Journal of Memory and Language 132 (2023) 104447

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

When time shifts the boundaries: Isolating the role of forgetting in children's changing category representations

Melina L. Knabe[a,*], Christina C. Schonberg[b], Haley A. Vlach[a]

[a] University of Wisconsin-Madison, USA
[b] IXL Learning, USA

ARTICLE INFO

Keywords:
Categorization
Forgetting
Memory
Generalization
Cognitive development

ABSTRACT

In studies of children's categorization, researchers have typically studied how encoding characteristics of exemplars contribute to children's generalization. However, it is unclear whether children's internal cognitive processes alone, independent of new information, may also influence their generalization. Thus, we examined the role that one cognitive process, forgetting, plays in shaping children's category representations by conducting three experiments. In the first two experiments, participants ($N_{Exp1} = 37$, $M_{age} = 4.02$ years; $N_{Exp2} = 32$, $M_{age} = 4.48$ years) saw a novel object labeled by the experimenter and then saw five new objects with between one and five features changed from the learned exemplar. The experimenter asked whether each object was a member of the same category as the exemplar; children saw the five new objects either immediately or after a 5-minute delay. Children endorsed category membership at higher rates at immediate test than at delayed test, suggesting that children's category representations became narrower over time. In Experiment 3, we investigated forgetting as a key mechanism underlying the narrowing found in Experiments 1 and 2. We showed participants ($N_{Exp3} = 34$, $M_{age} = 4.20$ years) the same exemplars used in Experiments 1 and 2; then, either immediately or after a 5-minute delay, we showed children seven individual object features and asked if each one had been part of the exemplar. Children's accuracy was lower after the delay, showing that they did indeed forget individual features. Taken together, these results show that forgetting plays an important role in changing children's newly-learned categories over time.

Introduction

Categorization is critical for cognition and development. Learning categories allows children to structure and understand the world, thus affording inferences to new experiences (Harnad, 2005; Horst & Simmering, 2015). A key process in successfully forming categories is *abstraction*—that is, determining the underlying features that category members do and do not have in common (e.g., Son et al., 2008; Vlach, 2014). Because category exemplars are rarely encountered *en masse*, children typically encode individual exemplars as they are encountered, and then abstract across those experiences to form categories (e.g., Diesendruck & Peretz, 2013; Graham et al., 2012; Lawson & Fisher, 2011). Prior research has examined how encoding new exemplars leads to changes in children's abstraction and inferences (e.g., Booth, 2014; Gelman & Markman, 1987; Goldenberg & Sandhofer, 2013; Graham et al., 2012; Lawson & Fisher, 2011; Perry et al., 2010; Sloutsky et al., 2007; Vlach et al., 2012; Vlach & Sandhofer, 2011). In the present study, we examined whether other factors—independent of experience with multiple category exemplars—contribute to changes in children's generalization. Specifically, we examined whether category representations naturally shift *over time*, independent of new learning, and whether this shift could be due to basic cognitive processes, such as forgetting.

Most research on children's categorization has examined how new learning affects category representations. For instance, researchers have examined how children's generalizations change as they encounter additional exemplars, depending on the complexity (e.g., Son et al., 2008), variability (e.g., Lawson & Fisher, 2011; Perry et al., 2010), and number (e.g., Ross et al., 1986; Twomey et al., 2014) of exemplars presented. Researchers have also examined how changes in task features can influence children's generalization (e.g., Flack & Horst, 2018; Gathercole et al., 1995; Samuelson et al., 2009; Samuelson & Smith, 2000; Schonberg et al., 2020; Yoshida & Smith, 2003). This work has revealed that children are sensitive to different types of

* Corresponding author at: UW-Madison Department of Educational Psychology, 1025 W Johnson, St. Madison, WI 53706, USA.
E-mail address: mknabe@wisc.edu (M.L. Knabe).

https://doi.org/10.1016/j.jml.2023.104447
Received 12 August 2022; Received in revised form 16 June 2023; Accepted 17 June 2023
Available online 3 July 2023

information—perceptual, conceptual, and causal—that are provided about exemplars (e.g., Booth, 2014; Booth & Waxman, 2002; Diesendruck & Peretz, 2013; Graham, Booth, & Waxman, 2012; Nelson, Frankenfield, Morris, & Blair, 2000; Opfer & Bulloch, 2007; Sloutsky, Kloos, & Fisher, 2007; Smith, Jones, Yoshida, & Colunga, 2003; Gelman and Markman, 1987). Finally, children's use of this information in generalization changes across infancy and early childhood (Fisher & Sloutsky, 2005; Hayes & Rehder, 2012; Sloutsky et al., 2007; see Fisher, 2015, for a review).

Researchers have also observed that category representations start out somewhat broad and become narrower and more refined with more category experience (e.g., Quinn et al., 2001; Smith & Kemler, 1977; Unger & Fisher, 2019). For example, 3–4 month-old infants show more differentiated categories than neonates; they can differentiate between shapes such as circles and triangles (Quinn et al., 2001). This increased differentiation is thought to be a result of infants' increasing experience with the tested categories (e.g., Quinn, 2004; Quinn et al., 2001; Quinn & Johnson, 2000; Quinn & Tanaka, 2007). Moreover, real-world learning experiences help children refine their category representations. Children who attended a zoo-based educational summer camp showed greater differentiation among animal categories than children who attended a control camp, even after accounting for children's pre-camp category representations (Unger & Fisher, 2019; see also Vales et al., 2020).

Because research to date has focused on how children's learning of additional information and/or exemplars changes abstraction, less is known about whether factors beyond additional learning experiences lead to shifts in generalization. We hypothesize that basic cognitive processes—independent of encountering new information about a category—can alter the way that children abstract and generalize. One internal mental process that plays an important role in cognition is forgetting. As time passes after exposure to information, memory for the details of that information begins to diminish according to a curvilinear pattern, often referred to as a forgetting curve (Cepeda et al., 2006; Ebbinghaus, 1885/1913). That is, forgetting causes shifts in learners' ability to retain and retrieve information over time. As most categories in the real world are learned via exposures to single instances spaced apart in time, it is important to understand the role that forgetting across time plays in shifting children's category representations.

Research suggests that forgetting of these features between exemplar presentations can affect children's categorization and generalization (e.g., Vlach et al., 2012, 2014; Vlach & Sandhofer, 2012; for a review, see Vlach, 2014; 2019). In these studies, researchers have manipulated the amount of time that elapses between presentation of exemplars during learning (Slone & Sandhofer, 2017; Vlach et al., 2008), as well as between learning and test (e.g., Perry et al., 2016), or both (Vlach et al., 2012). Because children rapidly forget new information, researchers commonly use short time delays (e.g., 5 min) to assess learning and forgetting across time (Horst & Samuelson, 2008; Karaman & Hay, 2018; Vlach et al., 2012). These studies have found that allowing children time to forget between multiple learning events is helpful for generalization. According to the *forgetting-as-abstraction* account (Vlach, 2014), forgetting supports generalization because it allows for category-irrelevant features to diminish in memory. At the same time, memory for category-relevant features is maintained because these features are encountered with each new exemplar, reactivating their representations and thus slowing forgetting. As time passes, irrelevant features are forgotten and relevant features are more likely to be remembered. Thus, when children are engaging in the process of abstraction and generalization, they are more likely to remember the relevant features for categorization.

Although the existing literature suggests that forgetting is a mechanism underlying children's categorization and generalization, the effects of forgetting and new learning remain enmeshed. That is, forgetting as a mechanism has always been studied in tandem with multiple instances of new learning. Thus, the goal—and novel theoretical contribution—of the present study is to examine whether forgetting alone leads to changes in category representations in the absence of new learning. Indeed, this work serves to build previous theories, such as forgetting-as-abstraction theory, in a new direction.

Given what we know about how forgetting shapes learning, there are two competing possibilities for how forgetting might shift categorization. On the one hand, as time passes after an exposure to an exemplar, category representations could become broader. That is, as individual details of an exemplar are forgotten, the category representation becomes less specific and able to accommodate a wider variety of items. Indeed, there is evidence that across early development, generalization becomes more flexible with the passage of time (e.g., Borovsky & Rovee-Collier, 1990; see also Barr & Brito, 2013; McGaugh, 2000).

On the other hand, as time passes after exposure to an exemplar, category representations might become narrower. If some details of the exemplar are retained but most details are forgotten, a learner may not recognize a subsequent exemplar as a category member if it contains only the forgotten details. If a learner is relying on the presence of a few retained details when deciding how to categorize new exemplars, they may generalize more narrowly, only identifying objects that contain those specific details as category members. For example, in one study (Wojcik, 2017), 2-year-olds' category representations became narrower with time: they were less likely to generalize a category label to a novel exemplar after a 1-week delay relative to a 1-minute delay. This research converges with findings from other studies in which children were presented with additional learning events; the pattern of broad-to-narrow is commonly found in research examining shifts in children's boundaries (e.g., Quinn et al., 2001; Quinn & Johnson, 2000; Unger & Fisher, 2019; Vales et al., 2020).

Thus, in the present study, we investigated whether children's forgetting of exemplar features led to shifts in their subsequent generalizations. In Experiments 1 and 2, we asked children to categorize novel objects either immediately after exposure to an exemplar or after a 5-minute delay. We hypothesized that children's generalizations would become narrower after a delay, due to their forgetting of the initial exemplar. In Experiment 3, we examined whether forgetting of exemplar features occurred during the 5-minute delay. If so, this would provide evidence that forgetting was the mechanism underlying our findings from Experiments 1 and 2.

Experiment 1

Data Availability

All stimuli, data, and analysis code for the experiments are available at https://osf.io/h5wz9/.

Method

Participants. Three- to five-year-old children ($N = 37$, $M_{age} = 4.02$ years, range = 3.15–5.27 years, 18 girls, 19 boys) were recruited from preschools and childcare centers in a Midwestern US city to participate in this study. To determine a sample size, effect sizes were gathered from previous studies examining children's generalization at various time delays ($ds > 0.55$, $\eta_p^2 > .396$; e.g., Slone & Sandhofer, 2017; Vlach et al., 2012). A power analysis for a linear mixed effects model that includes two main effects, as well as $\alpha = .05$ and $d = 0.55$, yielded a sample size of 24 participants to achieve 80% power (Judd et al., 2016; Westfall et al., 2014). We chose this developmental period because children's memory and generalization change rapidly during the preschool years (e.g., Diesendruck & Peretz, 2013; Gathercole, 1998; Huang-Pollock et al., 2011; Lawson & Fisher, 2011), which means there is a high likelihood of observing variability across participants. The sample was 16.2% Asian/Asian-American, 75.6% White, and 8.1% multiracial; 5.4% of the sample was Hispanic, and 94.6% was not Hispanic. 97.3% of participants had at least one parent who had completed a 4-year college degree or higher. Children received a book for participating in the study. An additional 2 children participated in the study but were removed from the final sample for being off task.

Materials and Stimuli. Twelve sets of images of novel object categories (six animate, six inanimate) were created using Microsoft PowerPoint (see Fig. 1). Novel objects and words were used to minimize the role of prior experience and to model children's categorization of newly acquired information (e.g., a baby hearing the word "dog" during the first encounter with a dog). Animate and inanimate object categories were modeled after stimuli used in previous research on categorization (e.g., Horst & Hout, 2016; Sloutsky et al., 2007; Sumner, DeAngelis, Hyatt, Goodman, & Kidd, 2019), and animacy was denoted through facial features (e.g., eyes, mouth). Each category consisted of one multicolored exemplar with shape-based features and five test items. Each test item had one, two, three, four, or five shape-based feature changes relative to the exemplar. The colors used in the stimuli changed randomly across items so that shape would be the only reliably-changing feature. Five features were selected randomly to change, in a random

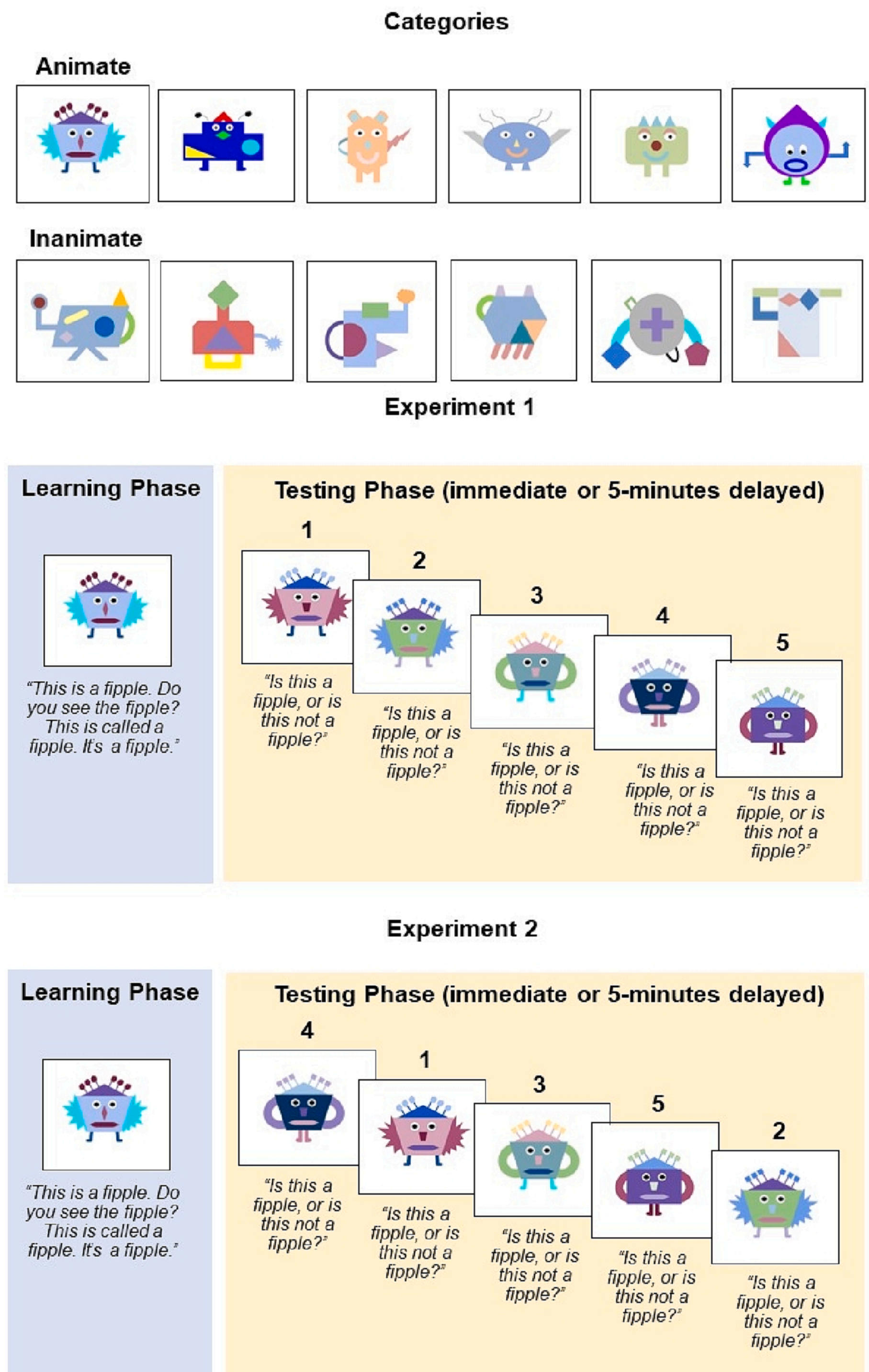

Fig. 1. *Stimuli and experimental procedure for Experiments 1 and 2, Note.* Twelve total categories (6 animate, 6 inanimate) were assigned novel names. Children saw one exemplar during the learning phase and were shown 5 test items during the testing phase (either immediately or after a 5-minute delay). In Experiment 1, the number of feature changes increased sequentially from 1 to 5. In Experiment 2, the number of feature changes was random on each trial.

order, across test items. For the animate categories, the eyes or mouth never changed, as they are central cues to animacy (e.g., Anderson, Meagher, Welder, & Graham, 2018; Looser & Wheatley, 2010). Feature changes were additive, such that items with more than one feature change retained the changed features of the preceding items. For example, if the exemplar had a circle and the one-feature-change object had a triangle in its place, the rest of the feature-change objects also contained a triangle. Visual stimuli were processed and presented in a Microsoft Office PowerPoint presentation on an iPad (height: 11.03 in., width: 8.46 in., display: 12.9 in.).

Each novel object category was paired with a novel label (e.g., "boskot," "fipple") from the Novel Object and Unusual Name (NOUN) database (Horst & Hout, 2016). Two sets of category-label pairs and category orders were used; in each set, labels were randomly paired with objects, and object categories were randomly ordered.

Design. This experiment used a within-subjects design; children participated in both an immediate test phase and a 5-minute delayed test phase. The six categories that were seen in the immediate and delayed conditions were different. Independent variables of interest included condition (immediate vs. delayed), the number of feature changes from the original training exemplar (1–5), and trial order. The dependent variable of interest was whether participants endorsed a test trial as belonging to the exemplar category ("Yes" = 1, "No" = 0). All analyses were conducted in R (version 3.5.2) using the packages *lme4* (Bates et al., 2015), *lmerTest* (Kuznetsova et al., 2017), *emmeans* (Lenth, 2020), and *MuMIn* (Bartoń, 2020).

Procedure. Children were tested alone in a quiet room or a quiet corner of their classroom. There were two phases of the experiment: (1) a familiarization phase and (2) a learning and testing phase.

Familiarization phase. Children were first trained in the experimental procedure using familiar objects. Using an iPad to display the images, the experimenter showed a picture of a car for 10 s, labeling it four times (e.g., "This is a car! Do you see the car? This is called a car. It's a car."). Then, the experimenter said "Now I'm going to show you more pictures and ask if they are cars. If it is a car, I want you to say 'yes, it's a car.' If it's not a car, I want you to say 'no, it's not a car.' The experimenter then showed five images, one at a time, asking "Is this a car, or is this not a car?" for each one. To teach children that "no" was an acceptable response in this paradigm, three images were cars and two images were not cars. In the case of an incorrect response, children were corrected and given the opportunity to try again until they were correct.

Learning and testing phase. After the familiarization phase, children were taught the names of 12 novel object categories (see Fig. 1). For each category, the exemplar was shown for 10 s and labeled four times by the experimenter. Only a single exemplar was shown before the test items because we were interested in whether time alone, absent additional learning, affected children's category boundaries. Test items were shown in order of the number of features changed—that is, the first item had one feature changed from the exemplar and the fifth item had five features changed from the exemplar. For each test item, the experimenter asked the child "Is this a [label], or is this not a [label]?" and recorded the child's response. If the child responded "I don't know", their response was coded as a failure to endorse the test item as part of the category.

Categories were either tested immediately, or after a 5-minute delay to determine how the passage of time affects children's categorization. For *immediate test* categories, five test items were shown immediately after the exemplar. For *delayed test* categories, there was a 5-minute delay between when children were shown the exemplar and when they were shown the five test items; during the delay, children played with Play-Doh. A 5-minute delay has been shown to be sufficient for children to forget during category-learning tasks (e.g., Vlach et al., 2008). Two presentation orders were created by randomly assigning three animate and three inanimate objects to the immediate condition and three animate and three inanimate objects to the delayed condition. Children were randomly assigned to one of the two orders to counterbalance item and condition assignments across children. In both presentation orders, children were presented with the immediate test categories first and the delayed categories second.

The entire experiment took approximately 15 min to complete.

Results

We were interested in whether children's categorization of new objects would change after a delay. Specifically, we hypothesized that children would be less likely to endorse an item as part of the category after the 5-minute delay. Therefore, a logistic generalized linear model was constructed to assess the extent to which condition and number of feature changes from the first category exemplar affected children's likelihood—or log odds—of endorsing an exemplar as part of the category ("Yes" = 1 or "No" = 0). The maximal model was as follows: glmer (Endorsement ~ Condition + Number of Feature Changes + (1 + Condition + Number of Feature Changes | Participant). The fixed effect of condition was contrast coded ("Immediate" = −0.5, "Delayed" = 0.5). The fixed effect of number of feature changes was repeated coded, such that each feature change was compared to the prior feature change (2 vs. 1 changes, 3 vs. 2 changes, 4 vs. 3 changes, 5 vs. 4 changes; see Schad et al., 2020 for a review of contrast coding). We used this approach because we predicted that learners would compare each new item with the previous item to determine the category boundary. That is, if children endorsed an item as a "wug" after three feature changes, they might compare a new item with four feature changes to the previous one to determine whether the new item is also a "wug".

We first constructed a model with a by-subject random intercept only and added either a by-subject random slope for condition, a by-subject random slope for item, or a by-subject random slope for condition and item. We then assessed how model fit changed using pairwise likelihood ratio tests (see Supplementary Material for model output). The model fit was improved with the addition of a by-subject random slope for condition, $\chi^2(2) = 83.91$, $p < .001$, in comparison to a by-subject intercept only model. Similarly, model fit was improved with the addition of a by-subject random slope for item, $\chi^2(14) = 34.83$, $p = 0.002$. However, the best model fit included a by-subject random slope for item and condition, $\chi^2(18) = 34.44$, $p = 0.01$. The addition of age did not significantly improve fit of the final model $\chi^2(1) = 2.34$, $p = 0.13$, and did not change the pattern of the results. Thus, we report the model without participant age below.

The logistic effects model revealed a significant main effect of condition, $B = -1.02$, $SE = 0.50$, $z = -2.03$, $p = 0.04$; $OR = 0.36$, 95% $CI = 0.13–0.97$, such that children were 64% less likely to endorse a category member in the delayed condition than the immediate condition

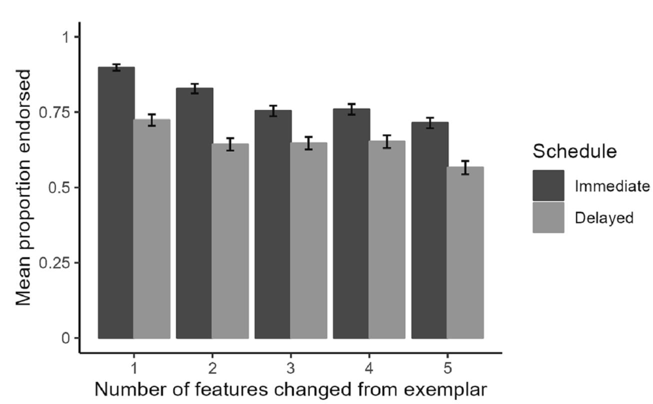

Fig. 2. Children's mean category endorsements in Experiment 1, *Note.* Mean category endorsements across number of feature changes for the immediate and delayed condition in Experiment 1. Generalized linear mixed effects models revealed a significant main effect of condition, such that children were less likely to endorse a category exemplar after the 5-minute delay.

($M_{immediate}$ =.79, *SE* =.053, $M_{delayed}$ =.61, *SE* =.053; see Fig. 2 for mean category endorsements across condition and number of feature changes). There was also a significant main effect of number of feature changes: Participants were 74% less likely to endorse the exemplar when it had 5 vs. 4 feature changes from the initial category exemplar, B = -1.33, SE = 0.41, z = -3.28, p =.001; OR = 0.26, 95% CI = 0.12–0.59 ($M_{\text{FourFeatures}}$ =.68, SE =.056; $M_{\text{FiveFeatures}}$ =.62, SE =.056).

Taken together, these results show that children are sensitive to objects' feature changes and are less likely to consider an object as a category member if it shares fewer features in common with the initial exemplar. Furthermore, these findings suggest that children's category representations became narrower after they were given time to forget the exemplar during the delay: Children were less likely to endorse objects' category membership when they were tested after a delay compared to when they were tested immediately.

Experiment 2

The results in Experiment 1 suggest that children's generalizations of category boundaries become narrower over time, such that children endorsed category membership for a narrower range of objects when they were tested after a 5-minute delay as opposed to immediately after exposure to the exemplars. However, in Experiment 1, feature changes were always presented in order (i.e., the first test item always had one changed feature and the fifth test item always had five changes). Therefore, it is possible that children's generalizations were changing as a function of how many exemplars within a category they had already judged—perhaps they were more likely to say "no" in later trials due to fatigue or boredom, rather than due to a real decision about an item's category membership. Experiment 2 addressed this possibility by randomizing the order of the test trials for each category. That is, the number of features that changed were no longer presented sequentially as in Experiment 1.

Method

Participants. Three- to five-year-old children were recruited from preschools and childcare centers in a Midwestern US city to participate in this study (N = 32, M_{age} = 4.48 years, range = 3.13 – 5.53 years, 18 girls, 14 boys). Because this experiment was identical in design to Experiment 1, we used the same power analysis to determine sample size and the same rationale for the age range of participants. The sample was 6.5% Asian/Asian-American, 74.2% White, and 16.1% multiracial; 3.4% of the sample was Hispanic, and 96.6% was not Hispanic. 93.4% of participants had at least one parent who had completed a 4-year college degree or higher. Children were recruited and compensated in the same way as described in Experiment 1. An additional 3 children participated in the study but were removed from the final sample for being off task.

Materials, Stimuli, and Design. Same as Experiment 1.

Procedure. The procedure was the same as that of Experiment 1, except that test items were shown in a random order (as opposed to being ordered by the number of features changed from the original exemplar).

Results

In this experiment, the test stimuli were randomly ordered to determine whether children's category representations became narrower due to the pattern in the changing object features in Experiment 1 or due to the passage of time. We conducted the same analyses as in Experiment 1: a logistic generalized linear model predicting children's likelihood (log odds) of endorsing an exemplar as part of the category ("Yes" = 1 or "No" = 0). The maximal model was as follows: glmer (Endorsement ~ Condition + Number of Feature Changes + (1 + Condition + Number of Feature Changes | Participant). Like Experiment 1, the fixed effect of condition was contrast coded ("Immediate" = −0.5, "Delayed" = 0.5) and the fixed effect of number of feature changes was treatment coded (1 feature change serving as the reference group). We chose to treatment code the number of feature changes because, unlike Experiment 1, items were not presented in order of the number of features changed. Instead of comparing each item to the one they saw previously, children likely compared each new item to the item that most closely resembled the original category exemplar (i.e., the item with 1 feature change).

We first constructed a model with a by-subject random intercept only and added a by-subject random slope for condition, a by-subject random slope for item, or a by-subject random slope for condition and item. We again assessed how model fit changed using pairwise likelihood ratio tests (see Supplementary Material for model output). The model fit was improved with the addition of a by-subject random slope for condition, $\chi^2(2) = 81.83$, $p < .001$. However, model fit was not improved with the addition of a by-subject random slope for item, $\chi^2(14) = 10.23$, $p = 0.74$. Next, we added by-subject random slopes for item and condition, which did not yield a better model fit than including a by-subject random slope for condition only, $\chi^2(18) = 18.79$, $p = 0.40$. Although model fit was not improved significantly, we used a "keep it maximal" approach (Barr et al., 2013) and included the model with the maximum random effects structure. The addition of age did not significantly improve fit of the final model, $\chi^2(1) = 1.35$, $p = 0.24$, and did not change the pattern of the results. Thus, we report the model without participant age below.

As in Experiment 1, the logistic effects model revealed a significant main effect of condition, B = -0.92, SE = 0.35, z = -2.62, p = 0.008; OR = 0.39, 95% CI = 0.20–0.79, such that children were 61% less likely to endorse a category member in the delayed condition than the immediate condition ($M_{immediate}$ =.74, SE =.048; $M_{delayed}$ =.62, SE =.049; see Fig. 3 for mean category endorsements across condition and number of feature changes). There was also a significant main effect of number of feature changes: Participants were 67% less likely to endorse the exemplar when it had 5 feature changes vs. 1 feature change, even when the order of the feature-changed objects was randomized, B = -1.10, SE = 0.26, z = -4.33, p <.001; OR = 0.33, 95% CI = 0.20–0.54 ($M_{\text{OneFeature}}$ = .75, SE = .052; $M_{\text{FiveFeatures}}$ =.60, SE =. 052).

In sum, these results revealed that children's category generalizations became narrower across time. Furthermore, children were sensitive to the changing features of the objects and were not simply less likely to endorse category membership as they progressed through the test phase.

Experiment 3

In Experiments 1 and 2, we found that children's category

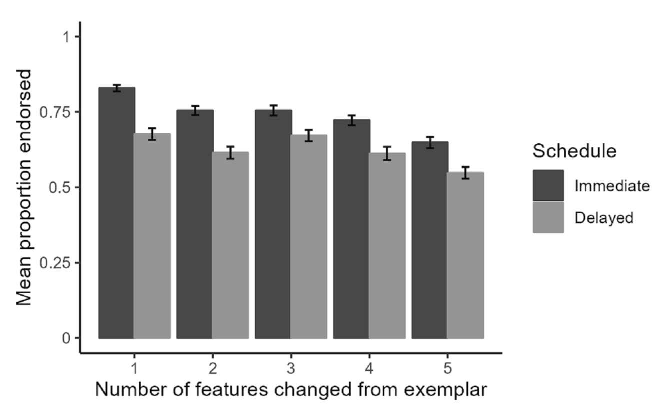

Fig. 3. Children's mean category endorsements in Experiment 2, *Note. Mean category endorsements across number of feature changes for the immediate and delayed condition in Experiment 2. Generalized linear mixed effects models revealed a significant main effect of condition, such that children were less likely to endorse a category exemplar after the 5-minute delay.*

generalizations become narrower over time. Why? We hypothesized that this is due to children forgetting the features of the original exemplar for each category across the 5-minute delay. That is, as children forget more features from the original exemplar, they have fewer features in memory that can be compared to features of a new exemplar. With fewer features available to compare, the likelihood that these features will overlap with those of a new exemplar decreases, leading them to reject new exemplars as category members. Experiment 3 was designed to determine whether children forget exemplar features between the immediate and 5-minute delayed tests. If so, this would be evidence for our hypothesis that children's forgetting of exemplar features leads them to generalize more narrowly.

Method

Participants. Three- to five-year-old children were recruited from preschools and childcare centers in a Midwestern US city to participate in this study (N = 34, M_{age} = 4.20 years, range = 3.00–5.14 years, 17 girls, 17 boys). We conducted a power analysis using the R package *pwr* (Champely, 2020) to estimate a sample size for detecting a hypothesized medium effect (Cohen's d =.5) at 80% power, which showed that a sample size of 34 would be sufficient given the design of this experiment. Because the goal was to investigate forgetting as a potential mechanism underlying children's behavior in Experiments 1 and 2, the same age range was recruited for this experiment. The sample was 3.0% Native American/Alaska Native, 11.8% Asian/Asian-American, 70.6% White, and 14.7% multiracial; 9.1% of the sample was Hispanic, and 90.9% was not Hispanic. 94.1% of participants had at least one parent who had completed a 4-year college degree or higher. Children were recruited and compensated in the same way as described in Experiment 1. An additional 6 children participated in the study but were removed from the final sample for being off task.

Materials and Stimuli. In this experiment, we used the same 12 category exemplars as in Experiment 1. To construct the test trials, four individual features were selected from each exemplar and inserted onto separate PowerPoint slides (see Fig. 4). These features maintained their color, shape, and size from the exemplar, but were centered on the screen for consistency across trials. In addition, three PowerPoint slides were created that contained one feature that was *not* part of the category exemplar. These distractor items (i.e., lures) allowed us to analyze children's incorrect responses at test in more depth. The same category labels from Experiments 1 and 2 were used, and the stimuli sets were constructed in the same manner (i.e., creating two sets of randomized category orders and category-label pairs).

Procedure. Children were tested alone in a quiet room or a quiet corner of their classroom. There were two phases of the experiment: (1) a familiarization phase and (2) a learning and testing phase.

Familiarization phase. The familiarization phase was the same as Experiments 1 and 2, with the following exception: after showing the child the picture of the car, the experimenter displayed seven object features, one at a time, and asked the child to identify whether each one had been part of the car ("Was this part of the car, or was this not part of the car?"). Four of the seven images were features that had been part of the exemplar car. In the case of an incorrect response, children were corrected. The experimenter recorded the number of incorrect responses made by each child during this phase.

Learning and testing phase. The learning and testing phases were the same as Experiments 1 and 2, with the following exception: at test, children viewed seven test trials. For each test trial, the experimenter asked the child "Was this part of the *[label]* or was it not part of the *[label]*?" and recorded the child's response. All children were presented with objects and features in the same order. The entire experiment took approximately 15 min to complete.

Results

We hypothesized that children in Experiments 1 and 2 showed different generalization patterns at immediate and delayed test because they had forgotten individual object features. To test this hypothesis, we examined how children's feature memory changed over time. We first compared children's overall accuracy in identifying whether a feature was present or absent in the exemplar between the immediate and delayed tests (see Fig. 5). Indeed, children were significantly more accurate when tested immediately (M = 25.59, SD = 4.04) than after a 5-minute delay (M = 20.92, SD = 3.69), $t(36)$ = 6.582, p <.001, Cohen's d = 1.08. Furthermore, children's accuracy was significantly higher than chance (21/42 trials correct) at the immediate test, $t(36)$ = 6.993, p <.001, Cohen's d = 1.15, but was not significantly different from chance at the delayed test, $t(36)$ = 1.009, p =.319, Cohen's d = 0.17.

In addition to examining children's overall accuracy, we were interested in how their specific patterns of hits (correctly saying that a feature had been part of the exemplar), correct rejections (correctly saying that a feature had not been part of the exemplar), false alarms (incorrectly saying that a feature had been part of the exemplar), and

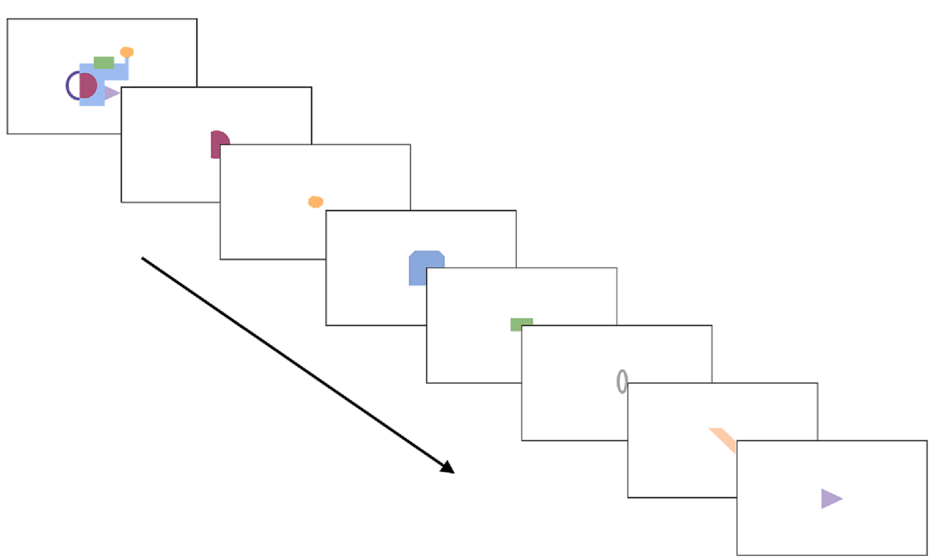

Fig. 4. Example of one category from Experiment 3 (exemplar and seven feature-memory trials), *Note. For this selected category, trials 1, 2, 4, and 7 contain features from the exemplar; the other trials consist of lures that were not part of the exemplar.*

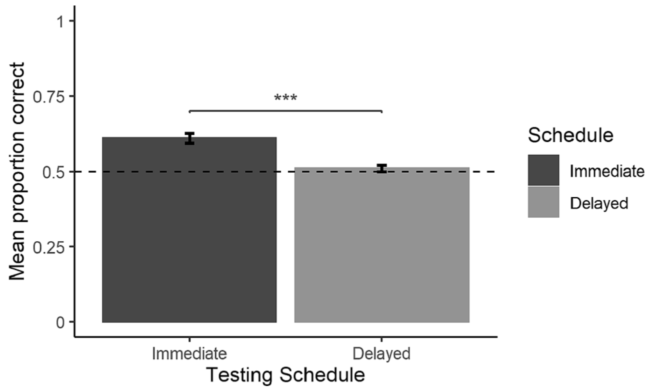

Fig. 5. Children's performance on memory tests for object features at immediate and delayed test, *Note. *** p <.001. Error bars represent standard errors.*

misses (incorrectly saying that a feature had not been part of the exemplar) changed over time. If children forgot features that had been present, we would expect a decrease in hits and a corresponding increase in misses after a delay. Because these data were nonnormally distributed—particularly false alarms, which were right-skewed, and correct rejections, which were left-skewed—we conducted Wilcoxon signed rank tests. These results revealed that children had significantly more hits at the immediate test and had significantly more misses at the delayed test (see Table 1). Finally, we examined whether there were age effects by conducting a regression with overall accuracy as the outcome variable and testing condition and participant age as predictor variables; testing condition significantly predicted performance (β = -0.10, SE = 0.02, t = -5.43, p <.001) but participant age did not (β = 0.03, SE = 0.02, t = 1.95, p = 0.06).

In brief, we found that children responded with more hits at the immediate test than delayed test, and more misses at the delayed test than immediate test; rates of correct rejections and false alarms did not significantly differ across testing conditions. This pattern of hits and misses shows that children forgot more object features after the delay.

Discussion

We investigated whether the passage of time alone, after one learning event, shifted children's category boundaries. Strikingly, the results of these experiments showed that new learning about the category did not need to occur in order for category representations to change: children made different judgments about objects' category membership depending on whether they were asked immediately after viewing an exemplar, or after a 5-minute delay. Specifically, children's category boundaries become narrower across time. We hypothesized that children's forgetting of specific object features might be responsible for this narrowing, and a follow-up study confirmed that children did forget individual features across a short time delay.

Why did categories become narrower, rather than broader, over time? Our hypothesis is that narrowing is due to children's forgetting of object features during the delay. For instance, suppose that a child sees an exemplar that has wavy arms and a triangle and circle on its body. Over time, the child forgets some of these features, perhaps remembering only the triangle. If they are presented with a new object that has wavy arms and a circle on its body, they might reject this object as a category member because it lacks a triangle—the only feature they remember—even though the other features overlap with the exemplar. More generally, when children remember only a subset of an exemplar's features and these features are absent from the tested objects, the child may think that the test object did not share any features with the exemplar. This apparent lack of overlapping features may lead them to reject new objects as category members. In Experiment 3, we found that children reliably forgot specific object features after a delay. Thus, the forgetting of object features is one likely mechanism underlying children's category narrowing.

These findings provide further support for theoretical frameworks implicating forgetting as a key mechanism in children's categorization and generalization. For instance, according to the forgetting-as-abstraction theory (Vlach, 2014), forgetting that occurs between category exemplar presentations facilitates children's ability to remember relevant features of categories. The current work adds to this theory by expanding the scope at which forgetting contributes to children's categorization. That is, this research is the first to show that children's general forgetting of category exemplars, absent of additional learning events, may facilitate abstraction by narrowing category boundaries. Although we observed that forgetting led to narrower category boundaries in this research, it is possible that forgetting could lead to differing behavior across timescales. For instance, over extended periods of time, children might view a new category member as an entirely new category because they forgot all features that are diagnostic for the category. That is, too much forgetting might deter children's generalization and/or cause categorization errors. Future research should examine children's categorization and generalization over longer timescales, such as days, weeks, etc., to test these possibilities.

These findings also raise theoretical questions about research paradigms used to test children's categorization. If time and forgetting leads to more narrow generalization, this makes research paradigms that introduce both additional learning events and opportunities to forget more difficult to interpret. That is, the extent to which new learning and forgetting each influenced changes in children's category representations may be unclear. Consequently, it is possible that forgetting alone could explain results from prior studies. For instance, forgetting might explain why children show broader categorization behavior when trained on a large sample of exemplars or on exemplars that are highly variable (e.g., Lawson & Fisher, 2011; Perry et al., 2010). Due to children's memory constraints, seeing many exemplars may result in more forgetting, leading to changes in categorization behavior. Future work should therefore incorporate controls to isolate the influence of forgetting and new learning, such as a forgetting-only baseline where no new learning is introduced.

A next step for researchers is to examine how additional cognitive processes lead to shifts in children's generalization of categories. For instance, in addition to forgetting, consolidation may be a critical mental process that underlies shifts from narrow to broad categories. According to Fuzzy Trace Theory (Brainerd & Reyna, 2004), learners encode verbatim and gist memory traces in parallel, but these traces undergo different forgetting rates due to consolidation processes. Verbatim traces refer to precise memories for the surface form (or exact details of the information) and are forgotten more quickly. In contrast,

Table 1
Children's test trial responses in Experiment 3.

Outcome	Immediate			Delayed			V	p
	M	SD	Median	M	SD	Median		
Hits	12.95	5.49	13	9.46	7.39	8	479	<.001***
Correct Rejections	12.65	4.77	14	11.46	6.03	13	279.5	.183
False Alarms	5.30	4.77	4	6.11	6.01	5	181.5	.442
Misses	11.05	5.49	11	13.92	7.30	15	118	.004**

gist traces refer to less precise memories for background information and are forgotten more slowly. Due to the higher rates of forgetting for verbatim traces, they are less likely to survive the consolidation process. It is possible that in our study, children's category judgments were based primarily on verbatim traces—that is, memory for each exemplar's specific features—because their representations had only just begun to consolidate. We predict that a longer delay between learning and testing would result in the persistence of primarily gist-based memory traces, which could in turn lead to broader category representations. Therefore, future research should investigate whether consolidation is a mechanism that shifts category representations from narrow to broad.

In addition, researchers should seek to identify when internal mental processes may play a minimal role. In the experiments presented here, we focused on memory and generalization during category *acquisition*, using only items that were new to children. For newly-learned categories, relevant and irrelevant features are likely to be forgotten at similar rates because the learner has not had enough experience to strongly differentiate them; thus, forgetting plays an important role in shaping new category representations. We predict that processes such as forgetting may play a smaller role in shifting the boundaries of established categories. For these types of categories, key features likely already have priority in memory, and so the forgetting rate for these features should be slow—meaning that forgetting should not shift category boundaries in the same way in familiar as opposed to novel categories.

This work also has implications for real-world categorization contexts, such as a young child learning about the category 'dog'. When children acquire new categories, their endorsement of new objects as part of that category likely changes across short intervals of time; this shift in category boundary may be explained by children's memory for key features. For example, a child may call all four-legged creatures a 'dog' in the morning, but then only use 'dog' to refer to the golden retriever at home later that night. This narrower category boundary may be due to the child's forgetting of key features of dogs, such as their general body shape and number of legs, and/or stronger memory for other features, such as a caregiver referring to this animal at home as a 'dog'. Indeed, forgetting of key features over time is likely one explanation for why we observe dynamic category representations early in development.

One limitation of this study is that we do not know what caused forgetting. Children saw all category exemplars before the 5-minute break and delayed test (see Perry et al., 2016; Slone & Sandhofer, 2017; Vlach et al., 2012 for similar designs). Was it the passage of time alone that caused children's forgetting? Or was it both the passage of time and interference from other, non-task relevant, learning? This is an open question in memory research, as one can never truly isolate time from new learning and vice versa (for discussions of this issue, see Altmann & Gray, 2002). Indeed, the field still does not know why forgetting occurs in any situation (see decay vs. interference debate; Altmann & Schunn, 2012; Anderson, 2003; Portrat, Barrouillet, & Camos, 2008; Darby & Sloutsky, 2015; Kail, 2002). In brief, the results from Experiment 3 demonstrate that forgetting occurred in this paradigm, but we will need to overcome significant theoretical and methodological barriers to determine whether it is time, interference, or both that contribute to forgetting.

In sum, the present study provides evidence that children's categorization does not shift solely because they learn new information. There are basic cognitive processes, such as forgetting, that shape children's category representations. Thus, category learning – a central aspect of cognitive development – emerges from a dynamic interaction between basic cognitive processes of learners and the learning events encountered in the environment. It is therefore critical for future research to examine how forgetting, as well as other internal processes (e.g., consolidation, retrieval, prediction), interact to shape young children's categorization. Because these cognitive processes underlie learning of other categories (e.g., social categories), as well as learning in other domains, this line of work will reveal how children make sense of the world around them.

CRediT authorship contribution statement

Melina L. Knabe: Formal analysis, Visualization, Writing – review & editing. **Christina C. Schonberg:** Conceptualization, Data curation, Methodology, Investigation. **Haley A. Vlach:** Funding acquisition, Resources, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data Availability Statement: All raw data, code, analyses, and stimuli are available at: https://osf.io/h5wz9/.

Acknowledgements

We thank the research assistants and members of the UW-Madison Learning, Cognition, and Development lab for their assistance with this project. We are especially thankful to our participants and their families for their contributions to this research.

Funding

This research was funded by NIH R21HD092867-01 to H. Vlach & G. Lupyan, the James S. McDonnell Foundation Human Cognition Scholar Award to H. Vlach, and NIH T32 DC05359 to S. Ellis-Weismer (supported C. Schonberg).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jml.2023.104447.

References

Altmann, E. M., & Gray, W. D. (2002). Forgetting to remember: The functional relationship of decay and interference. *Psychological Science, 13*(1), 27–33. https://doi.org/10.1111/1467-9280.00405

Altmann, E. M., & Schunn, C. D. (2012). Decay versus interference: A new look at an old interaction. *Psychological Science, 23*(11), 1435–1437. https://doi.org/10.1177/0956797612446027

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language, 49*(4), 415–445. https://doi.org/10.1016/j.jml.2003.08.006

Anderson, N., Meagher, K., Welder, A., & Graham, S. A. (2018). Animacy cues facilitate 10-month-olds' categorization of novel objects with similar insides. *PLoS ONE, 13*(11), e0207800.

Barr, R., & Brito, N. (2013). From Specificity to Flexibility: Early Developmental Changes in Memory Generalization. In P. J. Bauer, & R. Fivush (Eds.), *The Wiley Handbook on the Development of Children's Memory* (pp. 453–479). John Wiley & Sons Ltd.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal.* https://doi.org/10.1002/9781118597705.ch20

Bartoń, K. (2020). *MuMIn: Multi-Model Inference (1.43.17)*. [Computer software].

Booth, A. E. (2014). Conceptually coherent categories support label-based inductive generalization in preschoolers. *Journal of Experimental Child Psychology, 123*, 1–14. https://doi.org/10.1016/j.jecp.2014.01.007

Booth, A. E., & Waxman, S. R. (2002). Word learning is 'smart': Evidence that conceptual information affects preschoolers' extension of novel words. *Cognition, 84*(1), B11–B22. https://doi.org/10.1016/S0010-0277(02)00015-X

Borovsky, D., & Rovee-Collier, C. (1990). Contextual Constraints on Memory Retrieval at Six Months. *Child Development, 61*(5), 1569–1583. https://doi.org/10.1111/j.1467-8624.1990.tb02884.x

Brainerd, C. J., & Reyna, V. F. (2004). Fuzzy-trace theory and memory development. *Developmental Review, 24*(4), 396–439. https://doi.org/10.1016/j.dr.2004.08.005
Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132* (3), 354–380. https://doi.org/10.1037/0033-2909.132.3.354
Champely, S. (2020). *pwr: Basic Functions for Power Analysis* (1.3-0) [R]. https://CRAN.R-project.org/package=pwr.
Darby, K. P., & Sloutsky, V. M. (2015 Apr). The cost of learning: Interference effects in memory development. *J Exp Psychol Gen., 144*(2), 410–431. https://doi.org/10.1037/xge0000051
Diesendruck, G., & Peretz, S. (2013). Domain differences in the weights of perceptual and conceptual information in children's categorization. *Developmental Psychology, 49* (12), 2383–2395. https://doi.org/10.1037/a0032049
Ebbinghaus, H. (1885/1913). Memory: a contribution to experimental psychology. Translated by H.A. Ruger and C.E. Bussenius. In *German, Ueber das gedächtnis: Untersuchen zur experimentellen psychologie)*. New York: Teachers College, Columbia University.
Fisher, A. V. (2015). Development of Inductive Generalization. *Child Development Perspectives, 9*(3), 172–177. https://doi.org/10.1111/cdep.12127
Fisher, A. V., & Sloutsky, V. M. (2005). When Induction Meets Memory: Evidence for Gradual Transition From Similarity-Based to Category-Based Induction. *Child Development, 76*(3), 583–597.
Flack, Z. M., & Horst, J. S. (2018). Two sides to every story: Children learn words better from one storybook page at a time. *Infant and Child Development, 27*(1), e2047.
Gathercole, S. E. (1998). The development of memory. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 39*(1), 3–27. https://doi.org/10.1017/S0021963097001753
Gelman, S. A., & Markman, E. M. (1987). Young Children's Inductions from Natural Kinds: The Role of Categories and Appearances. *Child Development*, *58*(6), 1532–1541. JSTOR. https://doi.org/10.2307/1130693.
Goldenberg, E. R., & Sandhofer, C. M. (2013). Same, varied, or both? Contextual support aids young children in generalizing category labels. *Journal of Experimental Child Psychology, 115*(1), 150–162. https://doi.org/10.1016/j.jecp.2012.11.011
Graham, S. A., Booth, A. E., & Waxman, S. R. (2012). Words Are Not Merely Features: Only Consistently Applied Nouns Guide 4-year-olds' Inferences About Object Categories. *Language Learning and Development, 8*(2), 136–145. https://doi.org/10.1080/15475441.2011.599304
Harnad, S. (2005). Cognition is categorization. In C. Lefebvre, & H. Cohen (Eds.), *Handbook of categorization*. Elsevier.
Hayes, B. K., & Rehder, B. (2012). The Development of Causal Categorization. *Cognitive Science, 36*(6), 1102–1128. https://doi.org/10.1111/j.1551-6709.2012.01244.x
Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behavior Research Methods, 48*(4), 1393–1409. https://doi.org/10.3758/s13428-015-0647-3
Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, 13(2), 128-157. https://doi.org/10.1080/15250000701795598.
Horst, J. S., & Simmering, V. R. (2015). Category learning in a dynamic world. *Frontiers in Psychology, 6*, 46. https://doi.org/10.3389/fpsyg.2015.00046
Huang-Pollock, C. L., Maddox, W. T., & Karalunas, S. L. (2011). Development of implicit and explicit category learning. *Journal of experimental child psychology, 109*, 321–335. https://doi.org/10.1016/j.jecp.2011.02.002
Judd, C. M., Westfall, J., & Kenny, D. A. (2016). Experiments with more than one random factor: Designs, analytic models, and statistical power. *Annual Review of Psychology*. https://doi.org/10.1146/annurev-psych-122414-033702
Kail, R. (2002). Developmental change in proactive interference. *Child Development, 73* (6), 1703–1714. https://doi.org/10.1111/1467-8624.00500
Karaman, F., & Hay, J. F. (2018). The longevity of statistical learning: When infant memory decays, isolated words come to the rescue. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(2), 221. https://doi.org/10.1037/xlm0000448
Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software, 82*(13), 1–26.
Lawson, C. A., & Fisher, A. V. (2011). It's in the sample: The effects of sample size and sample diversity on the breadth of inductive generalization. *Journal of Experimental Child Psychology, 110*(4), 499–519. https://doi.org/10.1016/j.jecp.2011.07.001
Lenth, R. (2020). *emmeans: Estimated Marginal Means, aka Least-Squares Means* (1.5.0) [Computer software]. https://CRAN.R-project.org/package=emmeans.
Looser, C. E., & Wheatley, T. (2010). The tipping point of animacy: How, when, and where we perceive life in a face. *Psychological Science, 21*(12), 1854–1862. https://doi.org/10.1177/0956797610388044
McGaugh, J. L. (2000). Memory—A Century of Consolidation. *Science, 287*(5451), 248–251. https://doi.org/10.1126/science.287.5451.248
Nelson, D. G. K., Frankenfield, A., Morris, C., & Blair, E. (2000). Young children's use of functional information to categorize artifacts: Three factors that matter. *Cognition, 77* (2), 133–168. https://doi.org/10.1016/S0010-0277(00)00097-4
Opfer, J. E., & Bulloch, M. J. (2007). Causal relations drive young children's induction, naming, and categorization. *Cognition, 105*(1), 206–217. https://doi.org/10.1016/j.cognition.2006.08.006
Perry, L. K., Axelsson, E. L., & Horst, J. S. (2016). Learning what to remember: Vocabulary knowledge and children's memory for object names and features. *Infant and Child Development, 25*(4), 247–258. https://doi.org/10.1002/icd.1933
Perry, L. K., Samuelson, L. K., Malloy, L. M., & Schiffer, R. N. (2010). Learn Locally, Think Globally: Exemplar Variability Supports Higher-Order Generalization and Word Learning. *Psychological Science, 21*(12), 1894–1902. https://doi.org/10.1177/0956797610389189
Portrat, S., Barrouillet, P., & Camos, V. (2008). Time-related decay or interference-based forgetting in working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(6), 1561. https://doi.org/10.1037/a0013356
Quinn, P. C. (2004). Development of Subordinate-Level Categorization in 3- to 7-Month-Old Infants. *Child Development, 75*(3), 886–899. https://doi.org/10.1111/j.1467-8624.2004.00712.x
Quinn, P. C., & Johnson, M. H. (2000). Global-Before-Basic Object Categorization in Connectionist Networks and 2-Month-Old Infants. *Infancy, 1*(1), 31–46. https://doi.org/10.1207/S15327078IN0101_04
Quinn, P. C., Slater, A. M., Brown, E., & Hayes, R. A. (2001). Developmental change in form categorization in early infancy. *British Journal of Developmental Psychology, 19* (2), 207–218. https://doi.org/10.1348/026151001166038
Quinn, P. C., & Tanaka, J. W. (2007). Early development of perceptual expertise: Within-basic-level categorization experience facilitates the formation of subordinate-level category representations in 6- to 7-month-old infants. *Memory & Cognition, 35*(6), 1422–1431. https://doi.org/10.3758/BF03193612
Ross, G., Nelson, K., Wetstone, H., & Tanouye, E. (1986). Acquisition and generalization of novel object concepts by young language learners*. *Journal of Child Language, 13* (1), 67–83. https://doi.org/10.1017/S0305000900000301
Samuelson, L. K., Schutte, A. R., & Horst, J. S. (2009). The dynamic nature of knowledge: Insights from a dynamic field model of children's novel noun generalization. *Cognition, 110*(3), 322–345. https://doi.org/10.1016/j.cognition.2008.10.017
Samuelson, L. K., & Smith, L. B. (2000). Children's Attention to Rigid and Deformable Shape in Naming and Non-Naming Tasks. *Child Development, 71*(6), 1555–1570. https://doi.org/10.1111/1467-8624.00248
Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language, 110*, Article 104038. https://doi.org/10.1016/j.jml.2019.104038
Schonberg, C. C., Russell, E. E., & Luna, M. L. (2020). Effects of past language experience and present language context on the shape bias in Spanish-English bilingual children. *Developmental Science, 23*(2). https://doi.org/10.1111/desc.12879
Slone, L. K., & Sandhofer, C. M. (2017). Consider the category: The effect of spacing depends on individual learning histories. *Journal of Experimental Child Psychology, 159*, 34–49. https://doi.org/10.1016/j.jecp.2017.01.010
Sloutsky, V. M., Kloos, H., & Fisher, A. V. (2007). When Looks Are Everything: Appearance Similarity Versus Kind Information in Early Induction. *Psychological Science, 18*(2), 179–185. https://doi.org/10.1111/j.1467-9280.2007.01869.x
Smith, L. B., Jones, S. S., Yoshida, H., & Colunga, E. (2003). Whose DAM account? Attentional learning explains Booth and Waxman. *Cognition, 87*(3), 209–213. https://doi.org/10.1016/s0010-0277(02)00236-6
Smith, L. B., & Kemler, D. G. (1977). Developmental trends in free classification: Evidence for a new conceptualization of perceptual development. *Journal of Experimental Child Psychology, 24*(2), 279–298. https://doi.org/10.1016/0022-0965(77)90007-8
Son, J. Y., Smith, L. B., & Goldstone, R. L. (2008). Simplicity and generalization: Short-cutting abstraction in children's object categorizations. *Cognition, 108*(3), 626–638. https://doi.org/10.1016/j.cognition.2008.05.002
Sumner, E., DeAngelis, E., Hyatt, M., Goodman, N., & Kidd, C. (2019). Cake or broccoli? Recency biases children's verbal responses. *PloS one, 14*(6), e0217207.
Twomey, K. E., Ranson, S. L., & Horst, J. S. (2014). That's More Like It: Multiple Exemplars Facilitate Word Learning. *Infant and Child Development, 23*(2), 105–122. https://doi.org/10.1002/icd.1824
Unger, L., & Fisher, A. V. (2019). Rapid, experience-related changes in the organization of children's semantic knowledge. *Journal of Experimental Child Psychology, 179*, 1–22. https://doi.org/10.1016/j.jecp.2018.10.007
Vales, C., States, S. L., & Fisher, A. V. (2020). Experience-Driven Semantic Differentiation: Effects of a Naturalistic Experience on Within- and Across-Domain Differentiation in Children. *Child Development, 91*(3), 733–742. https://doi.org/10.1111/cdev.13369
Vlach, H. A. (2014). The Spacing Effect in Children's Generalization of Knowledge: Allowing Children Time to Forget Promotes Their Ability to Learn. *Child Development Perspectives, 8*(3), 163–168. https://doi.org/10.1111/cdep.12079
Vlach, H. A. (2019). Learning to remember words: Memory constraints as double-edged sword mechanisms of language development. *Child Development Perspectives, 13*, 159–165. https://doi.org/10.1111/cdep.12337
Vlach, H. A., Ankowski, A. A., & Sandhofer, C. M. (2012). At the same time or apart in time? The role of presentation timing and retrieval dynamics in generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(1), 246–254. https://doi.org/10.1037/a0025260
Vlach, H. A., & Sandhofer, C. M. (2011). Developmental differences in children's context-dependent word learning. *Journal of Experimental Child Psychology, 108*(2), 394–401. https://doi.org/10.1016/j.jecp.2010.09.011
Vlach, H. A., Sandhofer, C. M., & Bjork, R. A. (2014). Equal spacing and expanding schedules in children's categorization and generalization. *Journal of Experimental Child Psychology, 123*, 129–137. https://doi.org/10.1016/j.jecp.2014.01.004
Vlach, H. A., Sandhofer, C. M., & Kornell, N. (2008). The spacing effect in children's memory and category induction. *Cognition, 109*(1), 163–167. https://doi.org/10.1016/j.cognition.2008.07.013
Vlach, H., & Sandhofer, C. M. (2012). Fast Mapping Across Time: Memory Processes Support Children's Retention of Learned Words. *Frontiers in Psychology, 3*. https://doi.org/10.3389/fpsyg.2012.00046
Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General, 143*(5), 2020–2045. https://doi.org/10.1037/xge0000014

Wojcik, E. H. (2017). 2.5-year-olds' retention and generalization of novel words across short and long delays. *Language Learning and Development, 13*, 300–316. https://doi.org/10.1080/15475441.2016.1275644

Yoshida, H., & Smith, L. B. (2003). Known and novel noun extensions: Attention at two levels of abstraction. *Child Development, 74*(2), 564–577. https://doi.org/10.1111/1467-8624.7402016

Gathercole, V. C. M., Cramer, L. J., Somerville, S. C., & de Haar, M. J. op. (1995). Ontological categories and function: Acquisition of new names. *Cognitive Development*, *10*(2), 225–251. https://doi.org/10.1016/0885-2014(95)90010-1.