org/10.2307/1130476

Younger, B. A. (1993). Understanding category members as "the same sort of thing": Explicit categorization in ten-month infants. *Child Development*, *64*(1), 309–320. https://doi.org/10.1111/j.1467-8624.1993.tb02912.x

Younger, B. A., & Cohen, L. B. (1986). Developmental change in infants' perception of correlations among attributes. *Child Development*, *57*(3), 803–815. https://doi.org/10.2307/1130356

Younger, B. A., & Furrer, S. D. (2003). A comparison of visual familiarization and object-examining measures of categorization in 9-month-old infants. *Infancy*, *4*(3), 327–348. https://doi.org/10.1207/S15327078IN0403_02

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Sučević, J., Althaus, N., & Plunkett, K. (2022). Discovering category boundaries: The role of comparison in infants' novel category learning. *Infancy*, 27, 533–554. https://doi.org/10.1111/infa.12459