

# **The role of labels and motions in infant category learning**

Jelena Sučević<sup>1</sup>, Nadja Althaus<sup>2</sup>, Kim Plunkett<sup>1</sup>

<sup>1</sup>Department of Experimental Psychology, University of Oxford, Anna Watts

Building, Radcliffe Observatory Quarter, Oxford, OX2 6GG, UK

<sup>2</sup>School of Psychology, University of East Anglia, Norwich Research Park, Norwich,

NR4 7TJ, UK

The authors declare no conflict of interest.

## **Author Note**

Correspondence regarding this article should be addressed to Jelena Sucevic,

Department of Experimental Psychology, University of Oxford, Anna Watts Building,

Radcliffe Observatory Quarter, Oxford, OX2 6GG, UK

E-mail: jelena.sucevic@psy.ox.ac.uk

## Introduction

During visual object category formation, do infants use transient information such as the auditory labels and visual motions that may accompany objects, in addition to their inherent and permanent properties, such as shape and colour, to identify categories? And if true, how are these transient properties included in the representations of the object categories themselves? For a full mastery of these kinds of categories, an infant must encode both permanent features such as the shape and size of objects, but also their transient features, such as characteristic motions and labels. Given the transient nature of labels and motions, they are unlikely to provide input throughout the event when infants pay attention to the visual object. Nevertheless, both labels and motions are highly salient: labels are known to maintain infants' attention to the objects they accompany (Althaus & Westermann, 2016) and can even override visual information during category formation (Plunkett, Hu, & Cohen, 2008); motion is thought to represent one of the most salient features of objects (Wronski & Daum, 2014) and a preference for moving over static objects is present already in newborns and very young infants (Dannemiller, 2000; Slater, 1989).

To determine the impact of labels and motions on novel visual category learning, we developed a novel gaze-contingent eye-tracking task, which provides a more realistic learning scenario where motions and labels are triggered by infants' fixations upon specific objects. In a series of three experiments, we show that both category-specific auditory labels and category-specific visual motions support categorisation of novel visual objects in 10-month-olds. Moreover, we demonstrate that both labels and motions engender a similarity-driven pattern of exploration of category exemplars which is not apparent when the exemplars are presented in silence.

To form a category, an infant must come to understand that members of the same category can look very different, and yet they are still the same kind of a thing: despite

substantial differences in their appearance, the Mastiff, the Labrador and the Chihuahua all count as dogs. An infant may also notice that some categories are characterised by a certain type of motion: birds flap their wings, kangaroos hop, dogs wag their tails.

*The impact of labels on visual category learning*

A substantial body of research has explored the effects of auditory labels on infants' category formation, providing evidence that labels may profoundly shape categorisation (Althaus & Mareschal, 2014; Althaus & Plunkett, 2015; Fulkerson & Waxman, 2007; Robinson & Sloutsky, 2007a; Sloutsky & Robinson, 2008; Waxman & Braun, 2005; Waxman & Markow, 1995). In some cases, labels were shown to facilitate category learning (e.g. Fulkerson & Waxman, 2007), while labels were also shown to hinder categorisation (Deng & Sloutsky, 2015).

Waxman and colleagues investigated the impact of labels on infants' category learning in a series of studies (for an overview, see Waxman & Gelman, 2010). For instance, Fulkerson and Waxman (2007) familiarised infants with a set of objects accompanied by a novel label or a tone. In a subsequent test phase, infants showed evidence of category learning when familiarised with objects accompanied by a label, but not when accompanied by a tone. Furthermore, this effect was found only when the labels were presented consistently: mixed labelling of the familiarisation objects failed to yield evidence of categorisation (Waxman & Braun, 2005). Waxman and colleagues proposed that labels facilitate object categorisation and argued that labels act as an invitation to form categories. Several recent studies found evidence that a potential mechanism driving the effect of labels is by labels directing infants' attention to commonalities between items that share the same label (Althaus & Mareschal, 2014; Althaus & Plunkett, 2016). For instance, Althaus and Plunkett (2016) presented 12-month-old infants with a set of exemplars of a novel category, accompanied by a label or in silence. To elucidate whether different object processing

strategies were present in the two experimental conditions, the authors analysed infants' looking patterns. They discovered that in the label condition infants' attention was directed towards features shared by the familiarisation objects, while infants' exploration in the silent condition was driven by perceptual variability. Althaus and Plunkett (2016) argued that labels induce a persistent focus on commonalities, help maintain attention to objects during familiarisation, and thereby facilitate categorisation.

Moreover, several studies have provided evidence of the power of labels to realign category boundaries (Althaus & Westermann, 2016; Gliozzi, Mayor, Hu, & Plunkett, 2009; Plunkett et al., 2008). For instance, Plunkett and colleagues (2008) presented infants with a set of visual stimuli that could be grouped into one large category or split into two small categories. They found that infants formed two novel categories when exemplars were presented in silence and also when accompanied by two category-specific labels (Exps. 2 and 3). However, when these very same exemplars (from both categories) were accompanied by only one label, infants grouped all the items together into one category (Exp. 5). Labels can also produce the opposite effect: presenting infants with an ambiguous set of stimuli accompanied by two labels can result in dividing these items into two categories (Althaus & Westermann, 2016).

Other studies have revealed that labels can impede infants' ability to learn novel categories. In a series of studies, Sloutsky and colleagues tested infants' performance in a novel category learning task with category exemplars presented in silence or accompanied by labels (Sloutsky & Napolitano, 2003; Sloutsky & Robinson, 2008). They discovered that infants learned better when the category exemplars were presented in silence and that labels in fact interfered with learning. They proposed that limitations in processing capacities lead to infants' restricting attentional resources to one sensory domain. The authors interpret these results as an auditory dominance effect (Robinson & Sloutsky, 2004). Presenting visual

categories together with labels results in *auditory overshadowing*: auditory information is processed, but visual information is not processed to the same extent as when visual exemplars are presented in silence. As a result, category learning is impeded. An alternative view proposes that the auditory overshadowing stems from difficulties with the level of discriminability of the visual and auditory stimuli (Noles & Gelman, 2012).

Althaus and Plunkett (2015) offered a possible reconciliation of these contradictory findings on the impact of labels on visual category learning. They systematically manipulated the timing of the label onset in relation to the onset of the visual stimulus by presenting labels at the same time as the visual stimulus (synchronous labels) as Sloutsky and colleagues had done, or 1 second after the visual stimulus onset (asynchronous labels). Althaus and Plunkett (2015) found that synchronous labels hindered category learning, whereas asynchronous labels did not, revealing that the timing of the label is a critical parameter determining the impact of labels on novel category learning. The authors reasoned that asynchronous presentation allows infants to establish a representation of the visual information prior to processing the label itself.

In sum, labels represent highly salient and transient features that capture infants' attention during learning and may impact the process of visual category formation. Several studies have argued that labels achieve their effect by drawing infants' attention to the similarities between exemplars of the same category.

### *Encoding perceptual features and motion*

Motion represents another salient and transient feature which, in contrast to labels, is within the same sensory domain as static visual features. Sensitivity to the correlations between motion and static features is considered crucial for infants' developing knowledge

about the motion properties of a wide range of categories (Rakison, 2004). Some studies provide evidence that already by 7 months of age, infants are able to integrate information about the perceptual features of objects and motion (Pauen & Träuble, 2009) and to map a specific type of motion to an object (Markson & Spelke, 2006). Pauen and Träuble (2009) familiarised 7-month-old infants with a moving object and a static object (an animal or a ball). At test, infants were presented with the same two objects, but both objects remained static. Infants looked longer at the object that previously moved, suggesting they show an expectation for the object to move again. To investigate infants' ability to learn about self-propelled objects, Markson and Spelke (2006) presented 7-month-old infants with two moving toys, one being a wind-up self-propelled toy and the other toy being moved by an experimenter. Following familiarisation, when presented with both toys together stationary, infants expressed more interest in the self-propelled wind-up toy, suggesting that they learned the mapping between objects and their respective motions. Interestingly, both studies found that infants' learning was less reliable when objects lacked faces and body parts (e.g. removing eyes or fur from a bear in Pauen & Träuble, 2009) or when the objects were novel (Markson & Spelke, 2006). These studies suggest that from an early age, infants can rapidly learn motion properties associated with objects. Moreover, they highlight the importance of existing category knowledge in encoding motion and static information: motion properties were more robustly encoded for exemplars of familiar animate categories.

While these studies demonstrate that young infants readily learn about the motion properties of individual objects, they do not address how this might affect the process of organising objects into categories. Given how rapidly infants form these mappings, one might expect that motion should contribute to organising novel objects into categories.

A recent study by Deng and Sloutsky (2015) supported this idea and showed that motion can facilitate infants' category learning. They familiarised 8- to 12-month-old infants

with novel creatures from one of two categories each of which had a family-resemblance structure. Infants were familiarised with exemplars from a single category either in silence, accompanied by a specific type of motion (moving feet) or a label. To test for category formation, infants were presented with a prototype of the familiarised category and a novel prototype of another category (not presented during familiarisation). The results revealed that infants exposed to motion demonstrated a preference for the novel prototype, providing evidence of category learning, whereas those exposed to labels failed to demonstrate any novelty preference. Infants who were presented with the category exemplars in silence also failed to show any evidence of category formation. Infants' looking patterns suggested that the presence of motion encourages more distributed attention during learning, and the authors proposed that this in turn facilitates category learning (Deng & Sloutsky, 2015).

In contrast to the reported facilitatory effects of motion on category formation, some studies suggest that the integration of static and motion information is a difficult task for infants (Burnham, Vignes, & Ihlen, 1988; Rakison, 2004; Rakison & Poulin-Dubois, 2002). To examine infants' ability to memorise moving colour-shape compounds, Burnham and colleagues (1988) presented 4-, 7- and 10-months-old infants with differently coloured shapes, e.g. a red square and a yellow circle. The shapes were either stationary or moving (vertical or horizontal displacement of 20 degrees in each direction) during familiarisation. At test, infants were presented with a familiarised stimulus and a novel one (infants familiarised with moving shapes were presented with moving shapes at test, while infants familiarised with stationary shapes were tested with stationary shapes). Infants' looking preference for the novel shape was used as an index of successful memory, i.e. the ability to encode static features and a type of motion. They discovered that all age groups were successful in remembering static compounds, but not moving compounds. Similarly, Rakison and Poulin-Dubois (2002) explored infants' ability to detect correlations between motion trajectories and

perceptual features of objects. They presented infants with two objects (each consisting of a body and two arm-like parts), each characterised by a distinct motion trajectory. For example, infants were habituated to a red round object that moved up-and-down and a rectilinear blue object that moved in a curvilinear fashion. At test, infants were presented with a familiar compound stimulus and a novel compound stimulus that violated previous object-motion correlations, e.g. the red object moving up-and-down and a blue object also moving up-and-down. While 10-month-old infants did not show a systematic looking preference, eighteen-month-old infants expressed longer looking at the novel as compared to the familiar compound. This suggested that 10-month-old infants failed to detect violations of the correlations between static features and motion trajectories, while eighteen-month-old infants were able to detect these violations.

In a related study, Rakison (2004) investigated the age at which infants become able to detect these correlations between motion and static features in a category formation context. Infants were presented with a set of moving objects from two categories. Objects from each category were always presented with a category-specific motion trajectory. After infants habituated to the training set, they were presented with three test trials. In each test trial, one moving object was displayed: a congruent event (the motion-category combination was identical to the one presented during training), an incongruent event (violating the previously presented motion-category relationship), or a novel event (a novel object moving in a novel manner). The amount of looking time to the three test events was used as an index of infants' ability to encode correlations between static features and motion trajectory. Rakison (2004) found no difference in the looking time in 14-month-olds; 18-month-olds showed more interest in the novel event, but with no difference in the looking time to the congruent and the incongruent event; 22-month-olds looked longer at the incongruent and novel event as compared to the congruent event. Rakison (2004) concluded that infants show

no evidence of encoding any features at 14-months, but infants do show *some* ability to encode isolated features by 18 months of age. Only at 22 months of age, were infants able to generalise learned category-motion relations to novel instances.

Taken together, the existing research suggests that motion attracts infants' attention to objects and that infants learn mappings between objects and motions, but that they can struggle to encode static-motion bindings in novel objects.

These studies raise the question of the role of motion in category learning. Does motion just highlight attention to category exemplars, or does it become part of the defining characteristics of the category? Determining the impact of motion on acquiring a single category does not provide a complete answer to these questions. Instead, it is necessary to determine how motion impacts learning to differentiate two categories. If motion is contrastive, i.e. a category-defining feature, it could be used as a cue to help solve the categorisation problem and facilitate category learning. In contrast, if motion has a solely attention-arousing function, the process of learning two categories might not benefit from it. To address this issue, in the present study we examined how motion impacts infants' ability to differentiate highly similar objects into two categories.

### *Overview of the study*

To determine further the nature of the impact of label and motion features on the formation of visual object categories, labels and motions were introduced independently as category-specific features, in addition to the category exemplars themselves. They occurred in a gaze-contingent way, i.e. were triggered by fixations to the appropriate familiarisation stimuli. We ensured that the duration and timing of the category-specific labels and motions were identical. Each member of a category was associated with the same motion. Labels and motions bore an arbitrary relationship to the respective category. The labels and motions were

redundant insofar as the static visual information (item features such as size and shape) was, in principle, sufficient to extract the structure of these categories.

Experiment 1 explored the impact of labels on learning novel visual categories. Experiment 2 assessed the role of motion features in category formation. Finally, Experiment 3 explored the formation of these novel visual categories in the absence of any transient features, i.e. when only static visual information is presented.

In contrast to many categorisation experiments with infants, in this study we presented multiple familiarisation stimuli at the same time, simulating the more cluttered environment to which infants are typically exposed.

Another important innovation introduced in Experiments 1 and 2 is the use of a gaze-contingent familiarisation phase. Instead of using a pre-determined order of presentation of category exemplars and associated transient stimuli, motions and labels were triggered by infants' fixations upon specific objects. In a scenario with simultaneous presentation of multiple exemplars, a gaze-contingent design provides a way to present a label unambiguously for the relevant object, i.e. it enables a clear correspondence between say a labelling event and the fixation of an object. Moreover, the gaze-contingent design provides a more realistic scenario for learning: Infants play a more active role in designing their own learning experience. Within such gaze-contingent learning settings, infants quickly discover that they are in control, with consequent positive effects on their performance (Wang et al., 2012). The gaze-contingent approach also enables tracking of infants' patterns of fixation within the multiple, simultaneous presentation scenario.

If labels support the formation of categories and focus attention on the commonalities between objects, as suggested by previous research, labels might be particularly useful for discovering category-relevant information in a scenario with multiple items presented at the same time. Therefore, we expected that infants in Experiment 1 would show robust evidence

of category learning. The impact of motion on category learning is less well understood. If motion can also be used as a contrastive feature, the presence of motion might support category learning in Experiment 2. However, if motion serves only to maintain attention to objects, the use of motions as features to differentiate 2 categories may fail. In both Experiments 1 and 2, the presence of labels and motions inevitably increases the processing load on our infant participants. Given that familiarising infants with four items simultaneously imposes high processing demands, the presence of labels or motions may overload the system and impede the process of novel visual category learning. Experiment 3 reduces these processing demands by presenting the same objects in silence and without movement. To enable comparison with the relevant body of research (most studies were conducted with 8-12-month-old infants), the experiments were conducted with 10-month-old infants.

## EXPERIMENT 1

Experiment 1 investigated the impact of labels on visual category formation. Infants were presented with a set of exemplars from two novel categories, and each exemplar was accompanied by a category-specific label. Four exemplars were presented simultaneously in each familiarisation trial. To ensure a clear correspondence between an object and a label, we employed a gaze-contingent eye-tracking paradigm: infants' fixations triggered the presentation of a category-specific label.

Category learning involves a process of continuous information accrual throughout a sequence of familiarisation trials. A category representation is expected to emerge gradually across the experiment, and mature with increased exposure. Therefore, we probed category formation at different points in the familiarisation process. We introduced an experimental design with interleaved familiarisation and test blocks, which enabled us to test categorisation

on several occasions, each defined by increasing levels of familiarisation.

## Method

### Participants

Twenty-six 10-month-old infants took part in this experiment (12 girls,  $M_{age} = 10.22$ ,  $age\ range = 9.69 - 11.04$  months)<sup>1</sup>. Three additional participants were tested, but excluded from the analysis due to insufficient looking time accumulated in the first familiarisation block (accumulated looking time was more than 2 SD below the mean). All participants were full-term babies with no known health conditions, came from monolingual British-English families and were recruited at the local maternity ward. The study was approved by the University of Oxford Medical Sciences Ethics Committee. Of the 26 participants in the final sample, only 22 contributed to the data for the final testing block. Four participants were excluded from the analysis of the final testing block, due to low looking times ( $n = 2$ ) and not completing the experiment ( $n = 2$ ).

### Stimuli

A set of novel objects was designed for the purposes of this study. Coloured and textured 3-dimensional-looking objects represented novel creatures (called Sukis). The set of Sukis was inspired by stimulus properties used in several categorisation studies (Mather & Plunkett, 2011; Plunkett et al., 2008; Younger, 1993; Younger & Cohen, 1986). Younger's original stimulus set consisted of drawings of novel animal-like creatures. All creatures consisted of the same set of features (neck, legs, ears and a tail), which varied in size and could take one of five levels (for instance, legs varied in five levels of length, ears could take one of five levels of separation). Whereas individual features were not prognostic, correlations amongst the features provided the basis for category formation (e.g., long necks

---

<sup>1</sup> Sample size was determined based on previous studies with the same age-group and using similar category structure.

co-occurred with short legs). Infants at 10-months of age were able to encode these feature correlations and use them to form 2 distinct categories (Younger & Cohen, 1983).

The key novelty introduced in the Suki set is that it contained test stimuli with feature levels not seen during familiarisation. While the test items in Younger's stimulus set contained feature levels that were previously presented during familiarisation, by extending feature variability in the Suki set, the test items were completely novel.

As illustrated in Figure 1, each Suki consisted of four features: body shape, number of antennae, size of the hands and length of the legs. Each feature could vary incrementally along a scale of 7 levels. A set of 24 Sukis was created with levels of one feature being predictive for levels on other features, inviting participants to form two categories (for the full specification of the stimulus set, see Table A1). To illustrate, Sukis with long legs always had small hands, whereas Sukis with a rounded body had fewer antennae. 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, Figure 1b), and two category prototypes (each consisting of average levels on each feature for its category, i.e. items with values 2222 and 6666 on four relevant features). In addition, completely novel, out-of-category objects which comprised the same features as other objects, but organised in a completely different manner were presented in the final trial of each test phase (see Figure 2c). All objects were 450 x 450 pixels and depicted against a 5% grey background.

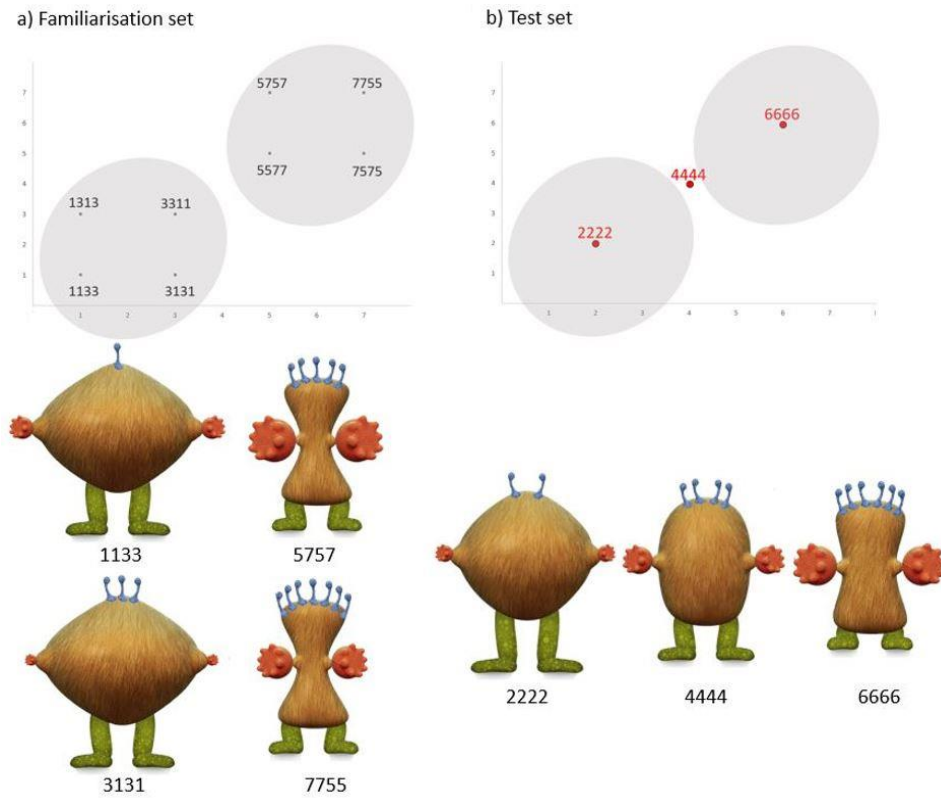


Figure 1. Schematic depiction of the category structure and stimuli examples of **(a)** a familiarisation set and **(b)** a test set: Each stimulus can be described in a 4-dimensional perceptual space where each dimension represents one feature, i.e. antenna, hands, body or legs (category 1 is defined by low values on features, and category 2 by high values on features).

## Experimental design

Familiarisation and test blocks were interleaved, to test category formation at different levels of familiarity. There were three familiarisation blocks, each followed by a test block (Figure 2a).

**Familiarisation.** Each of the familiarisation blocks consisted of four gaze-contingent familiarisation trials. Four items were presented simultaneously in each trial, two exemplars from each category. Each trial began with the presentation of a central animated star (an attention getter) that remained on screen until infants fixated it, thereby initiating the display of four items for a fixed duration of 16,000 ms.

Eight exemplars were presented in each familiarisation block (four exemplars from

each category, see Appendix for details). The same four exemplars were presented in the first and third familiarisation trial (and the remaining four in the second and fourth trial). The positioning of the items and categories on the screen was counterbalanced, i.e. exemplars of both categories were presented in various locations across trials; thus, location could not be used as a cue for category membership.

When infants accumulated 500 ms of looking at an exemplar, the exemplar was labelled with one of the two novel category-specific labels: "perto" and "dulver" (both labels were 400 ms long). If infants failed to accumulate 500 ms of looking, no label was played. After the label was played and the infant still maintained her fixation to the object, the label was not repeated. Only if the infant moved her fixation away from the object and then back to the object, was the label played again. There was no restriction to the number of times a fixation could trigger the label of the same object within a trial. Label-category associations were counterbalanced across participants, i.e. for half of the participants, one category was associated with the label "perto", and for the other half of participants was associated with the label "dulver". To emphasise the gaze-contingent nature of the task and provide a clear correspondence between the label and the object, the fixated object became more illuminated while the label was played.

**Test.** Immediately after each familiarisation block, infants were presented with a test block. Each test trial started with the presentation of an animated star for 2,000 ms, followed by a pair of test items that remained on screen for 6,000 ms.

Two trials were category formation test trials in which one of the prototypes of the two categories (items 2222 and 6666) was paired with an item that consisted of the average of the features of both categories (i.e., the inter-category item, 4444). The order of the trials was

balanced using a Latin square design<sup>2</sup>. The choice of the category prototype (2222 or 6666) presented first in each test block was balanced across participants. The third test block was identical to the first test block for half of the participants, whereas others saw identical items as in the second test block.

The final test trial was always a novelty preference test in which one of the familiarisation items from the previous familiarisation phase was presented along with a novel, out-of-category object.

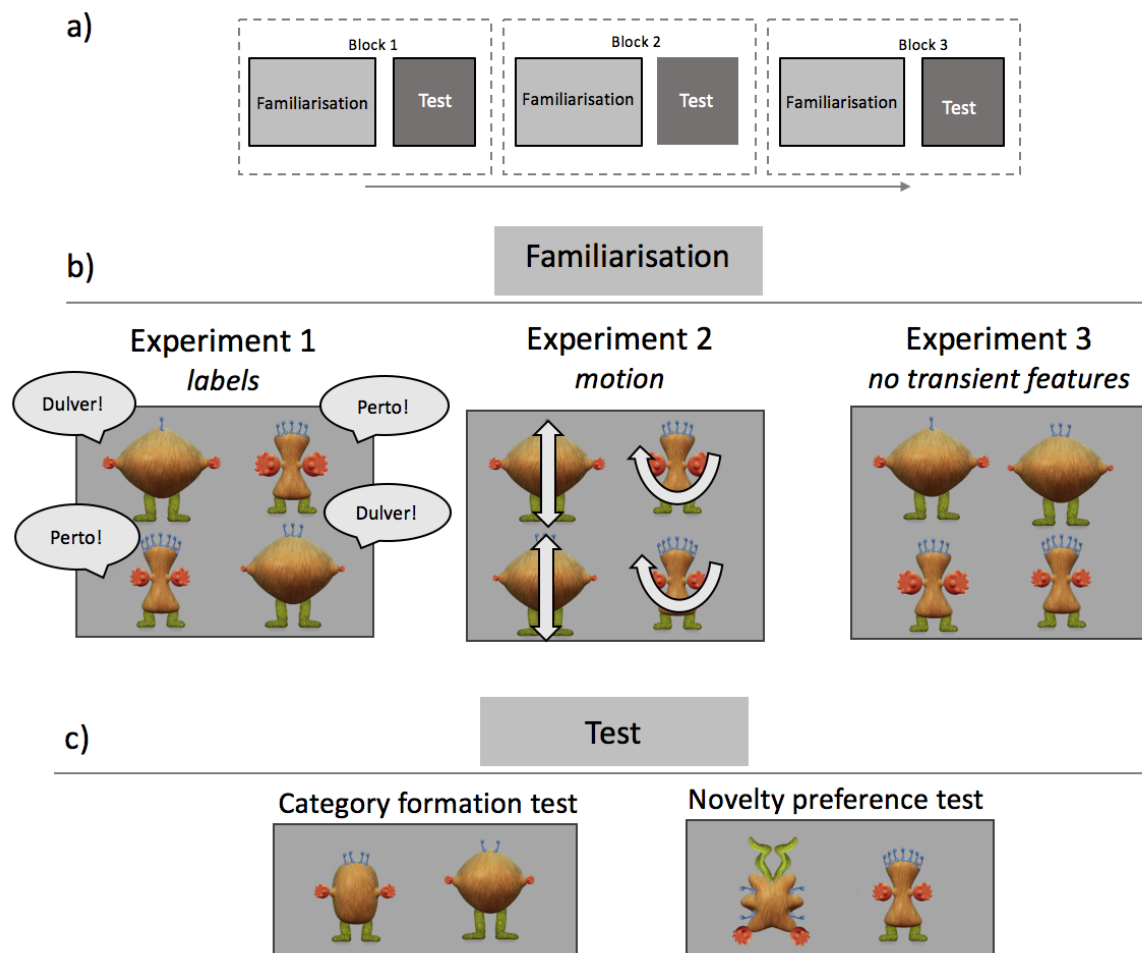


Figure 2. Schematic depiction of the experimental design (a). Characteristics of the familiarisation phase in 3 experiments: items were accompanied by category-specific labels in

<sup>2</sup> An additional test trial was presented. In this trial, the two category prototypes were presented side by side to test whether infants express a preference for either of the two categories. This trial was presented to enable comparison to another set of related experiments.

Experiment 1, accompanied by category-specific motions in Experiment 2, and presented in silence in Experiment 3 **(b)**. Illustrations of test trials **(c)**.

## Experimental procedure

After written consent was obtained from the carer, the infant was seated on their carer's lap approximately 60 cm from a 1920x1080 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 was recorded using a Tobii TX300 Eye Tracker (Tobii, Stockholm, Sweden) with a 120 Hz sampling frequency. The testing session started with a nine-point calibration procedure. The study was run with the custom Matlab stimulus presentation software PresentMate based on the Psychophysics Toolbox. In case infants became distracted, the experimenter could talk to them via a microphone between trials in order to attract infants' attention back to the screen. Infants' behaviour was also monitored via a centrally-located camera above the screen.

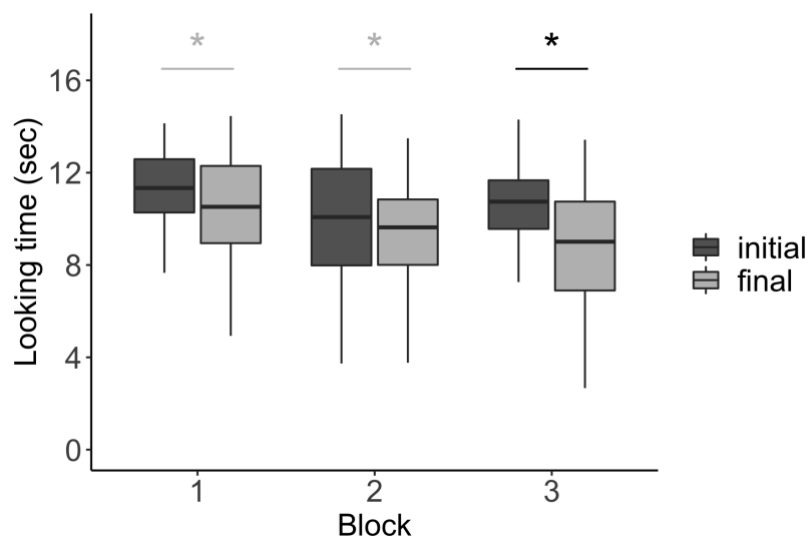
## Results

### Familiarisation

To explore the amount of looking time over the course of the experiment, mean total looking time during familiarisation was submitted to a one-way ANOVA with a within-subject factor *Block* (1, 2 and 3). The analysis revealed a near significant effect of *Block*,  $F(2, 42) = 2.66$ ,  $p = .08$ ,  $\eta_g^2 = .06$ . Pairwise comparisons using paired t-tests with Bonferroni corrections revealed no difference in the mean looking time between any of the blocks, *Block* 1 ( $M = 43.62$ ,  $SD = 8.15$ ), *Block* 2 ( $M = 38.08$ ,  $SD = 10.92$ ), *Block* 3 ( $M = 38.48$ ,  $SD = 7.83$ ). There was no difference in the amount of looking time to the exemplars from each of the two categories,  $t(25) = -.79$ ,  $p = .43$ ,  $d = .25$ .

To determine whether there was a decrease in the looking time within each of the three familiarisation blocks, mean looking time in the first two trials was compared to the mean looking time in the last two trials of each familiarisation block. Figure 3 depicts the findings. There was a significant decrease in the mean looking time in the final familiarisation block and a trend in the first and the second familiarisation blocks, *Block 1*<sup>3</sup>:  $Z = -1.81$ ,  $W = 247$ ,  $p = .07$ ,  $d = .35$ , *Block 2*:  $t(25) = -1.83$ ,  $p = .08$ ,  $d = .27$ , *Block 3*:  $t(21) = -2.94$ ,  $p < .01$ ,  $d = .68$ .

To summarise, the analyses of the looking time during familiarisation blocks suggested that infants became increasingly familiar with the category exemplars over the course of familiarisation. The lack of a significant drop in looking times across blocks might be driven by the presence of labels. Previous studies have demonstrated that infants tend to look longer at the visual stimuli and show less decrease in looking times when the visual stimuli are accompanied by labels (Althaus & Mareschal, 2014; Plunkett et al., 2008). In addition, the present experiment employed a gaze-contingent experimental paradigm, which has been shown to have positive effects on infants' engagement in a task (Wang et al., 2012).



<sup>3</sup> As Shapiro test revealed that the data was not normally distributed, the data was analyzed with the non-parametric Wilcoxon's Signed Rank test.

Figure 3. Mean looking time in the initial phase (first two trials) and in the final phase (mean looking time in the last two trials) of each familiarisation block of Experiment 1 (black asterisk depicts  $p < .05$ , grey asterisk depicts  $p < .08$ ).

### Category formation test

To test for category formation, infants were presented with two items, one of the category prototypes and an inter-category item (representing the overall average of both categories). Preference for the inter-category item was used as an index of infants' category formation. If infants formed two categories, we expected they would spend more time looking at the inter-category item because that item should appear less familiar and more surprising (cf. Younger, 1985; Plunkett et al., 2008). If infants formed no categories, we would expect no systematic looking preferences.

Preference scores were calculated by dividing the time spent looking at the inter-category item by the total time spent looking at both objects. The scores in the two trials were averaged and a one-way ANOVA with the within-subject factor *Block* (1, 2, and 3) revealed no significant effect, suggesting that there was no difference in the preference scores between blocks,  $F(2, 42) = .15$ ,  $p = .85$ ,  $\eta_s^2 = .004$ . Planned comparisons of the mean preference in each test block against chance (0.5) revealed a preference for the inter-category item in all three test blocks, *Block 1*:  $t(25) = 2.45$ ,  $p = 0.022$ ,  $d = .48$ ; *Block 2*:  $t(25) = 2.16$ ,  $p = 0.041$ ,  $d = .42$ ; *Block 3*:  $t(21) = 2.36$ ,  $p = 0.028$ ,  $d = .50$ , see Figure 4a. As in Plunkett et al., (2008), a preference for the inter-category item constitutes evidence for the formation of 2 categories.

### Novelty preference test

The final trial in each test block served to validate that infants' looking preferences were driven by novelty preference. Infants were presented with one of the familiarisation items along with a novel item. If infants' looking was driven by novelty preference, we would expect they would spend more time looking at the novel item. In contrast, if infants

displayed a familiarity preference, it was expected they would look more at the previously familiarised item.

Mean looking preferences in the novelty preference test trials were calculated by dividing the time spent looking at the novel item by the total looking time at both items. A one-way ANOVA with within-subject factor *Block* (1, 2, and 3) revealed no effect of *Block*,  $F(2, 42) = 1.61$ ,  $p = .21$ ,  $\eta_g^2 = .04$ . Planned comparisons of the mean preference in each test block against chance (0.5) indicated that there was a trend towards a novelty preference in the first block and a significant novelty preference in the second and third blocks (*Block* 1:  $t(25) = 1.74$ ,  $p = 0.095$ ,  $d = .34$ ; *Block* 2:  $t(25) = 2.49$ ,  $p = 0.019$ ,  $d = .49$ ; *Block* 3:  $t(21) = 3.20$ ,  $p = 0.004$ ,  $d = .68$ , Figure 4b).

To summarise, the analyses of the looking preferences in the category formation test trials showed that infants had learnt labelled categories already after one block of familiarisation and expressed this same pattern of preference in all test blocks. Novelty preference tests support this conclusion by revealing that infants' looking in all three test blocks was driven by novelty preference.

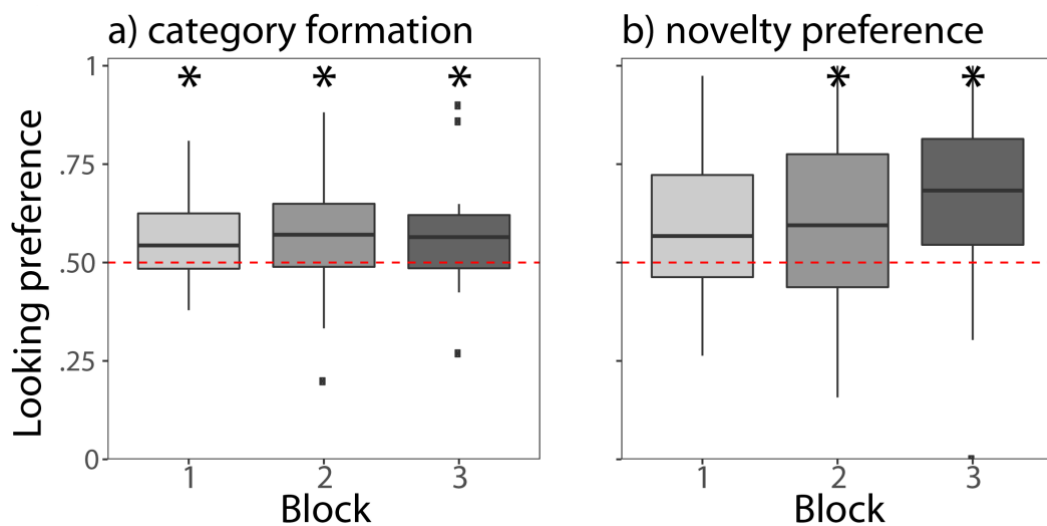


Figure 4. Mean looking preference scores in: (a) the category formation test (red dashed line depicts expected preference score if performance was at chance, scores above chance indicate preference for the inter-category item, suggesting infants formed two categories); (b) the novelty preference test (preference scores above 0.5 indicate preference for the novel item; values less than 0.5 indicate familiarity preference, black asterisk depicts  $p < .05$ ).

## Discussion

Experiment 1 explored infants' ability to form novel visual categories with an entirely novel set of stimuli inspired by stimuli used in several previous categorisation studies (Plunkett et al., 2008; Younger, 1993; Younger & Cohen, 1986). Infants were familiarised with exemplars of two categories accompanied by category-specific labels. Results revealed that infants were successful in forming two categories already after a single block of familiarisation. Infants showed evidence of category learning throughout the three blocks of testing. Moreover, the results revealed that infants are able to learn novel categories from a multi-item display with four items presented at the same time, when category exemplars are labelled in a gaze-contingent manner.

Despite the processing load posed by presenting four items at the same time, labels did not interfere with learning, and evidence of category formation was observed already after one block of familiarisation. This is in line with findings of Plunkett, Hu and Cohen's (2008) Experiment 3 which revealed that infants are able to form two categories in the presence of category-specific labels. The present study extends these findings and demonstrates that infants show evidence of category learning even when presented with more complex categories. While previous studies presented infants with the overlapping feature levels in familiarisation and at test (Plunkett et al., 2008; Younger & Cohen, 1983), the stimuli set used in the present study comprised of an extended feature range. This provided the opportunity to test category formation with a completely novel set of feature levels at test. The fact that infants show evidence of category learning in this case suggests that they are able to apply category knowledge not only to previously familiarised features, but also to generalise this information to entirely novel feature levels.

The present set of results contrasts with Deng and Sloutsky's (2015) finding that labels have detrimental effects on infants' ability to organise novel items into a single category. Differences in contexts in which labels were presented in the two studies might be responsible for these diverging effects. Deng and Sloutsky presented the label embedded in a carrier phrase ("Look! This is a Flurp"). In contrast, in the present study labels were presented in isolation as citation forms ("Perto!"). Perhaps labels presented embedded in a phrase are not as salient as labels presented in isolation. To detect the embedded label, an infant must segment the speech signal, detect the label, and make inferences about its' relation to the regularities of the visual input, while at the same time processing the available visual information. As a result, this might pose an additional processing load compared to labels presented in isolation. Importantly, in the present experiment labels were not presented synchronously with the visual information, which is found to have detrimental effects on categorisation (Althaus & Plunkett, 2015).

While spoken labels represent salient transient features processed in a different sensory domain to the to-be-categorised objects, motion is processed within the same modality as the visual features. To determine whether there are uniform effects of labels and motions on categorisation when tested under similar conditions, in the following experiment we explored the impact of motion on categorisation.

## EXPERIMENT 2

This experiment investigated the impact of motion on novel category learning. Infants were presented with the same visual categories and exposed to an identical experimental design as in Experiment 1, with the exception that instead of labels, in Experiment 2 exemplars were accompanied by category-specific motions during familiarisation.

### Method

#### Participants

A total of 27 10-month-old infants took part in this study (13 girls,  $M_{age} = 10.37$ ,  $age\ range = 9.49 - 11.07$ ). All participants were full-term babies with no known health conditions, recruited at the local maternity ward and came from monolingual British-English families. The study was approved by the University of Oxford Medical Sciences Ethics Committee. Two infants were excluded from the analysis due to insufficient looking time accumulated during familiarisation (more than 2 SD below the mean). One additional participant was excluded from the analysis of the final block as their session was ended at the beginning of the third block due to fussiness.

#### Experimental design

The stimulus set and experimental design were the same as in Experiment 1, with the exception that during familiarisation exemplars were accompanied by a category-specific motion: jumping or rocking.

**Familiarisation.** Once the infant fixated one of the four displayed items, the fixated item began moving and continued to move for a total of 400 ms, irrespective of whether the infant maintained her fixation on the object. Motion duration was set to 400 ms to equate that of the duration of labels in Experiment 1. As in Experiment 1, while moving, the fixated object became illuminated, to emphasise the gaze-contingent nature of the task. Two

category-specific types of motion were used: Familiarisation items could either rock (angle change of 2 degrees every 16ms, ranging from -10 to + 10 by 2 steps of degrees) or jump (vertical displacement ranging from -25 to +25 with steps of 5 degrees every 16ms – refresh rate). The motion-category association was counterbalanced across infants, i.e., for half of the participants, one category was associated with rocking whereas for the other half of the participants, this category was associated with jumping. The distribution of exemplars across trials and their locations was identical to Experiment 1.

**Test.** After each familiarisation block, infants were presented with a test block. Test blocks were identical to Experiment 1. All objects remained static during these test trials.

**Motion Test.** Since motion may be a highly salient feature of the category representation, omitting this feature, i.e., presenting a static test, might prevent infants from expressing evidence of category learning. To test for this possibility, in addition to the standard category formation tests, at the end of the final test, infants' category formation was evaluated with two trials in which the two category prototypes were presented side-by-side. Both prototypes were animated from the onset of the trial. In one trial, items were animated with a movement specific to one of the categories (e.g. rocking), while in the other trial they moved in a fashion associated with the other category (e.g. jumping). Hence, in both trials one item had a category-congruent motion and the other item moved in a way incongruent with its category. Again, each trial started with the display of an animated star, which remained on screen until fixated by the infant. Then, the presentation of the two prototypes followed for 6,000 ms. The position of the two prototypes and movement type were counterbalanced across participants. The final trial was a novelty preference trial, in which one of the familiarisation items was presented along with a novel, out-of-category object.

## Results

### Familiarisation

To explore the amount of looking time across the three familiarisation blocks, the mean looking time was submitted to a one-way ANOVA with a within-subject factor *Block*. The analysis revealed a near significant main effect of *Block*,  $F_g(1.57, 36.16) = 3.05$ ,  $p = .07$ ,  $\eta_g^2 = .04$ .<sup>4</sup> Pairwise comparisons using paired t-tests with Bonferroni corrections revealed no difference in the mean looking time between *Block 1* ( $M = 41.51$ ,  $SD = 7.97$ ) and *Block 2* ( $M = 41.88$ ,  $SD = 10.38$ ,  $p = 1$ ), or between *Block 1* and *Block 3* ( $M = 38.26$ ,  $SD = 10.80$ ,  $p = .46$ ). In contrast, there was a significant difference in the looking times in *Block 2* and *Block 3* ( $p = .026$ ). There was no difference in the amount of looking time to the exemplars from each of the two categories,  $t(24) = 1$ ,  $p = .32$ ,  $d = .26$ .

To investigate whether there was a decrease in looking time within each familiarisation block, mean looking time in the initial two trials was compared to the mean looking time in the final two trials in each familiarisation block (see Figure 5). A significant decrease in the mean looking time was present only in the first familiarisation block, with a trend in the second block, *Block 1*:  $t(24) = 2.22$ ,  $p = .04$ ,  $d = .51$ , *Block 2*:  $t(24) = 1.87$ ,  $p = .07$ ,  $d = .35$ , *Block 3*:  $t(23) = .92$ ,  $p = .36$ ,  $d = .22$ .

To summarise, the analyses of the familiarisation looking times revealed that there was some decrease in the mean looking time over the course of familiarisation. Further analyses suggested that this was the case in the first two familiarisation blocks, suggesting that infants become increasingly familiar with these categories.

---

<sup>4</sup> Mauchly's test indicated violation of the sphericity assumption,  $W = 0.73$ ,  $p = .03$ , therefore, the degrees of freedom were corrected using the Greenhouse-Geiser estimates of sphericity ( $\epsilon = .786$ ).

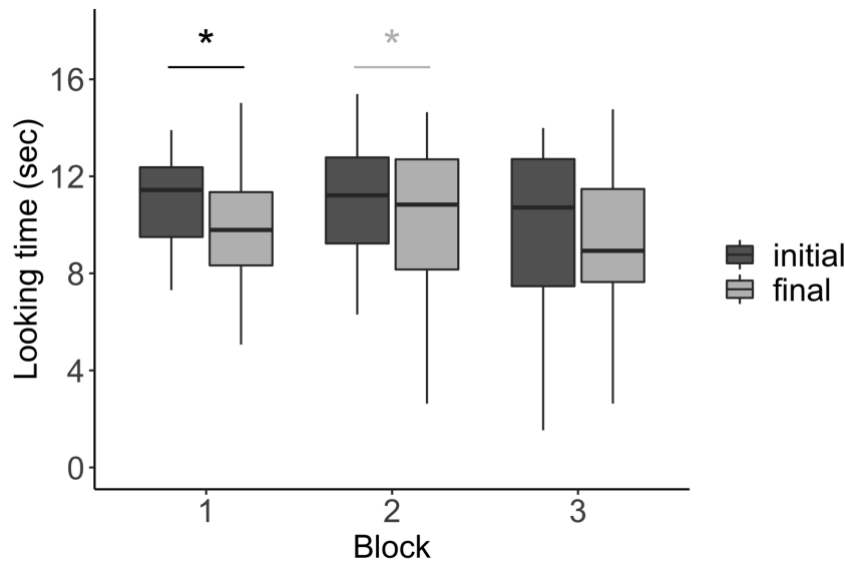


Figure 5. Mean looking time in the initial familiarisation phase (mean looking time in the first two trials) and in the final familiarisation phase (mean looking time in the last two trials) in each familiarisation block of Experiment 2 (black asterisk depicts  $p < .05$ , grey asterisk depicts  $< .08$ ).

### Category formation test

Following each familiarisation block, category formation was tested with two trials, as in Experiment 1. Preference for the inter-category item was calculated by dividing looking time to the inter-category item by the total time looking at both objects, and averaged for the two trials. As shown in Figure 6a, infants showed no preferences in any of the three test blocks, suggesting that they failed to form any categories, *Block 1*<sup>5</sup>:  $Z = -.32$ ,  $W = 137$ ,  $p = .74$ ,  $d = .06$ ; *Block 2*:  $t(24) = .68$ ,  $p = .50$ ,  $d = .14$ ; *Block 3*:  $t(23) = .15$ ,  $p = .88$ ,  $d = .03$ .

### Novelty preference test

To validate that infants were engaged in the task, the final trial in each test block was a novelty preference trial. The novelty preference scores were submitted to a one-way ANOVA with a within-subject factor *Block* (1, 2, and 3). The analysis revealed no effect of *Block*,  $F(2, 46) = .56$ ,  $p = .57$ ,  $\eta_s^2 = .02$ . Planned comparisons revealed that infants expressed

<sup>5</sup> Shapiro test revealed that the data was not normally distributed, and the data was analysed using the non-parametric Wilcoxon's Signed Rank test.

a novelty preference only in the first test block, with a trend in the second block, Figure 6b, *Block 1*<sup>6</sup>:  $Z = -2.52$ ,  $W = 257$ ,  $p = .01$ ,  $d = .51$ , *Block 2*:  $t(24) = 1.79$ ,  $p = .08$ ,  $d = .36$ ; *Block 3*:  $t(23) = .86$ ,  $p = .39$ ,  $d = .17$ .

To summarise, category formation tests revealed no evidence of category learning.

The pattern of preferences in the novelty preference tests revealed a significant novelty preference in the first block, gradually decreasing in the following two blocks.

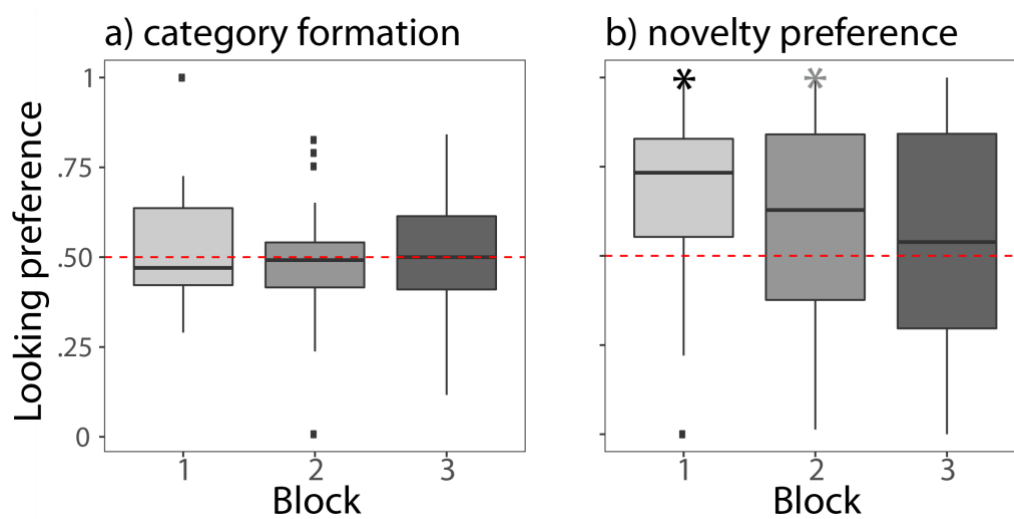


Figure 6. Mean looking preference scores in **(a)** the category formation test (red dashed line depicts expected preference score if performance was at chance (0.5), scores above chance indicate preference for the inter-category item, suggesting infants formed two categories). **(b)** the novelty preference test (preference scores above 0.5 indicate preference for the novel item; values less than 0.5 indicate familiarity preference, black asterisk depicts  $p < .05$ , grey asterisk depicts  $p < .08$ ).

### Motion test

After completing three interleaved familiarisation and test blocks, infants were presented with an additional test block. This block consisted of three trials. In the first two trials, the prototypes of the two categories were displayed side-by-side, and were coupled

<sup>6</sup> Shapiro test revealed that the data was not normally distributed, and the data was analysed using the non-parametric Wilcoxon's Signed Rank test.

with the same motion pattern. In one trial, both items displayed the rocking motion, and in the other both items displayed the jumping motion, so that in both trials one item was presented with a consistent motion and one item with an inconsistent motion pattern. The final trial the same novelty preference trial described above.

Preference for the prototype presented with the incongruent motion was calculated by dividing looking time to this incongruent compound by the total looking time to both prototypes, averaged for the two trials and compared against chance (0.5). A one-sample *t*-test revealed no preference,  $t(15) = .03$ ,  $p = .97$ ,  $d = .009$ . There was no systematic preference for the novel item in the novelty preference trial,  $t(15) = 1.45$ ,  $p = .17$ ,  $d = .36$ .

The lack of evidence of category learning, or association of motion with the relevant prototypes, when the test items remained dynamic might suggest that infants failed to learn these categories. However, since there is no novelty preference in the final novelty preference task, it seems more likely that infants disengaged from the task by the time the dynamic test took place.<sup>7</sup>

## Discussion

Experiment 2 explored infants' ability to learn visual categories in the presence of category-specific motions. Infants were familiarised with four items simultaneously, with items moving when they were fixated. Each category was associated with a specific motion: one category with jumping and the other rocking. Category formation tests followed each familiarisation block and revealed no evidence of category learning. Infants also failed to show evidence of category learning in the Motion test where category prototypes were presented together with congruent and incongruent motions. These findings are consistent with those reported by Rakison & Poulin-Dubois (2002) and Rakison (2004) who also found

---

<sup>7</sup> In line with this claim is the fact that only 16 out of 25 participants successfully completed this block.

insensitivity to correlations between motion trajectories and perceptual features of objects in this age group.

On the other hand, the results are surprising, particularly given that motion has been shown to facilitate visual category learning in infants of the same age (Deng & Sloutsky, 2015). Why would motion facilitate category formation in one study, but hinder it in another? There are two key differences in the experimental designs of the two studies that might be responsible for these diverging results. The first key difference between the two studies is that, in the current study infants were asked to differentiate high-similarity items into two categories while Deng and Sloutsky (2015) familiarised infants with exemplars from one category. Thereupon, the two experiments pose different tasks: Splitting similar looking objects into two categories requires infants to discover a boundary between the two categories, whereas learning a single category might draw on a different process such as a focus on similarities, rather than differences.

Perhaps more importantly, in Experiment 2, motion was a *contrastive feature*: different types of motion were related to different categories. In such a case, infants could in principle use the category-specific motion to help solve the categorisation problem (as they did for labels in Experiment 1). In contrast, infants were presented with one motion in Deng and Sloutsky (2015). It is therefore possible that motion serves a purely attention-arousing role in 10-month-old infants: motion can attract and maintain infants' attention, but cannot be used as a contrastive feature to form contrasting categories.

Another possibility is that infants are able to use motion as a contrastive feature, but introducing motion increased the processing load. As a consequence, motion may have led to extending the learning process so that additional familiarisation experience is needed to form categories. In order to understand in more detail how infants deal with the stimuli during learning, we turn to an analysis of gaze behaviour during familiarisation.

### IMPACT OF LABELS AND MOTION ON LOOKING TIMES DURING FAMILIARISATION

Infants in both Experiments 1 and 2 were given the same opportunity to explore the category exemplars, i.e. familiarisation time was equated across the two experiments. However, offering the infants equal opportunities to inspect the objects during familiarisation does not guarantee that they will spend similar amounts of time doing so. The present analysis was conducted to explore the possibility that the observed difference in categorisation effects was driven by differing levels of familiarisation in the two experiments.

To test whether infants accumulated similar amounts of looking time in Experiments 1 and 2, mean total looking times were submitted to a two-way ANOVA with a between-subject factor *Experiment* (1 and 2) and within-subject factor *Block* (1, 2 and 3). The results revealed a significant main effect of *Block*,  $F(2, 88) = 4.77, p = .01, \eta_g^2 = .04$ . However, there was no main effect of *Experiment* nor an *Experiment*  $\times$  *Block* interaction (see Figure 7a). These results indicate that infants demonstrated similar levels of looking times in both experiments. High looking times across the familiarisation blocks in Experiment 1 are consistent with previous findings which indicate that auditory stimuli help maintain infants' attention towards visual stimuli in a categorisation task (Althaus & Plunkett, 2016; Baldwin & Markman, 1989; Plunkett et al., 2008; Robinson & Sloutsky, 2007; Sloutsky & Robinson, 2008). The results indicate that motion has a similar effect on infants' attention and engagement as labels, and suggest that attention to the stimuli as expressed in looking times during familiarisation cannot provide an explanation of the differences in the observed performance in category formation tests in Experiments 1 and 2.

Next, we compared the mean duration of fixations during familiarisation. The mean look duration was calculated as the average amount of time infants spent looking at a

familiarisation object before moving their gaze away from it. The mean look duration was averaged across all looks within a trial and across four trials of each familiarisation block. The mean look duration was submitted to a two-way mixed ANOVA with a between-subject factor *Experiment* (1 and 2) and within-subject factor *Block* (1, 2, and 3). There was a significant main effect of *Experiment*, revealing a longer look duration in Experiment 2 (motion),  $F(1, 44) = 18.89$ ,  $p < .001$ ,  $\eta_g^2 = .23$  (see Figure 7b). There was no effect of *Block* nor an *Experiment*  $\times$  *Block* interaction. Given that the overall amount of looking time was similar across the two conditions, the longer mean look duration in Experiment 2 was accompanied by fewer fixations. A two-way mixed ANOVA with a between-subject factor *Experiment* (1 and 2) and within-subject factor *Block* (1, 2, and 3) on the mean number of fixations revealed significant main effects of *Experiment*,  $F(1, 44) = 30.23$ ,  $p < .001$ ,  $\eta_g^2 = .337$ , and *Block*,  $F(1, 88) = 6.176$ ,  $p < .001$ ,  $\eta_g^2 = .035$ . There was no *Experiment*  $\times$  *Block* interaction. On average, infants made 9.5 fixations in Experiment 1, and 12.7 fixations in Experiment 2. There was a decrease in the number of fixations across blocks (*Block 1*:  $M = 11.9$ ; *Block 2*:  $M = 10.8$ ; *Block 3*:  $M = 10.8$ ).

In sum, while there was no difference in the total looking time accumulated during familiarisation, infants in Experiment 2 had longer mean look duration (and made fewer fixations). Greater mean look duration in Experiment 2 than Experiment 1 suggests that infants attended longer to objects that were moving than when the objects were labelled. Importantly, recall that both labelling and motion events lasted for 400 ms. This result therefore implies that even though beyond the event duration both trial types (label and motion) were identical, infants' gaze remained longer on objects in Experiment 2.

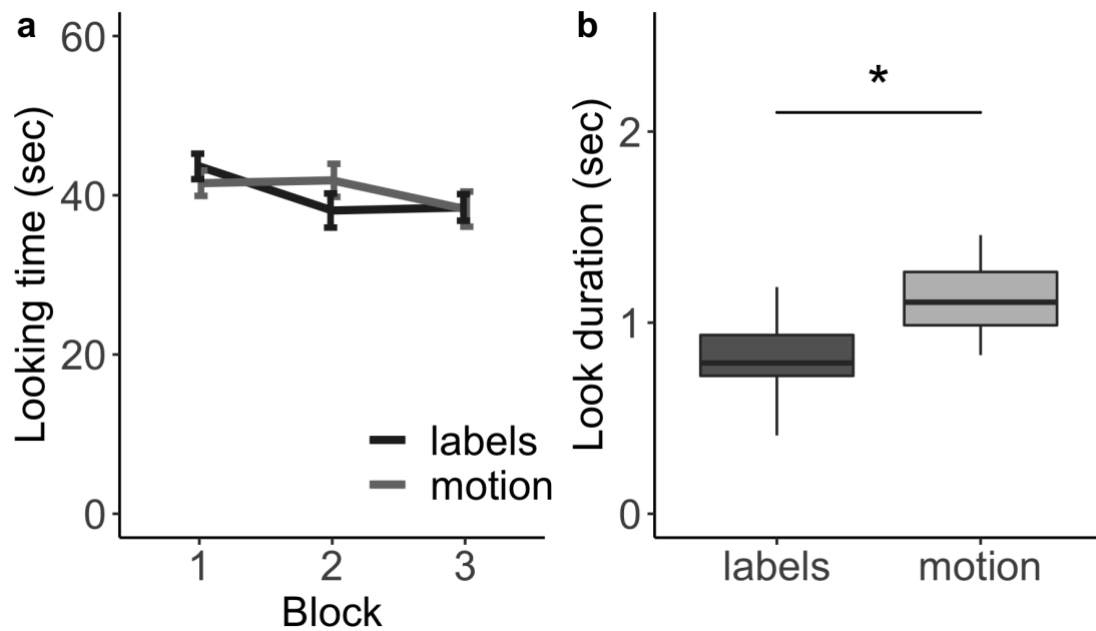


Figure 7. Comparison of mean looking times (a) and mean look duration (b) in Experiment 1 (label) and Experiment 2 (motion): Infants accumulate similar amount of looking time in both experiments, but mean look duration is longer in Experiment 2.

### Visual exploration during familiarisation

Familiarising infants with four items simultaneously provides them with an opportunity to compare multiple items while all of them are visually accessible. Infants may fixate objects at random or they may engage in a more systematic exploration of the 4 objects appearing on the screen. We examine 2 possibilities: namely, that infants tend to move their eyes to an object from the same category they just inspected, or that they move their eyes to a different category. The former would result in a prevalence of within-category shifts whilst the latter would result in more between-category shifts. The aim of the present analysis is to determine whether either of the 2 shift types is more prevalent, as compared to random exploration of 4 objects displayed on the screen.

Previous research suggests that shared labels highlight similarities between objects, but studies revealing this effect presented infants with one item at a time (e.g. Althaus & Plunkett, 2016). If labels direct attention towards similarities when multiple items are

simultaneously available, within-category shifts are expected to be more frequent. We also test whether between- or within-category shifts are more prevalent in the presence of motion. In addition, we ask whether exploration patterns are stable across familiarisation blocks or whether they change with increasing levels of familiarity. Such analyses have the potential to provide additional insights into the learning process.

First, we determine whether there was a difference in the mean number of shifts between objects in Experiments 1 and 2. We calculated the total number of shifts in each trial: a shift was defined as a transition of looking from one familiarisation object to another, i.e. 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 x 450 pixels square surrounding the object. The mean number of shifts was calculated by averaging the number of shifts across four familiarisation trials of each familiarisation block.

The mean number of shifts was submitted to a two-way mixed ANOVA with a between-subject factor *Experiment* (1 and 2) and within-subject factor *Block* (1, 2, and 3). The results revealed significant main effects of *Experiment*,  $F(2, 44) = 27.74$ ,  $p < .001$ ,  $\eta_g^2 = .32$ , and *Block*:  $F(2, 88) = 6.72$ ,  $p = .002$ ,  $\eta_g^2 = .04$  (see Figure ). There was no *Experiment* x *Block* interaction. Infants in Experiment 1 (label) made more shifts as compared to infants in Experiment 2 (motion), even though they accumulated similar amounts of looking time in the two experiments. This finding was to be expected since, as reported earlier, infants mean look duration in Experiment 2 was longer than Experiment 1 (see Figure 7b). The main effect of *Block* indicates that the number of shifts declined over the course of both experiments.

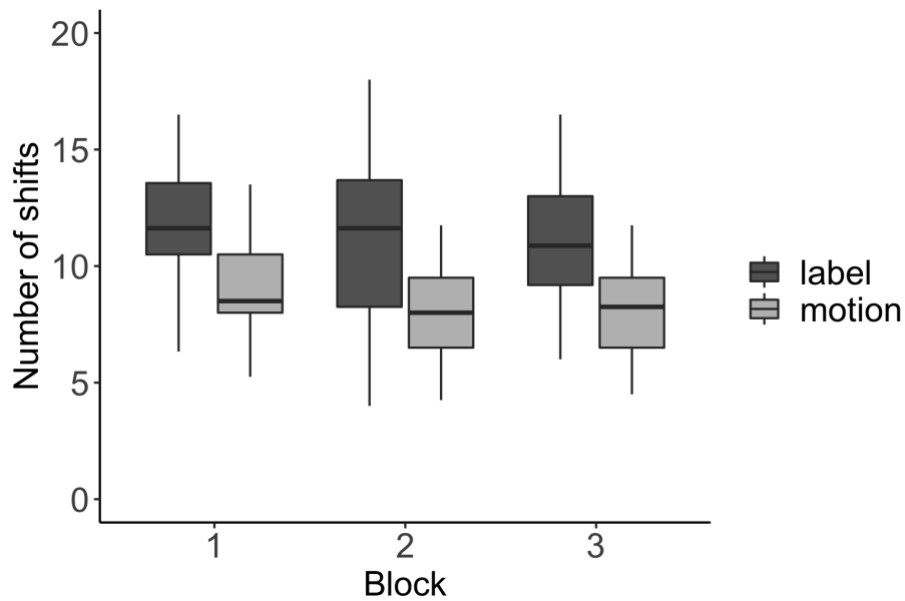


Figure 8. Mean number of shifts over the course of three familiarisation blocks in Experiment 1 (label) and Experiment 2 (motion).

Familiarisation trials were identical across the two experiments (with the exception of the presence of labels or motion), and infants accumulated similar amounts of looking time. Nevertheless, infants shifted their gaze less in Experiment 2, supporting the finding that motion is a highly salient feature (e.g. Aslin & Shea, 1990). Fewer gaze shifts and longer looks in the presence of motion show that motion not only attracts attention, but that infants also attend longer to objects that have moved.

Next, we explore whether within- or between-category shifts were more likely to occur during familiarisation. The mean number of within- and between-category shifts was calculated for each familiarisation block. To control for differences in the total number of shifts, the proportion of between-category shifts was calculated by dividing the number of between-category shifts by the total number of shifts. To determine whether there was any systematic preference for within- or between-category shifts, one-sample t-tests were used to compare the mean shift rate of between-category shifts against chance. Note that chance-level

is 0.667 because between-category shifts are twice as likely to occur than within-category shifts in a display that contains two items from each category.

Figure 8 depicts the rate of between-category shifts over the course of three familiarisation blocks. In Experiment 1 (labels), infants consistently made fewer between-category shifts than expected by chance in all three blocks, suggesting that infants engaged in a similarity-focused pattern of exploration, *Block 1*:  $t(25) = -9.80, p < .001, d = 1.92$ , *Block 2*:  $t(25) = -11.33, p < .001, d = 2.22$ , *Block 3*:  $t(21) = -6.75, p < .001, d = 1.44$  (see Figure 8a).

To determine whether there was a change in the shift rate across the three familiarisation blocks in Experiment 1, the mean between-category shift rates were submitted to a one-way ANOVA with a within-subjects factor *Block* (1, 2, and 3). The analysis revealed a significant main effect of *Block*,  $F(2, 42) = 3.68, p = .03, \eta_g^2 = .09$ . Pairwise comparisons (with Bonferroni corrections for multiple comparisons) revealed that there was no difference between *Block 1* and *Block 2* ( $p = .1$ ), nor *Block 1* and *Block 3* ( $p = 1$ ), but there was a significant difference between *Block 2* and *Block 3* ( $p = .02$ ). This indicates that infants made more between-category shifts in the final block as opposed to the second block.

In Experiment 2 (motion), infants also expressed significantly fewer between-category shifts than expected by chance in all three blocks, again suggesting that infants engaged in a similarity-based pattern of exploration, *Block 1*:  $t(24) = -12.02, p < .001, d = 2.40$ , *Block 2*:  $t(24) = -8.90, p < .001, d = 1.78$ , *Block 3*:  $t(23) = -13.04, p < .001, d = 2.66$ , (see Figure 8b). A one-way ANOVA with a within-subjects factor *Block* (1, 2 and 3) revealed that there was no change in the shift rates across the three familiarisation blocks,  $F(2, 46) = .18, p = .83, \eta_g^2 = 0$ .

In summary, these results reveal that infants in both experiments make significantly more within-category shifts than expected by chance. The tendency to fixate an object that is from the same category as the previously fixated object suggests that infants were more likely to search for similarities, rather than differences amongst objects, and this was the case when familiarisation items were accompanied by labels and by motions.

This result is in contrast to the findings from the test phase reported above. How is it possible that infants show preference for within-category shifts during familiarisation, but not sensitivity to categories at test? It is important to note that, during familiarisation, objects moved or were labelled only after being fixated. Therefore, a shift occurred before the transient stimulus engaged, rather than in response to the onset of a motion event or a label event. Given that infants only launched their shifts on the basis of static visual information alone, the observed bias for within-category shifts in Experiment 2 (motion) reveals that motion did not interfere with or overload visual processing, otherwise the shift patterns should not differ from chance. More specifically, if infants failed to encode perceptual features of familiarisation items due to the presence of motion, they would not be capable of detecting perceptual similarities and express a preference for within-category shifts. A similar argument holds for Experiment 1 (labels): infants launched shifts of fixation on the basis of static visual information alone. Given that these shifts were responsive to the visual similarity of the previously fixated object, we may conclude that the labelling event did not result in auditory overshadowing of the visual features.

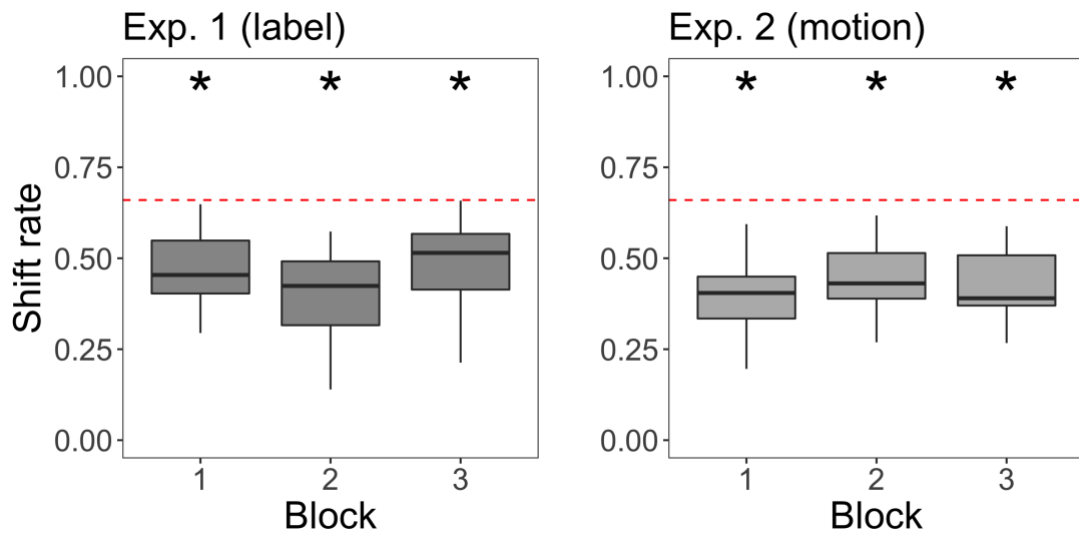


Figure 8. Proportion of between-category shifts over the course of three familiarisation blocks (red dashed line (0.67) depicts expected shifting rate if both shift types were equally likely).

### VISUAL EXPLORATION AT TEST

The previous analysis demonstrated that infants make significantly fewer gaze shifts during familiarisation when objects are accompanied by motion than in the presence of labels. In this section, we explore the patterns of visual exploration at test.

Given that the test phase was identical across the two experiments, any difference in visual exploration at test is likely to be driven by the effects of familiarisation experience. To characterise infants' visual exploration at test, we analysed 3 parameters: shift rates, mean looking duration and total looking time in category formation trials.

Infants made fewer gaze shifts during familiarisation but had longer mean look durations when objects were accompanied by motions than by labels. The present analysis compares the shift rates at test. Mean number of shifts was calculated by averaging the number of shifts across the two category formation trials of each test block. The mean number of shifts was submitted to a two-way mixed ANOVA with a between-subject factor *Experiment* (1 and 2) and within-subject factor *Block* (1, 2, and 3). The results revealed a significant main effects of *Experiment*, revealing the same pattern as in familiarisation:

infants in Experiment 2 (motion) made fewer shifts as compared to infants in Experiment 1 (label),  $F(1, 44) = 12.52$ ,  $p < .001$ ,  $\eta_g^2 = .14$ . There was no effect of *Block* nor an *Experiment x Block* interaction (see Figure 9a).

Next, we compared the mean look duration at test. The mean look duration was calculated as the average amount of time infants spent looking at an object before moving their gaze away from it. Mean look duration was obtained by averaging the duration of all looks within a trial (looks to either of the two test objects) and averaged across trials of each test block. The mean look duration was submitted to a two-way mixed ANOVA with a between-subject factor *Experiment* (1 and 2) and within-subject factor *Block* (1, 2, and 3). There was a significant main effect of *Experiment*, showing a longer look duration in Experiment 2 (motion),  $F(1, 44) = 14.86$ ,  $p < .001$ ,  $\eta_g^2 = .16$ . No effect of *Block* nor an *Experiment x Block* interaction were found (see Figure 9b).

Finally, we compared the total amount of looking time at test. The mean total looking time was averaged across trials of each test block, and submitted to a two-way mixed ANOVA with a between-subject factor *Experiment* (1 and 2) and within-subject factor *Block* (1, 2, and 3). There were no significant effects nor an interaction, suggesting that infants in Experiments 1 and 2 accumulated similar amounts of looking time at test (see Figure 9c).

In summary, even though test trials were identical across the two experiments, infants expressed different visual exploration patterns. While they accumulated similar amounts of looking time, infants shifted their gaze less and had longer looks in Experiment 2 (motion) as compared to infants in Experiment 1 (label). It seems likely that the gaze-contingent nature of the familiarisation phase enabled infants to discover that their looking triggers specific events, such as labels or motions. In turn, this might have affected their performance at a subsequent test, when there were no labels or motions presented. This result is in line with the results of a recent study which revealed that 6-8-month-old infants rapidly learn to perform

eye movements to trigger motion events and that infants quickly start to anticipate the effects of their eye movements (Wang et al., 2012).

The familiarisation phase in the Experiment 2 was gaze-contingent, and as a result, infants may have formed an anticipation of motion during familiarisation. By contrast, the test phase was not gaze-contingent. It is possible that infants' individual looks at the test objects were longer in Experiment 2 compared to Experiment 1 because they anticipated that objects should start moving.

Why is anticipation more apparent for motion than for labels even though both labels and motions were triggered in a gaze-contingent fashion? One possibility is that motion events occur in the same modality as the visual information about objects, while labels occur in a different modality. As a result, in Experiment 1 (labels), the visual input remains the same in familiarisation and at test (static in both phases), while in Experiment 2 (motion), the visual input contains static and dynamic information during familiarisation, and only static information at test. Familiarisation in Experiment 1 does not encourage infants to expect a visual change in the objects they fixate at test, whereas familiarisation in Experiment 2 encourages infants to expect some visual change at test. Longer fixations at test in Experiment 2 may simply reflect infants' anticipation of motion.

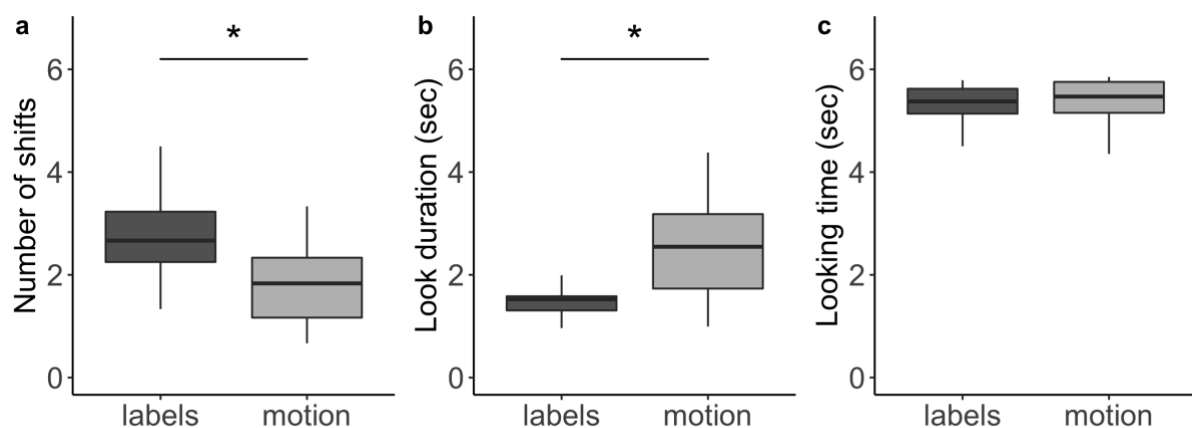


Figure 9. Visual exploration at test in Experiment 1 (labels) and Experiment 2 (motion): higher shift rates were present in Experiment 1 (a); shorter look duration in Experiment 1 (b); no difference in the mean amount of looking accumulated in Experiments 1 and 2.

### **First look as an index of category formation**

Visual exploration patterns in Experiment 2 (motion) revealed that infants made more within-category than between-category shifts during familiarisation, providing evidence of their sensitivity to the similarity structure of the categories. In addition, the gaze-contingent nature of the familiarisation phase motivated infants to anticipate that the objects will move once fixated. Driven by these anticipations of motion, infants maintained their looking for longer at the objects in both familiarisation and at test.

Given that test trials were not gaze-contingent, i.e. the test objects remained static, anticipation of motion that infants formed during familiarisation were violated at test. This means that looking behaviour over the course of the test trials might reflect different processes: at the very beginning of the test (immediately following familiarisation), infants' looking might be driven by the anticipation of motion; then, as there was no motion, later looking might reflect an extinction process: extinguishing the anticipation for the objects to move. As a result, novelty preference might not be the mechanism driving infants' looking at test.

One measure that is unaffected by this process is *the first look*. The first look in the test trial immediately following familiarisation is entirely driven by infants' familiarisation experience, and it occurs before there is any lack of gaze-contingent responsiveness from the objects. If infants' looking is driven by anticipation of motion rather than by novelty preference, as we assumed earlier, then the location of the first fixation should be the category prototype rather than the inter-category item, because the prototype is more similar to the categories seen during familiarisation, and therefore more likely to move.

To test this hypothesis, we examined whether the first look was more likely to be directed to the category prototype than to the inter-category item. The location of the first look (the prototype or the inter-category item) in the first test trial of each block was submitted to a binomial test.

In Experiment 1 (labels), the first look was equally likely to be directed to either test object (42 out of 74 first looks were to the prototype,  $p = .29$ ).

In Experiment 2 (motion), the location of the first look revealed a strong preference for the category prototype. The first look was to the category prototype in 69% of the trials (51 out of 74 first looks were to the prototype,  $p = .001$ ). In the first test block, 18 out of 25 first looks were to the prototype ( $p = .044$ ), in the second block 15 out of 25 ( $p = .42$ ), and 18 out of 24 in the third block ( $p = .022$ ), see Figure 11). This finding clearly demonstrates that infants' initial looking at test was driven by the anticipation that category prototypes will move. More importantly, this provides robust evidence that infants learned the two categories presented in Experiment 2 (motion).

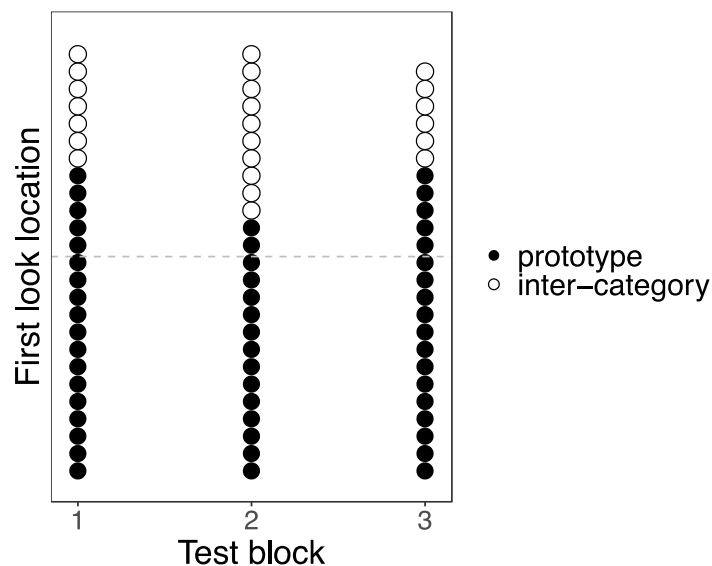


Figure 10. Location of the first look in the first test trial across 3 test blocks of Experiment 2 (dots represent individual data points).

To summarise, a first look analysis revealed infants' strong tendency to look first at the category prototype, likely because they were anticipating motion. This finding clearly indicates infants that they have an understanding of the category structure and can use distinctive motions as contrastive features to learn these categories.

These results also suggest that the standard category formation test based on novelty preference is not an appropriate tool for testing category learning in Experiment 2: gaze-contingent familiarisation motivates infants to form anticipations of motion and their looking at test is driven by this anticipation, which may compete with or overwrite novelty preference.

### EXPERIMENT 3

Experiment 2 tested infants' ability to form categories in the presence of category-specific motions, whilst Experiment 1 demonstrated that these categories are readily learnable in the presence of category-specific labels. Infants provided evidence of learning 2 categories after a single block of familiarisation in both experiments. Experiment 3 was conducted to explore whether infants can learn the same novel visual categories in silence, i.e. in the absence of any labelling or motion features. In this experiment infants were presented with the same visual categories as in Experiments 1 and 2, in three familiarisation blocks interleaved with three test blocks. As a cautionary note, the results of the present experiment cannot be compared directly to Experiments 1 and 2 since the familiarisation times are different. Pilot work indicated that infants do not sit for long enough when stimuli are played in silence. Consequently, familiarisation trials were shortened as compared to Experiments 1 and 2. Nevertheless, infants' performance in this experiment has the potential to evaluate their ability to form these novel visual categories in the absence of redundant transient stimuli

and to explore whether their visual search patterns during familiarisation show the same within-category preferences as found in Experiments 1 and 2.

Since the overall visual structure of the stimulus set in Experiment 3 is identical to that used in Experiments 1 and 2, we predict that infants will form 2 categories as indexed by a preference for the inter-category stimulus 4444 (see Figure 1). Such an outcome would parallel earlier findings for similarly structured category sets used by Younger (1985) and Plunkett et al (2008, Exp. 2).

## Method

### Participants

Thirty 10-month-old infants took part in this experiment (15 girls,  $M_{age} = 10.17$ ,  $age\ range = 9.60 - 10.69$ ). All participants were recruited at the local maternity ward and came from monolingual British-English families. All were full-term babies with no known health conditions. The study was approved by the University of Oxford Medical Sciences Ethics Committee. Two participants were excluded due to failure to establish tracking of their eye-movements and an additional one due to fussiness. Of the 27 participants in the final sample, 24 completed the second block and only 22 contributed to the data for the final of the three blocks.

### Stimuli and experimental design

The stimulus set was the same as in Experiments 1 and 2, but no additional transient features were presented. In addition, there was no gaze-contingent effects during familiarisation and familiarisation trials were 12,000 ms long, as compared to 16,000 ms in the previous experiments.

**Familiarisation.** There were four trials in each familiarisation block. Each trial was initiated with the presentation of an attention getter, followed by four exemplars presented simultaneously for 12,000 ms. The items remained on the screen for the duration of the trial.

**Test.** In each test trial, after an attention getter was presented for 2,000 ms, two test objects were presented simultaneously for 10,000 ms. The first two trials were category formation test trials in which the inter-category item (object 4444) and category prototype (2222 or 6666) were presented. The positions of the two objects were counterbalanced across the two trials. The third, novelty preference trial was always a novelty preference test in which one of the familiarisation items from the previous familiarisation phase was presented along with a novel, out-of-category object.

## Results

### Familiarisation

To explore the mean looking time over the course of the experiment, Figure 11. a one-way ANOVA with a within-subject factor *Block* revealed a significant main effect of *Block*,  $F(2, 42) = 11.78, p < .001, \eta_g^2 = .23$ . Pairwise t-tests with Bonferroni corrections for multiple comparisons revealed no difference in the amount of looking time between *Block 1* and *Block 2* (*Block 1*:  $M = 31.36, SD = 8.85$ , *Block 2*:  $M = 25.82, SD = 9.01$ ), but there was a significant decrease in *Block 3* ( $M = 19.19, SD = 10.18$ ), as compared to both *Block 1* ( $p < .001$ ) and *Block 2* ( $p = .023$ ). The observed reduction in looking is commonly observed in this type of experiment when stimuli are presented in silence (e.g. Plunkett et al., 2008).

Mean looking time in the initial two familiarisation trials was compared to the mean looking time in the final two trials to determine whether there was a decrease in the mean looking time within each familiarisation block, Figure 11. There was a significant decrease in the mean looking time in the first block, but not in the second and the third familiarisation

block, *Block 1*:  $Z = -2.22$ ,  $W = 281$ ,  $p = .03$ ,  $d = .43$ , *Block 2*:  $Z = -1.48$ ,  $W = 173$ ,  $p = .14$ ,  $d = .30$ , *Block 3*:  $Z = -.33$ ,  $W = 104$ ,  $p = .74$ ,  $d = .07$ .<sup>8</sup>

To summarise, these analyses reveal a reduction in the amount of looking time over the course of the three familiarisation blocks, suggesting that infants became increasingly familiar with the stimuli over the course of familiarisation.

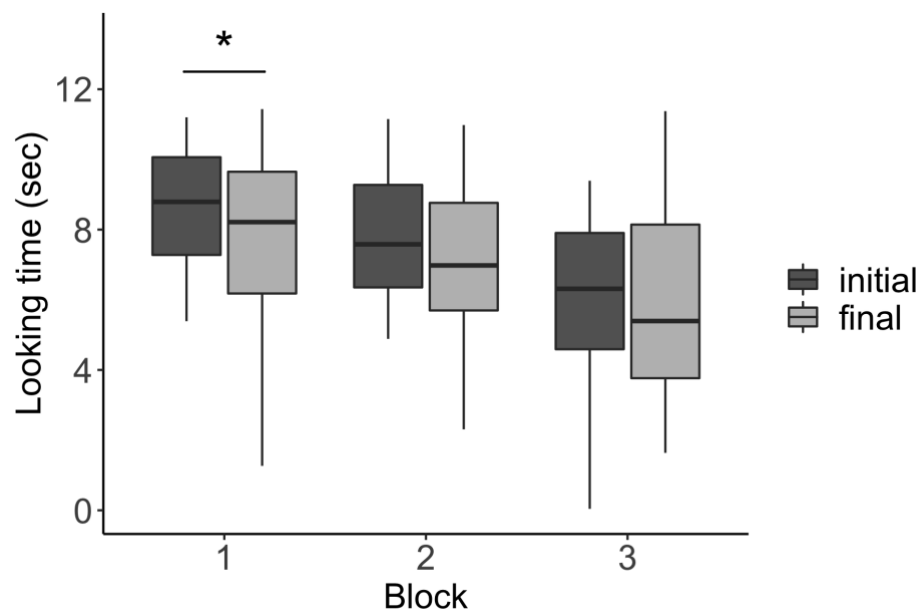


Figure 11. Mean looking time in the initial familiarisation phase (mean looking time in the first two trials) and in the final familiarisation phase (mean looking time in the last two trials) in each familiarisation block of Experiment 3 (black asterisk depicts  $p < .05$ ).

### Category formation test

Category formation was tested after each familiarisation block with two test trials. Preference score was calculated by dividing looking time to the inter-category item by the total looking time, and averaged for the two trials. A one-way ANOVA on the mean looking preference with a within-subject factor Block (1, 2, and 3) revealed a significant main effect

<sup>8</sup> Shapiro test revealed that the data was not normally distributed, and the data was analysed using the non-parametric Wilcoxon's Signed Rank test.

of Block,  $F(2, 42) = 4.04, p = .02, \eta_g^2 = .12$ . To perform planned comparisons, the average looking preference in each test block was compared against chance (0.5). There was a significant preference for the category prototype in the first block,  $Z = -1.99, W = 106, p = .046, d = .38$ <sup>9</sup>, followed by no preference for either of the test items in the second block,  $t(23) = -.22, p = .82, d = .04$ , and finally, there was a significant preference for the inter-category item in the final test block,  $t(21) = 2.21, p = .038, d = .47$  (see Figure 12a).

A first look analysis revealed no systematic preferences in the test blocks of Experiment 3. The first look was equally likely to be directed to either test object (33 out of 73 first looks were to the prototype,  $p = .48$ ).

### Novelty preference test

To validate that infants' looking at test was driven by novelty preference, the final trial was a novelty preference trial. One of the familiarisation items was presented along with a novel item. If infants' looking was driven by novelty preference, we would expect they would spend more time looking at the novel item.

A one-way ANOVA on the novelty preference scores with a within-subjects factor *Block* (1, 2 and 3) revealed no significant effect of Block,  $F(2, 42) = 1.24, p = .29, \eta_g^2 = .04$ . Planned comparisons revealed that infants expressed a novelty preference in the first block,  $t(26) = 3.003, p = .006, d = .58$ , no systematic preference in the second block,  $Z = .24, W = 159, p = .808, d = .052$ , and a novelty preference in the final test block,  $t(21) = 2.082, p = .049, d = .44$ , see Figure 12b.

In sum, category formation tests revealed that infants initially showed a preference for the category prototypes, followed by no preference, and finally a preference for the inter-

---

<sup>9</sup> Shapiro-Wilk test revealed that the data violated normality assumption, so the non-parametric Wilcoxon Signed Rank test was run.

category item in the third test block. These results indicate that infants initially formed a single category comprising of all familiarisation items, where the overall average – the inter-category item (4444) constituted the most familiar item and the category prototypes (2222 and 6666) were relatively peripheral. After further familiarisation experience, they formed two categories, with infants switching their novelty preference to the inter-category item (4444) between the first and the last familiarisation blocks (see Plunkett et al, 2008 and Younger, 1983 for similar switches between the ‘Broad’ and ‘Narrow’ conditions). This interpretation is further validated by the pattern of looking preferences in the novelty preference tests, which revealed that infants were in a novelty preference mode in the first and the last test blocks.

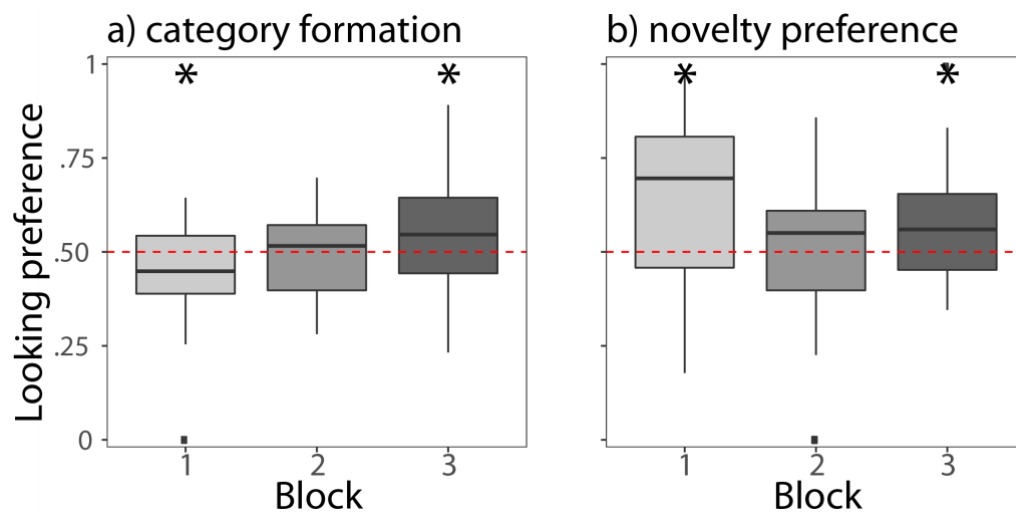


Figure 12. Mean looking preference scores in **(a)** test blocks (red dashed line depicts expected preference score if performance was at chance, scores above chance indicate preference for the inter-category item, suggesting infants formed two categories); **(b)** the novelty preference test (preference score above chance indicates novelty preference).

### Visual exploration during familiarisation

Experiment 1 and Experiment 2 revealed that in the presence of labels and motions, respectively, infants tend to make more within-category shifts during familiarisation. To test whether the same exploration pattern is present when infants learn these categories in silence,

mean between-category shift rates were compared against chance (0.67). The results revealed no systematic preference for either between- or within-category shifts, *Block 1*:  $t(26) = -1.49$ ,  $p = .15$ ,  $d = .29$ , *Block 2*:  $t(23) = -0.19$ ,  $p = .85$ ,  $d = .04$ , *Block 3*:  $t(21) = -0.89$ ,  $p = .38$ ,  $d = .19$ . In sum, these results suggest that there was no prevalent exploration pattern present during familiarisation in Experiment 3.

### Discussion

Experiment 3 explored infants' ability to learn the same novel visual categories as those used in Experiments 1 and 2, but in the absence of any motion or labelling features. The results revealed that, after the first familiarisation block, infants showed a preference for the category prototypes. After an additional block of familiarisation, infants showed no systematic preferences, and lastly, after three familiarisation blocks infants expressed a preference for the inter-category item. Novelty preference tests revealed that infants were in a novelty preference mode in the first and final test blocks. There was no prevalent exploration pattern present in infants' looking patterns during familiarisation. These results suggest that infants initially grouped all items into one category, and after additional familiarisation, reorganised the category space and differentiated these items into two categories.

Why do infants shift from forming one category to forming two categories? In the absence of labelling and motion features to highlight the similarity structure of category space, it is likely that infants initially encode overall similarities of all familiarisation items: items from both categories consist of the same features (i.e. all items have identically shaped hands, legs and antennae), and infants initially detect these similarities and use it as a basis for category formation. With more exposure to these categories, infants notice that feature levels are correlated (for instance, items with short legs always have many antennae) and this leads to a gradual emergence of feature-correlation-based categorisation. This is consistent with Younger & Cohen's (1986) discovery of a progression from detecting isolated features

at 4 months, followed by the perception of feature correlations of single items at 7 months, to feature-correlation-based categorisation at 10 months (Younger & Cohen, 1986). This gradual differentiation of categories based on the discovery of feature correlations has also been captured previously by computational models of semantic development (e.g. Plunkett, Sinha, Møller, & Strandsby, 1992; Rogers & McClelland, 2004).

Interestingly, in the present experiment infants showed no preference for a similarity- nor contrast-focused exploration during familiarisation. This is in contrast to the exploration patterns found in Experiments 1 and 2, which found a significant prevalence of similarity-focused exploration. Given that infants in all three experiments were presented with the same familiarisation items, it seems to be the case that labels and motion achieve a specific attention-directing effect that is not present when learning these categories in silence.

## GENERAL DISCUSSION

This study set out to explore the impact of auditory labels and motions on novel visual category learning. Experiment 1 investigated the impact of contrasting labels on infants' ability to form novel categories, while Experiment 2 explored how contrasting motions affect category formation. Lastly, Experiment 3 tested category formation when only static perceptual features were provided, i.e. in the absence of labels and motions.

All 3 experiments used interleaved train-test design, permitting an evaluation of the impact of stimulus exposure on category learning. Experiments 1 and 2 implemented a gaze-contingent presentation of the transient stimuli, with labels and motions triggered by infants' fixations upon specific objects. Standard preference scores at test in Experiment 1 revealed that category-specific labels supported category formation (indexed by a preference for the inter-category item) across all 3 train-test blocks. In Experiment 2, where category-specific motion was used, while the standard preference scores revealed no systematic effects for any

of the testing blocks, a first look analysis provided clear evidence of category learning in the first and final testing blocks. Lastly, in Experiment 3 infants succeeded in forming these categories when familiarised with items in the absence of any transient features, but not until the final block of testing.

This pattern of results demonstrates facilitating effects of contrastive labels and motions on visual category formation as compared to the presentation of the same set of objects in the absence of gaze-contingent transient features. In particular, infants learned the 2 categories already in Block 1 when the visual stimuli were accompanied by contrastive labels or motions, but failed to demonstrate evidence of learning the 2 categories in the silent condition until Block 3. It should be noted that the length of familiarisation trials in Experiment 3 was less than that in Experiments 1 and 2 (12 s vs 16 s) which may have contributed to the differences observed. However, infants also failed to show evidence of learning in Block 2 of Experiment 3, at the end of which they would have had considerably more exposure to the training stimuli. Taken together, these 3 experiments provide evidence that converges with earlier studies demonstrating the facilitative effect of labels on visual category formation (Fulkerson & Waxman, 2007; Waxman & Braun, 2005) and motion on visual category formation (Deng & Sloutsky, 2015). To our knowledge, these experiments offer the first demonstration that *both* motions and labels facilitate categorisation within similar experimental settings.

#### *Patterns of visual exploration during familiarisation*

Presenting infants with four items simultaneously provides them with an opportunity to compare items while all of them are visually accessible. Our results revealed that infants engage in a similarity-focused exploration of the familiarisation objects, making more within-category shifts when these objects are accompanied by category-specific labels (Experiment 1) and by category-specific motions (Experiment 2). In contrast, in the absence of labels and

motion, infants tend to fixate objects at random (Experiment 3). This pattern of findings indicates that labels and motions promote a similarity-focused visual exploration of the category exemplars whereas viewing the same set of objects in silence does not.

The process of comparison has an important role in category learning. 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 a search for commonalities between compared items so that similarity detection is at the core of category learning (Gentner & Namy, 1999; Namy & Gentner, 2002). Our results revealed that the nature of comparison in a visual category learning task is mediated by the presence of additional transient features – when labels and motions accompany category instances infants tend to engage in a similarity-focused exploration, but this is not the case when these items are presented in silence.

The presence of a similarity-focused exploration in Experiment 1 is in line with existing research demonstrating that labels direct infants' attention to similarities between objects (Althaus & Mareschal, 2014; Althaus & Plunkett, 2016). While these studies presented infants with one item at-a-time, the present study supports and further extends this claim by showing that labels also highlight similarities when infants are presented with multiple items simultaneously. The attention-directing function of labels becomes even more relevant when infants are faced with more cluttered environments which pose high processing demands. By inducing similarity-focused exploration, labels guide infants' attention and contribute to category learning.

Experiment 2 revealed that the presence of contrastive motions also promotes similarity-focused exploration during familiarisation. To the best of our knowledge, this is the first study to reveal that contrastive motion can promote attention to similarities.

Visual exploration focusing on similarities does not seem to be necessary for learning the categories in these experiments. Infants in Experiment 3 did not show any systematic exploration patterns during familiarisation, but they were still able to learn the categories.

It is possible that heightened processing demands, introduced by the presence of labels and motions made categories more difficult to discover, and, as a result, infants had to perform more comparison to learn the categories. However, this is unlikely given that infants showed evidence of category learning already after 1 block of familiarisation in Experiments 1 and 2, whereas it took 3 blocks of familiarisation for infants to succeed in Experiment 3.

#### *Infants' learning in gaze-contingent settings*

Experiments 1 and 2 employed a gaze-contingent paradigm during familiarisation: labels and motions were triggered by infants' gaze. Instead of being a passive recipient of a pre-specified presentation of labelling and motion events, the gaze-contingent paradigm enables an infant to actively shape her learning experience. Indeed, infants are shown to be sensitive to the responsiveness of gaze-contingent paradigms and quickly learn that they are in control (Wang et al., 2012). The present study demonstrated that infants readily anticipate that fixating an object will trigger a response. This was particularly evident in Experiment 2 (motion), where patterns of visual exploration clearly reflected infants' expectations for the objects to move. This expectation about motion guided their looking not only during familiarisation, but also at test.

Introducing gaze-contingent eye-tracking paradigms to the study of infants' learning has multiple benefits. It provides active learning settings, as an infant decides when and where to look next. In addition, in the domain of infant categorisation research, this creates an opportunity to move beyond a traditional familiarisation-test experimental design in which learning outcomes are measured at the end of the experiment. Instead, a combination of eye-

tracking technology and gaze-contingent paradigms provides a means of measuring the learning process as it happens, during familiarisation as well as test.

In the present study, the patterns of visual exploration provided novel insights into infants' understanding of the category structure, over and above the standard measures based on the category formation tests. The observed shifting bias during familiarisation clearly showed that infants were sensitive to the similarity structure of the categories even during familiarisation. Complementing the standard categorisation tests with these visual exploration measures has the potential to provide a more refined understanding of infants' category learning.

The present study also shows that the use of gaze-contingent tasks can change the content of infants' learning, particularly by enabling expectations of stimulus responsiveness during familiarisation. As a consequence, a non-gaze contingent test based on novelty preference, as implemented in the present set of experiments, can be rendered inadequate as a test of category formation. This test is based on the assumption that infants' looking is driven by novelty preference. However, if gaze-contingent familiarisation sets up expectations about motion that drive infants' looking, these expectations may mask or compete with the novelty preference in a non-gaze-contingent test, resulting in the unsystematic patterns of overall preference observed in Experiment 2.

One way the testing methodology could be adapted to achieve a successful test of novelty preference is to implement a gaze-contingent design in the category formation test trials. If infants' looking is driven by expectations about motion, presenting dynamic test objects that respect or violate these expectations can provide another measure of infants' understanding of these categories. We attempted to do this in the very final test block in Experiment 2. In this test, infants were presented with prototypes of both categories at the same time: one had a category-congruent motion and the other moved in a way incongruent

with its category. However, these trials were included at the very end of the experiment, and by that time many infants had already begun to disengage from the task. If infants encoded which motion is related to which category, the incongruent prototype-motion combination should attract more attention as it violates infants' expectations. Given the present set of findings, this test has the potential to provide another test of category formation, i.e. infants' ability to generalise this information to novel category instances, if administered immediately after familiarisation.

### *The role of labels in visual category learning*

An important manipulation in the present study was that labels (and motions) were entirely redundant. Perceptual features provided sufficient information for category formation. Nevertheless, despite being redundant, labels had a strong impact on category learning.

The present study revealed that introducing labels as additional stimuli modulated the process of category formation (Experiment 1) as compared to when the same categories are presented alone (Experiment 3). Given that infants showed evidence of category learning already after one block of familiarisation in Experiment 1 (labels) but not until the third block in Experiment 3, it is tempting to claim that labels facilitate learning. However, this claim is not fully supported by the data, because of the differences in experimental designs implemented in Experiments 1 and 3 (see earlier discussion). Nevertheless, the present study contributes to the growing body of research demonstrating the impact of labels on categorisation (Althaus & Mareschal, 2014; Althaus & Plunkett, 2015; Plunkett, 2010; Plunkett et al., 2008; Waxman & Markow, 1995).

It is important to note the observed effect of labels does not indicate that infants formed label-object associations, i.e. learned labels for individual exemplars. The present

study did not investigate whether infants are able to learn names for exemplars from the two categories, but rather how the presence of category-specific labels affects the process of organising these exemplars into categories.

#### *The role of motion in visual category learning*

Experiment 2 demonstrated that infants were able to learn novel visual categories in the presence of contrastive motions. The analysis of infants' looking patterns during familiarisation suggested that infants were sensitive to the overall similarity structure of the categories in all 3 familiarisation blocks. The analysis of the first look at test provided clear evidence of category formation.

This result is in line with the existing research suggesting that motion is a highly salient feature, and infants readily form expectations about the movement of objects (Markson & Spelke, 2006; Pauen & Träuble, 2009). The present study demonstrates that infants are able to anticipate motion in the context of learning novel categories, and differentiate a set of highly similar objects into two categories. Experiment 2 provides evidence that contrastive motions do more than highlight attention to category exemplars and their relative similarity; motion can be a defining characteristic of a category and infants can use it as a contrastive feature to learn novel categories.

#### *Concluding remarks*

While some studies found that labels facilitate categorisation (Fulkerson & Waxman, 2007; Waxman & Braun, 2005), other studies found that labels can hinder category learning (Deng & Sloutsky, 2015; Sloutsky & Robinson, 2008). The present study demonstrates that both labels and motions can support category learning and can serve as contrastive features that facilitate the discovery of category boundaries. Both labels and motion attract attention and promote similarity-focused exploration.

The present study also provides some novel insights to the ongoing discussion about the status of labels in category learning: some researchers claim that labels have a privileged status as category markers (Waxman & Braun, 2005; Waxman & Markow, 1995), whereas others claim that labels act as features (Gliozzi et al., 2009; Plunkett et al., 2008). This set of findings demonstrates that labels are not the only type of features that can promote categorisation: motion features can also facilitate category learning. These findings pose the question as to whether both labels and motions may have a privileged status or instead, whether both serve as features that have the power to direct attention towards similarities between objects, which in turn facilitates learning. While both labels and motions facilitate category learning, the outcomes, i.e. behavioural indices of learning might be different – label-driven categories can be captured in infants' overall looking preferences, whereas motion-driven categories can be indexed in looking anticipation of motion events.

In conclusion, the experiments described here clearly demonstrated that additional stimuli such as labels and motion, despite their transient nature, have a strong impact on category formation. In particular, both types of stimuli can steer infants' exploration patterns and promote attention to similarities. Lastly, this study revealed advantages of using gaze-contingent paradigms to study infants' visual category learning and demonstrated how parameters of visual exploration can provide an index of category learning as it unfolds.

### **Acknowledgments**

We thank the families who took part in this study. We also thank Irina Lepadatu for her assistance with data collection, and Mihaela Duta for her help with implementing the gaze-contingent experimental design.

### References

- Althaus, N., & Mareschal, D. (2014). Labels Direct Infants' Attention to Commonalities during Novel Category Learning. *PLoS ONE*, 9(7), 99670.  
<https://doi.org/10.1371/journal.pone.0099670>
- Althaus, N., & Plunkett, K. (2015). Timing matters: The impact of label synchrony on infant categorisation. *Cognition*, 139, 1–9. <https://doi.org/10.1016/J.COGNITION.2015.02.004>
- Althaus, N., & Plunkett, K. (2016). Categorization in infancy: labeling induces a persisting focus on commonalities. *Developmental Science*, 19(5), 770–780.  
<https://doi.org/10.1111/desc.12358>
- Althaus, N., & Westermann, G. (2016). Labels constructively shape object categories in 10-month-old infants. *Journal of Experimental Child Psychology*, 151, 5–17.  
<https://doi.org/10.1016/j.jecp.2015.11.013>
- Aslin, R. N., & Shea, S. L. (1990). Velocity Thresholds in Human Infants: Implications for the Perception of Motion. *Developmental Psychology*, 26(4), 589–598.  
<https://doi.org/10.1037/0012-1649.26.4.589>
- Baldwin, D. A., & Markman, E. M. (1989). Establishing Word-Object Relations: A First Step. *Child Development*, 60(2), 381–398. <https://doi.org/10.1111/j.1467-8624.1989.tb02723.x>
- Burnham, D. K., Vignes, G., & Ihsen, E. (1988). The effect of movement on infants' memory for visual compounds. *British Journal of Developmental Psychology*, 6(4), 351–360.  
<https://doi.org/10.1111/j.2044-835X.1988.tb01107.x>
- Dannemiller, J. L. (2000). Competition in Early Exogenous Orienting between 7 and 21 Weeks. *Journal of Experimental Child Psychology*, 76(4), 253–274.  
<https://doi.org/10.1006/JECP.1999.2551>
- Deng, W. S., & Sloutsky, V. M. (2015). Linguistic labels, dynamic visual features, and

- attention in infant category learning. *Journal of Experimental Child Psychology*, 134, 62–77. <https://doi.org/10.1016/j.jecp.2015.01.012>
- Fulkerson, A. L., & Waxman, S. R. (2007). Words (but not Tones) facilitate object categorization: Evidence from 6-and 12-month-olds. *Cognition*, 105, 218–228. <https://doi.org/10.1016/j.cognition.2006.09.005>
- 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](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](https://doi.org/10.1016/S0885-2014(99)00016-7)
- Giozzi, V., Mayor, J., Hu, J. F., & Plunkett, K. (2009). Labels as features (not names) for infant categorization: A neurocomputational approach. *Cognitive Science*, 33(4), 709–738. <https://doi.org/10.1111/j.1551-6709.2009.01026.x>
- Markson, L., & Spelke, E. S. (2006). Infants' Rapid Learning About Self-Propelled Objects. *Infancy*, 9(1), 45–71. [https://doi.org/https://doi.org/10.1207/s15327078in0901\\_3](https://doi.org/https://doi.org/10.1207/s15327078in0901_3)
- 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>
- Noles, N. S., & Gelman, S. A. (2012). Preschool-age children and adults flexibly shift their preferences for auditory versus visual modalities but do not exhibit auditory dominance. *Journal of Experimental Child Psychology*, 112, 338–350. <https://doi.org/10.1016/j.jecp.2011.12.002>
- Pauen, S., & Träuble, B. (2009). How 7-month-olds interpret ambiguous motion events:

Category-based reasoning in infancy. *Cognitive Psychology*, 59, 275–295.

<https://doi.org/10.1016/j.cogpsych.2009.06.001>

Plunkett, K. (2010). The Role of Auditory Stimuli in Infant Categorization Infant

Categorization. In L. Oakes, C. Cashion, M. Casasola, & D. Rakison (Eds.), *Infant perception and cognition: Recent advances, emerging theories, and future directions* (pp. 203–221). Oxford University Press. <https://doi.org/10.1093/acprof>

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>

Plunkett, K., Sinha, C., Møller, M. F., & Strandsby, O. (1992). Symbol Grounding or the Emergence of Symbols? Vocabulary Growth in Children and a Connectionist Net.

*Connection Science*, 4(3–4), 293–312. <https://doi.org/10.1080/09540099208946620>

Rakison, D. H. (2004). Infants' sensitivity to correlations between static and dynamic features in a category context. *Journal of Experimental Child Psychology*, 89, 1–30.

<https://doi.org/10.1016/j.jecp.2004.06.001>

Rakison, D. H., & Poulin-Dubois, D. (2002). You Go This Way and I'll Go That Way:

Developmental Changes in Infants' Detection of Correlations among Static and Dynamic Features in Motion Events. *Child Development*, 73(3), 682–699.

<https://doi.org/10.1111/1467-8624.00432>

Robinson, C. W., & Sloutsky, V. M. (2004). *Auditory Dominance and Its Change in the*

*Course of Development*. *Child Development* (Vol. 75). <https://doi.org/10.1111/j.1467-8624.2004.00747.x>

Robinson, C. W., & Sloutsky, V. M. (2007a). Linguistic Labels and Categorization in

Infancy: Do Labels Facilitate or Hinder? *Infancy*, 11(3), 233–253.

<https://doi.org/10.1111/j.1532-7078.2007.tb00225.x>

- Robinson, C. W., & Sloutsky, V. M. (2007b). Visual processing speed: effects of auditory input on visual processing. *Developmental Science*, 10(6), 734–740.  
<https://doi.org/10.1111/j.1467-7687.2007.00627.x>
- Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition : a parallel distributed processing approach*. MIT Press.
- Slater, A. (1989). Visual memory and perception in early infancy. In *Infant development*. (pp. 43–71). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
- Sloutsky, V. M., & Napolitano, A. C. (2003). Is a Picture Worth a Thousand Words? Preference for Auditory Modality in Young Children. *Child Development*, 74(3), 822–833. <https://doi.org/10.1111/1467-8624.00570>
- Sloutsky, V. M., & Robinson, C. W. (2008). The Role of Words and Sounds in Infants' Visual Processing: From Overshadowing to Attentional Tuning. *Cognitive Science*, 32(2), 342–365. <https://doi.org/10.1080/03640210701863495>
- Wang, Q., Bolhuis, J., Rothkopf, C. A., Kolling, T., & Knopf, M. (2012). Infants in Control: Rapid Anticipation of Action Outcomes in a Gaze-Contingent Paradigm. *PLoS ONE*, 7(2), 30884. <https://doi.org/10.1371/journal.pone.0030884>
- Waxman, S. R., & Braun, I. (2005). Consistent (but not variable) names as invitations to form object categories: new evidence from 12-month-old infants. *Cognition*, 95(3), B59–B68.  
<https://doi.org/10.1016/j.cognition.2004.09.003>
- Waxman, S. R., & Gelman, S. A. (2010). Different kinds of concepts and different kinds of words: What words do for human cognition. In D. Mareschal, P. C. Quinn, & S. E. G. Lea (Eds.), *Oxford series in developmental cognitive neuroscience. The Making of Human Concepts* (pp. 99–129). Oxford University Press.  
<https://doi.org/10.1093/acprof:oso/9780199549221.001.0001>
- Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence

from 12- to 13-month-old infants. *Cognitive Psychology*, 29(3), 257–302.

<https://doi.org/10.1006/cogp.1995.1016>

Wronski, C., & Daum, M. M. (2014). Spatial orienting following directional cues in infancy.

*Developmental Psychology*, 50(8), 2020–2029. <https://doi.org/10.1037/a0037155>

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. (1983). Infant Perception of Correlations among Attributes.

*Child Development*, 54(4), 858–867. <https://doi.org/https://doi.org/10.2307/1129890>

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. (1985). The Segregation of Items into Categories by Ten- Month-Old Infants.

Source: *Child Development*, 56(6), 1574–1583. <https://doi.org/10.2307/1130476>

### Appendix

Table A1. *Stimuli structure used in three familiarisation blocks (each number represents a value on each of the four features).*

Block	Category	Antenna	Hands	Body	Legs
1	1	1	1	3	3
	1	1	3	1	3
	1	3	1	3	1
	1	3	3	1	1
	2	5	5	7	7
	2	5	7	5	7
	2	7	5	7	5
	2	7	7	5	5
2	1	1	1	1	3
	1	1	1	3	1
	1	3	3	3	1
	1	3	3	1	3
	2	5	5	7	5
	2	5	5	5	7
	2	7	7	5	7
	2	7	7	7	5
3	1	1	3	1	1
	1	1	3	3	3
	1	3	1	1	1
	1	3	1	3	3
	2	5	7	5	5
	2	5	7	7	7
	2	7	5	5	5
	2	7	5	7	7