Consider the category: The effect of spacing depends on individual learning histories

Lauren K. Slone [*,1], Catherine M. Sandhofer

Department of Psychology, University of California, Los Angeles, Los Angeles, CA 90095, USA

A R T I C L E   I N F O

A B S T R A C T

The spacing effect refers to increased retention following learning instances that are spaced out in time compared with massed together in time. By one account, the advantages of spaced learning should be independent of task particulars and previous learning experiences given that spacing effects have been demonstrated in a variety of tasks across the lifespan. However, by another account, spaced learning should be affected by previous learning because past learning affects the memory and attention processes that form the crux of the spacing effect. The current study investigated whether individuals' learning histories affect the role of spacing in category learning. We examined the effect of spacing on 24 2- to 3.5-year-old children's learning of categories organized by properties to which children's previous learning experiences have biased them to attend (i.e., shape) and properties to which children are less biased to attend (i.e., texture and color). Spaced presentations led to significantly better learning of shape categories, but not of texture or color categories, compared with massed presentations. In addition, generalized estimating equations analyses revealed positive relations between the size of children's "shape-side" productive vocabularies and their shape category learning and between the size of children's "against-the-system" productive vocabularies and their texture category learning. These results suggest that children's attention to and memory for novel object categories are strongly related to their individual word-learning histories. Moreover, children's learned attentional biases affected the types of categories for which spacing facilitated learning.

* Corresponding author.
  E-mail address: laurenkslone@gmail.com (L.K. Slone).
[1] Current address: Department of Psychological and Brain Sciences, Indiana University–Bloomington, Bloomington, IN 47405, USA.

These findings highlight the importance of considering how learners' previous experiences may influence future learning.

Introduction

Spacing repetitions of learning instances over time promotes better memory and generalization than massing learning instances together in time (e.g., Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Childers & Tomasello, 2002; Donovan & Radosevich, 1999; Ebbinghaus, 1885/1913; Kornell & Bjork, 2008). This robust finding has been demonstrated in a variety of category-learning tasks across the lifespan (e.g., Kornell & Bjork, 2008; see also Vlach, 2014), suggesting that the advantages of spaced learning schedules are not tied to particular tasks and may be independent of individuals' previous learning experiences. However, the effects of spacing depend heavily on learners' online memory and attention (e.g., Appleton-Knapp, Bjork, & Wickens, 2005; Bjork & Bjork, 1992). Memory and attention, in turn, are heavily influenced by biases individuals acquire through past learning experiences (e.g., Castel, McCabe, Roediger, & Heitman, 2007; Hershler & Hochstein, 2009). It is unclear, therefore, whether spacing benefits learning of all types of information equally or whether spacing might preferentially facilitate acquisition of information to which learners' prior experiences bias them to attend.

The current study investigated how the memory and attention dynamics associated with the spacing effect are affected by the biases individuals acquire through their previous learning experiences. We did so by comparing the contents of children's early vocabularies (i.e., the first several hundred words children produce); such contents are indicators of the object properties to which children's experiences have biased them to attend. We examined the effect of spacing on children's learning of two types of novel object categories: categories organized by properties to which children's experiences have biased them to attend (i.e., shape) and properties to which they are less biased to attend (i.e., texture and color).

The spacing effect

The term *spacing effect* refers to the robust phenomenon whereby learning is increased via study that is spaced out compared with study that is massed together in time (Cepeda et al., 2006). One common explanation of the spacing effect is that, counterintuitively, spacing promotes learning by allowing time for forgetting during the interval between successive presentations. That is, forgetting an initial instance has been shown repeatedly to increase the potency of encoding subsequent instances (e.g., Bjork & Allen, 1970; Cuddy & Jacoby, 1982; see also Bjork, 2014). This is theorized to occur because the less accessible an item is in memory due to forgetting, the more the memory of that item is strengthened (learned) when it is successfully retrieved (Bjork & Bjork, 1992).

The spacing effect depends on memory and attention processes

Although spacing can enhance learning, the effect of spacing appears to be highly dependent on learners' online memory and attention. For instance, some forgetting during a spacing interval is optimal, but too much forgetting can be detrimental (e.g., Appleton-Knapp et al., 2005); if a learner completely forgets previous category instances during a spacing interval, it may be difficult to abstract relevant category features (Gagné, 1950). Importantly, the rate at which forgetting occurs is thought to depend on the strength of encoding (Bjork, 2014). Features that are less attended to, and therefore not as strongly encoded, may be forgotten more quickly than features that are better attended to.

Previous learning experiences affect memory and attention processes

The memory and attention processes that are important for spacing are heavily influenced by individuals' learning histories. Numerous studies find that prior experiences and domain-specific expertise affect both memory and attention across a variety of domains (e.g., Ericsson & Kintsch, 1995; Geoffrey, Brooks, & Allen, 1989; Hatano & Osawa, 1983; Lehmann & Ericsson, 1996). For example, when reading a passage about a baseball game, junior high school students' short- and long-term retention of the passage were dependent on students' prior knowledge of baseball (Recht & Leslie, 1988). In another study, individuals studied a word list containing animal names that represented American professional football teams (e.g., broncos; Castel et al., 2007). Individuals with high levels of knowledge about football correctly recalled more animal names, but also suffered from more intrusions of nonpresented animal names that represented football teams, compared with individuals with low levels of football knowledge. Such findings suggest that an individual's learning history can both facilitate and interfere with future memory and learning.

Research also suggests that the influence of prior experiences on memory and attention extends to categorization tasks. For instance, adults trained to be experts at identifying "Greebles"—novel objects that share a common spatial configuration—showed enhanced memory for specific Greebles, as well as better categorization of novel Greebles several months after training, compared with novices (Gauthier, Williams, Tarr, & Tanaka, 1998). In categorization tasks with young children, training with actions that draw attention to different properties of a category exemplar systematically changes the range of test items to which children extend category labels (Smith, 2005). Similarly, word labels that draw attention to different category boundaries affect the category prototypes infants form (Althaus & Westermann, 2016). Together with the aforementioned memory studies, this research demonstrates that past learning experiences can significantly affect memory and attention during future learning and categorization. The current study examined whether such effects of learning history on memory and attention influence whether and how learning benefits from spacing.

Children's early vocabularies and spaced learning

The current study examined one type of learning history, attentional biases reflected in children's early vocabularies, and its impact on the spacing effect in children's category learning. Because the term *category* often means different things to different people, we first describe how the term is used in this article; the term category refers to a set of distinct objects to which a common word label is applied. Similar definitions have been widely used in the psychology literature (e.g., Bruner, Goodnow, & Austin, 1956; Mervis & Rosch, 1981; Neisser, 1987). Several recent studies have examined the effect of spacing on children's category learning (e.g., Vlach, Ankowski, & Sandhofer, 2012; Vlach, Sandhofer, & Bjork, 2014; Vlach, Sandhofer, & Kornell, 2008). For instance, Vlach and colleagues (2008) examined the effect of spacing on 3-year-olds' acquisition of labels for novel object categories organized by similarity in shape. Children were presented with three category exemplars either in immediate succession (massed condition) or with 30-s play intervals between exemplars (spaced condition). The authors found that children correctly extended category labels to novel objects matching in shape more often following spaced presentations compared with massed presentations. Similarly, Vlach and colleagues (2012) also found a benefit of spacing on 2-year-olds' categorization of novel objects organized by similarity in shape.

Vlach and colleagues (2008, 2012, 2014) concluded from such studies that the forgetting induced by spacing promotes category induction. Specifically, they argued that forgetting facilitates categorization by promoting the memory of relevant category features and deterring the memory of irrelevant features. However, the evidence for this argument is based on English-learning children's generalization of labels for categories organized by similarity in shape, a feature that English-learning children are typically strongly biased to attend to by around 18 to 24 months of age (e.g., Gershkoff-Stowe & Smith, 2004; Smith, 2000, 2001; see "The shape bias and spaced learning" section below). It is possible that the effect of spacing might be influenced by learned attentional biases, such that spacing might not affect children's categorization of all types of information equally.

The shape bias and spaced learning

Children's learning histories may affect the impact of spacing on category learning by differentially promoting learning of shape-based object categories but not learning of categories organized by other dimensions. At around 18 to 24 months of age, English-learning children begin to show a bias to attend to object shape and map new words to object shape rather than to texture or color (e.g., Gershkoff-Stowe & Smith, 2004; Smith, 2000, 2001). This "shape bias" is thought to arise because children's early English vocabularies are dominated by object names that refer to solid things belonging to categories organized by shape (e.g., ball, cup). The shape bias is also observed in some other languages, although not as robustly (e.g., Gathercole & Min, 1997; Hahn & Cantrell, 2012; Waxman, Senghas, & Benveniste, 1997).

For instance, Samuelson and Smith (1999) asked adults to judge whether the 312 nouns on the MacArthur–Bates Communicative Developmental Inventory referred to a category of items (a) characterized by solidity or nonsolidity and (b) organized by similarity in shape or in material as well as (c) whether each noun was a mass or count noun. Samuelson and Smith found that young children's productive vocabularies consisted primarily of nouns judged by adults to be count nouns that name solid objects organized by similarity in shape (e.g., ball, cup). These so-called "shape-side" nouns (Perry & Samuelson, 2011) are thought to train children to attend primarily to objects' shapes, such that with increased shape-side vocabulary, children's shape bias becomes even more robust (see Smith & Samuelson, 2006).

When young children with a shape bias are learning shape-based categories, memory and attention dynamics may facilitate categorization. Specifically, children may attend less to features like texture and color, compared with shape, and therefore forget these features more quickly than shape during a spacing interval. Although this forgetting likely facilitates retrieval and abstraction of shape (Vlach et al., 2008), it may impair learners' ability to acquire categories with central properties other than shape.

Recent research suggests that the extent to which learning benefits from spacing depends on the degree to which the information to be learned is vulnerable to forgetting. For instance, spaced study of brand advertisements (Appleton-Knapp et al., 2005) resulted in adults' recall of significantly less information when the to-be-learned information was more vulnerable to forgetting (i.e., due to context variation). Thus, in contrast to previous assertions (Vlach et al., 2008, 2012, 2014), we contend that during the developmental period when children are likely to exhibit a shape bias—such that features other than shape are vulnerable to forgetting—spacing might not promote, and may even be detrimental to, children's ability to categorize by nonshape features.

Individual differences in children's word-learning histories

Nevertheless, there are individual differences in children's word-learning histories that might make some children more likely than others to attend to shape versus nonshape features. Perry, Axelsson, and Horst (2015) found that 2-year-olds who produced more shape-side nouns were more accurate at remembering objects' shapes than 2-year-olds who produced fewer shape-side nouns. In addition, Perry and Saffran (2016) found that toddlers who said more, compared with fewer, shape-side nouns were less affected in recognizing a target object whose color had been changed, suggesting that color information was less central to the object representations of children with larger shape-side vocabularies.

Moreover, although children's early vocabularies are dominated by shape-side nouns, children also learn "material-side" nouns—nouns judged by adults to be mass nouns that name nonsolids organized by similarity in material (e.g., water, pudding)—and "against-the-system" nouns—nouns that do not support the link between solidity and attention to shape (e.g., chalk and ice, both of which are solid but belong to categories organized by material; Perry & Samuelson, 2011). Knowing more material-side and against-the-system nouns, which label categories organized by properties other than shape, might make children more likely to attend to properties in addition to shape. For instance, Perry and Samuelson (2011) found that the more against-the-system nouns 1.5-year-olds had in their

productive vocabularies, the more likely they were to extend novel object labels based on similarity in material.

Thus, children's individual learning histories, reflected in the contents of their productive vocabularies, may predispose some children to attend to shape and/or nonshape features more so than other children. Such differences in attention to specific object features may result in individual differences in the effect of spacing on children's learning of shape and nonshape categories.

The current study

The current study investigated whether individuals' learning histories affect the role of spacing in children's category learning. In contrast to the contention of Vlach and colleagues (2008, 2012, 2014)—that the forgetting induced by spacing facilitates categorization by promoting the memory of *relevant* category features—we argue that the forgetting induced by spacing may facilitate categorization only by *particular* object features dependent on the past experiences of learners. We examined the effect of spacing on children's learning of novel shape, texture, and color categories. Spacing may promote categorization based on features to which children have learned to attend (i.e., shape) because such information is likely to be represented strongly enough in memory to be retrieved after a spacing delay. In contrast, spacing may impair children's ability to categorize by features to which they have learned to pay less attention (i.e., texture and color) because such features may be less easily recalled after a spacing delay. Because young children's attention to different object features has been shown to differ as a function of productive vocabulary, we also analyzed the contents of children's vocabularies—shape-side, material-side, and against-the-system nouns—and how these contents related to categorization performance.
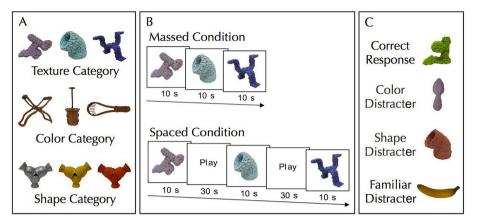
Method

Participants

The participants were 24 2- to 3.5-year-old monolingual English-speaking children (M_{age} = 30.5 months, SD = 4.7, range = 22–42). This relatively large age range was recruited to obtain a sample with a broad range of experiences in learning various types of nouns and a broad range of performances on the categorization task. Half of the children were randomly assigned to participate in the massed condition, and the other half were assigned to participate in the spaced condition. Power analyses based on the effect size (η_p^2 = .69) for the spacing effect in children's shape categorization reported in Vlach et al. (2008) indicated that a sample size of 12 infants per condition would yield power = 0.99. There was an equal number of boys and girls in each condition. Age did not differ significantly between conditions, $t(22)$ = 0.01, p = .990, d = 0.01. Children learned three category types (shape, texture, and color), each type on a separate day (within a span of 3–23 days, M = 10.7 days, SD = 5.8). Category order was counterbalanced across participants. Data from additional participants were excluded from the final sample due to inability to complete all three sessions (n = 9), not passing the pretest (n = 9), parents not returning the vocabulary questionnaire (n = 5), failure to understand the task (n = 5), or experimenter error (n = 1). Children were recruited from preschools in the Los Angeles area on the U.S. west coast and were given a small gift (a book) for their participation.

Stimuli

Children were presented with 12 novel object categories organized by similarity in shape (4 categories), texture (4 categories), and color (4 categories). Each category consisted of three exemplars that shared a central property (shape, texture, or color) and differed on the other dimensions. For instance, exemplars of a texture category had the same surface texture but varied in color and shape (see Fig. 1A). Exemplars of each category were given the same novel label (i.e., wug, fep, zav, blicket, binto, gorp, koba, dax, toma, coodle, tez, or biv). Category–label pairings and the order in which exemplars were presented were randomized for each participant.

Fig. 1. (A) Example stimuli sets used during the learning phase of the categorization task. (B) Learning phase procedure shown with texture category stimuli for both presentation timing conditions. (C) Example test phase stimuli for texture category.

At test, four objects were presented (see Fig. 1C). One object was a novel instance of the target category (e.g., wug). The second and third objects were novel objects that did not match the target property but did match a category exemplar on one nontarget perceptual feature (i.e., for texture categories, the second test object matched the color of one learning exemplar and the third test object matched the shape of a different learning exemplar); we refer to these objects as the *shape/texture/-color distracters* hereafter. The fourth object was a figurine of a familiar object (e.g., a toy cat, a shoe), the *familiar distracter,* which was equivalent in size to all of the other objects. Familiar objects for which children were likely to already have word labels (i.e., >75% of boys and girls at 24 months of age produce the label; Dale & Fenson, 1996) were included in the multiple-choice test to help us identify children who might not have understood the task; children who selected the familiar distracter on the majority of trials were excluded from our analyses due to failure to understand the task.

Design

The experiment had a 2 (Presentation Timing) × 3 (Category Type) design. Presentation timing (massed or spaced) was a between-participants factor, and category type (shape, texture, color) was a within-participants factor.

Procedure

Vocabulary assessment
Prior to the first session, parents completed a checklist of 312 early learned nouns that have been rated for solidity/nonsolidity, mass/count syntax, and organization by similarity in shape versus material (Samuelson & Smith, 1999). We used this noun checklist to analyze the contents of children's productive vocabularies—shape-side, material-side, and against-the-system nouns—and how these contents related to children's learning of the shape, texture, and color categories. Note that we did not include a color vocabulary measure because only three words on the checklist referred to categories organized by similarity in color but not shape or material (i.e., carrots, pickle, and pumpkin; Samuelson & Smith, 1999).

Comprehension pretest
Children completed a pretest prior to the first session to screen for task comprehension. The comprehension pretest and all subsequent categorization sessions were administered individually. A small book and set of keys were labeled one at a time and placed in front of each child. The child was asked

to offer the experimenter one object (e.g., "Can you hand me the *book*?"). After the child made a selection, the objects were presented again and the child was asked to select the other object (e.g., "Can you hand me the *keys*?"). The experimenter provided neutral feedback (e.g., "Thank you") after each selection. Only children who selected the correct object on both trials were asked to participate in the categorization sessions.

Categorization sessions

At the beginning of each categorization session, the child was told that he or she was going to play a game to learn about new toys. Two experimenters conducted each session: One experimenter kept track of the timing and order of presentations, keeping the objects out of the child's sight until presentation. During presentations, the second experimenter kept the object in the child's visual focus at all times, moving the object with the child's gaze if necessary to ensure equivalent looking times across all presentations. During each session, the child was presented with four novel object categories, each consisting of a learning phase and a test phase.

Learning phase. During the learning phase, three category exemplars were presented either in immediate succession (<1 s between exemplars; massed condition) or with 30-s play intervals between exemplars (spaced condition; see Fig. 1B). Play intervals consisted of reading books, completing puzzles, or playing with Play-Doh or stickers. In both timing conditions, each exemplar was presented for 10 s and was labeled three times (e.g., "Look at this *wug* toy! What a cool *wug* toy! You hold the *wug* toy."). An adjectival frame was used to inform the child that an attribute of the object was being labeled but did not cue attention to any specific property (shape, texture, or color).

Following the learning phase, there was a 2-min retention interval during which the child read a book, completed a puzzle, and/or played with Play-Doh or stickers. A 2-min interval was chosen because it (a) required children to access information from long-term memory during the test phase and (b) was short enough to allow children to stay on-task for the entire experiment.

Test phase. Test trials began immediately following the retention interval. The experimenter presented the four test items to the child simultaneously, saying, "Look at all of these." The child was given up to 30 s to explore the test objects. During this time, if the child did not attend to all of the objects, the experimenter touched each object one at a time, saying, "Look at this." After the child had attended to each test object, all four objects were gathered by the experimenter, mixed randomly, and pushed toward the child in a horizontal line. The experimenter then asked, "Can you hand me the *wug* toy?" and held out a hand onto which the child could place an object. Children were given neutral feedback ("Thank you") on their selections.

Results

The current study examined whether children's learning histories affect the role of spaced presentation timing in category learning. Specifically, we investigated the possibility that spaced presentation timing may be beneficial for children's learning of only particular types of categories (cf. Vlach et al., 2008, 2012) given the attentional biases engendered by children's previous word-learning experiences. Although participants' age was not of primary interest, we included age in all of our analyses due to the relatively large age range of our participants.

Does spacing promote learning of novel shape, texture, and color categories?

For each child, we calculated the number of correct responses for each category type (shape, texture, color). We first examined whether the order in which children were taught the different categories affected the number of correct responses for each category type. A 3 (Category Order: shape, texture, or color learned during the first session) × 3 (Category Type) repeated-measures analysis of variance (ANOVA) did not reveal any significant main effects or interactions ($Fs < 1.67$, $ps > .177$, $\eta_p^2 s < .137$). Therefore, category order was omitted from further analyses.

Fig. 2 shows the average number of correct responses (out of 4) for each category type by children in both presentation timing conditions. As can be seen in the figure, massed presentation timing resulted in near-chance performance (1 of 4 correct) for all category types, whereas spaced presentation timing resulted in different performances for the three category types, suggesting an interaction between presentation timing and category type. A 2 (Presentation Timing) $\times$ 3 (Category Type) repeated-measures ANOVA with age as a covariate confirmed a significant interaction between presentation timing and category type, $F(1, 21) = 7.86$, $p = .011$, $\eta_p^2 = .272$. There were no significant main effects or other significant interactions ($Fs < 2.36$, $ps > .139$, $\eta_p^2 s < .102$).

To investigate the interaction between presentation timing and category type, three planned independent-samples t tests were conducted (this and all future analyses were two tailed). Spaced presentations led to a significantly greater number of correct responses than massed presentations for shape categories, $t(22) = 2.30$, $p = .031$, $d = 0.94$, but not for texture categories, $t(22) = 1.04$, $p = .308$, $d = 0.43$, or color categories, $t(22) = 1.30$, $p = .207$, $d = 0.52$. Moreover, one-sample t tests revealed that only performance on shape categories by children in the spaced presentation timing condition was significantly above chance, $t(11) = 3.03$, $p = .006$, $d = 0.87$ (all other $ts < 1.78$, $ps > .103$, $ds < 0.51$). Thus, we replicated the spacing effect of Vlach and colleagues (2008, 2012) with the shape category but not with the texture or color category. The effect of spacing on children's categorization performance appears to depend on the category being learned.
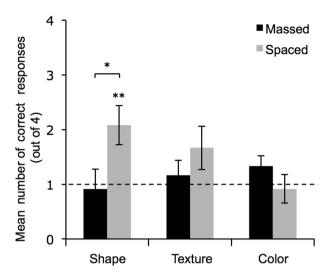
We also examined whether there were systematic patterns to children's errors (shape, color, texture, or familiar distracter choices). Different distracters were available for different categories (e.g., the shape distracter was available as an option only for texture and color categories), such that the differences in the number of target choices across categories could bias the number of times particular distracter items were chosen. Thus, rather than analyzing *numbers* of each type of distracter response, we instead calculated *proportions* to ask the following: Given that a distracter item was chosen, which was it? For each child, we calculated the proportions of each type of distracter response out of that child's total number of incorrect responses (see Fig. 3). Unlike shape, color, and texture distracters, which were options on 67% of trials, the familiar distracter was an option on 100% of trials; therefore, we analyzed familiar distracter responses in a separate analysis from the other distracter responses.

A univariate ANOVA on children's familiar distracter responses yielded no significant effects of age or presentation timing ($Fs < 1.64$, $ps > .214$, $\eta_p^2 s < .073$). As can be seen in Fig. 3, children chose the familiar distracter less often than predicted by chance, $t(23) = 3.57$, $p = .002$, $d = 0.73$. A 2 (Presentation Timing) $\times$ 3 (Novel Distracter Type: shape, color, or texture) repeated-measures ANOVA with age as a covariate yielded no significant main effects or interactions ($Fs < 2.00$, $ps > .148$, $\eta_p^2 s < .087$), suggesting that novel distracter responses did not differ significantly by presentation timing, distracter type, or age. As can be seen in Fig. 3, children did not chose shape, color, or texture distracters significantly more often than predicted by chance ($ts < 1.52$, $ps > .142$, $ds < 0.31$). Thus, although children were more likely to select a novel distracter compared with a familiar distracter on trials when they did not choose the target, there were no discernible patterns as to *which* novel distracter children chose.
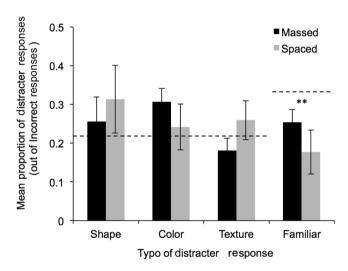
Previous word-learning experiences and categorization performance

To investigate why the effect of spacing varied by category type, we analyzed the contents of participants' vocabularies, which are markers of the object features to which children's early word-learning experiences have biased them to attend. First, we calculated the number of early learned nouns that parents indicated were in each child's productive vocabulary. Previous studies suggest that children with more than 50 nouns in their productive vocabularies are likely to exhibit a shape bias (Gershkoff-Stowe & Smith, 2004; Samuelson & Smith, 1999). The children in the current study produced between 128 and 311 ($M_{massed} = 239$, $SD = 50$, $M_{spaced} = 239$, $SD = 52$) of the 312 nouns on the checklist, suggesting that this population would be highly likely to exhibit a shape bias.

Next, we analyzed the particular types of nouns that comprised children's productive vocabularies. Children produced between 73 and 186 shape-side nouns ($M_{massed} = 146$, $SD = 28$, $M_{spaced} = 146$, $SD = 31$), between 8 and 24 material-side nouns ($M_{massed} = 17$, $SD = 6$, $M_{spaced} = 17$, $SD = 4$), and between 13 and 28 against-the-system nouns ($M_{massed} = 22$, $SD = 5$, $M_{spaced} = 22$, $SD = 4$). To investigate whether word-learning histories played a significant role in children's category learning, we examined

Fig. 2. Mean number of correct responses by category type for each of the two presentation timing conditions. Error bars indicate standard errors of the mean. The dashed line represents chance performance (1 of 4 correct). Asterisks indicate statistical significance between conditions (specified by the bracket), or compared with chance, at the .05 level (single asterisk) or .01 level (two asterisks).

Fig. 3. Mean proportion of shape, color, texture, and familiar distracter responses out of all incorrect responses for each of the two presentation timing conditions. Error bars indicate standard errors of the mean. The dashed lines represent chance performance (.22 for shape, color, and texture distracter responses; .33 for familiar distracter responses). Asterisks indicate performance (regardless of presentation timing condition, as described in the text) that is statistically below chance at the .01 level.

the relation between children's productive vocabulary contents—the number of shape-side, material-side, and against-the-system nouns they produced—and their shape, texture, and color categorization performances. We were also interested in the relative contributions of vocabulary contents and presentation timing to categorization performances as well as any interactions between vocabulary structure and presentation timing.

Statistical approach

To estimate these contributions and assess possible interactions, we first needed to identify the correct statistical model. The current count data violate several key assumptions of traditional statistical approaches like ordinary least squares (OLS) regression. For instance, OLS regression assumes that the outcome is a normally distributed continuous variable and that all observations are independent of one another (Berry, 1993). In contrast, the outcome variable in the current data—the number of correct responses for each category type—took on a limited range of values and tended to be right skewed. Thus, we treated the response on individual trials as binary (i.e., correctly chose the target or did not choose the target), such that each participant contributed four trials of data for each category type. Because these data fail the assumption of independence of observations, we used a regression-based model able to handle both correlated data and a variety of outcome variable distributions—generalized estimating equation (GEE) models (Liang & Zeger, 1986; Zeger & Liang, 1986).

Model specification. GEE models require several factors to be specified (Ballinger, 2004). For our binary outcome, we specified a binomial distribution and a logit link function. We chose an exchangeable correlation structure, which assumes that all four trials of each category are equally correlated with each other, for each participant (although GEE models are robust to misspecification of the correlation structure; Zeger & Liang, 1986).

Model selection. Predictors of shape, texture, and color performances were investigated in separate GEE models. Vocabulary measures were entered into separate models due to multicollinearity among these predictors (see Table 1). Model selection was carried out using the quasi-likelihood under the independence model information criterion (QIC; Pan, 2001). In this model selection process, the model with the lowest QIC score is judged to be the best fit, although absolute QIC score might not be meaningful.

We began by examining all possible main-effects models. Each model involved prediction of binary shape, texture, or color performances from a dummy-coded variable comparing the two levels of the presentation timing variable (massed and spaced), from continuous vocabulary variables (total nouns, shape-side nouns, material-side nouns, and against-the-system nouns), and/or from age. Once we found the best-fitting main-effects model, for models with more than one predictor we added an interaction term to the model and assessed its significance as well as its effect on the overall model fit. In all cases, models with an interaction term had larger QIC values (worse fit) and nonsignificant interaction coefficients. Table 2 presents five models of children's shape, texture, and color performances. Although only one of the models provides the best fit for a particular category—Model 1 for shape, Model 2 for texture, and Model 5 for color—the results of five models are shown for all three category types for purposes of comparison. For each model, coefficient estimates (*B*s), standard errors (*SE*s), odds ratios (*OR*s), and 95% confidence intervals (*CI*s) are provided for each predictor. We used Wald's chi-square test to assess the statistical significance of individual predictors.

Model interpretation. GEE models produce estimates of the population-averaged effect of each predictor on the outcome after adjusting for covariates and taking into account within-participants correlation. For interpretation, odds ratios are used rather than raw coefficients because the interpretation can be a more intuitive measure of likelihood. Odds ratios greater than 1 are interpreted as increasing the likelihood of an outcome, whereas odds ratios less than 1 are interpreted as

Table 1

Correlations between vocabulary measures and presentation timing.

Measure	1	2	3	4
1. Shape-side nouns	–			
2. Against-the-system nouns	.88[**]	–		
3. Material-side nouns	.77[**]	.89[**]	–	
4. Presentation timing	.02	−.07	.04	–

[**] $p < .01$.

Table 2

Model goodness of fit, coefficient estimates, and odds ratios from five logistic GEE models relating presentation timing and vocabulary measures to children's shape, color, and texture category performances.

Model	Independent variable	Shape			Texture			Color		
		QIC	B (SE)	OR (95% CI)	QIC	B (SE)	OR (95% CI)	QIC	B (SE)	OR (95% CI)
1	Presentation timing	118.74	1.36* (0.61)	3.89 (1.18, 12.83)	128.11	0.55 (0.49)	1.73 (0.66, 4.52)	116.84	−0.52 (0.40)	0.59 (0.27, 1.31)
	Shape-side nouns		0.02* (0.01)	1.02 (1.01, 1.04)		0.01 (0.01)	1.01 (0.99, 1.03)		0.00 (0.01)	1.00 (0.99, 1.02)
2	Against-the-system nouns	130.33	0.10 (0.07)	1.10 (0.97, 1.30)	124.64	0.12 (0.07)	1.13 (0.99, 1.29)	117.15	−0.01 (0.05)	0.99 (0.91, 1.09)
3	Material-side nouns	130.50	0.08 (0.06)	1.09 (0.97, 1.21)	128.43	0.06 (0.06)	1.07 (0.96, 1.19)	116.69	0.00 (0.04)	1.00 (0.93, 1.07)
4	Shape-side nouns	126.08	0.02* (0.01)	1.02 (1.00, 1.04)	127.69	0.01 (0.01)	1.01 (0.99, 1.03)	117.05	0.00 (0.01)	1.00 (0.99, 1.02)
5	Presentation timing	124.95	1.30* (0.59)	3.66 (1.14, 11.72)	128.64	0.55 (0.50)	1.73 (0.65, 4.62)	115.76	−0.52 (0.41)	0.59 (0.27, 1.32)

Note. GEE, generalized estimating equation; QIC, quasi-likelihood under the independence model information criterion; B, coefficient estimate; OR, odds ratio; CI, confidence interval.
* $p < .05$.

decreasing the likelihood of an outcome. Odds ratios equal to 1 are not significantly associated with either increased or decreased likelihood.

Shape category performance

A model with presentation timing and shape-side noun vocabulary as predictors provided the best fit (lowest QIC) to the shape categorization data (see Model 1 in Table 2). Both presentation timing and shape-side noun vocabulary were significant predictors of children's categorization by shape (Wald χ^2s > 4.98, ps < .026). The odds ratio for presentation timing in Model 1 is 3.89 with a 95% confidence interval of 1.18 to 12.83. Thus, we can be 95% confident that the odds of choosing the target on shape category trials were increased by between 18% and 1183% for the average participant in the spaced condition compared with the average participant in the massed condition, holding shape-side vocabulary constant. The odds ratio for shape-side vocabulary in Model 1 is 1.02 with a 95% confidence interval of 1.01 to 1.04. Thus, we can be 95% confident that the odds of choosing the target on shape category trials increased by between 1% and 4% (on average) for every 1-unit increase in shape-side vocabulary, adjusting for presentation timing condition. This is equivalent to saying that for every 10-unit increase in shape-side vocabulary, we can be 95% confident that the odds of choosing the target on shape category trials increased by between 10% and 48%, adjusting for presentation timing condition.

These findings are consistent with our previous finding that spaced presentation timing facilitated categorization by shape. Moreover, presentation timing is a significant predictor of shape categorization even when controlling for children's shape-side noun vocabulary. These findings also support the importance of the size of children's shape-side noun vocabularies for children's shape category learning. Specifically, although as children produce more words their vocabularies naturally tend to take on the structure of the noun checklist (i.e., mostly shape-side nouns), shape-side noun vocabulary size was a better predictor of shape category learning than was total noun vocabulary. This finding suggests that what may matter most for learning shape categories is not the total number of nouns children know but rather the number of nouns that fall into the shape-side classification.

Texture category performance

A model with against-the-system nouns as the only predictor provided the best fit to the texture categorization data (see Model 2 in Table 2). Nevertheless, against-the-system vocabulary accounted for only a marginal amount of the variance in texture categorization performances (Wald χ^2 = 3.50, p = .061). The odds ratio for against-the-system nouns in Model 2 is 1.13 with a 95% confidence interval of 0.99 to 1.29. Thus, we can be 95% confident that the odds of choosing the target on texture category trials increased by between −1% and 29% (on average) for every 1-unit increase in against-the-system vocabulary. These findings suggest that knowing more words that go against the typical classification system of solid objects belonging to categories organized by shape may help children attend to and categorize solid objects based on their textures.

Models 3 and 4 of texture performance, with only a material-side vocabulary predictor and only a shape-side vocabulary predictor, respectively, are included in Table 2 for comparison purposes. As can be seen from comparing the results of Models 2 to 4, although shape-side, material-side, and against-the-system vocabulary sizes were highly correlated (see Table 1), these vocabulary measures appear to account for different amounts of variance in children's categorization performances.

Color category performance

A model with only presentation timing provided the best fit to the color categorization data (see Model 5 in Table 2). Nevertheless, presentation timing did not account for a significant amount of the variance in color categorization performances (Wald χ^2 = 1.64, p = .201). This finding is consistent with our previous finding that spaced presentation timing does not have a significant effect on color categorization. Moreover, in contrast to our findings with shape and texture categories, the types of nouns children produced did not affect children's color category learning.

Discussion

Spaced presentations of category exemplars facilitated children's acquisition of shape categories but showed no benefit compared with massed presentations for texture and color categories. These results are consistent with previous research demonstrating that spacing may promote children's categorization by shape (Vlach et al., 2008, 2012); however, the current results also demonstrate that spacing does not benefit all types of category learning equally.

Why did spacing benefit only shape category learning?

One explanation comes from research demonstrating that forgetting and retrieval difficulty may promote children's categorization based on shape (e.g., Vlach et al., 2008, 2012, 2014). According to such an account, spacing allows for forgetting of object features. Therefore, learners must engage in more difficult retrieval to access those features during subsequent exemplar presentations when presentations are spaced compared with when presentations are massed (Vlach et al., 2012). When features from prior presentations are successfully retrieved, memory for those features is strengthened and those features are subsequently forgotten more slowly. However, features that are not successfully retrieved continue to be forgotten and at a faster rate than features that were retrieved (Bjork & Bjork, 1992; see also Bjork, 2014). In sum, spacing promotes forgetting, and therefore more difficult retrieval, than does massing, resulting in stronger memory for information that is successfully retrieved.

Following from such an account, spacing may preferentially benefit children's shape category learning because children's attentional biases influence memory for object features. Children in the current study had productive vocabularies dominated by object names that refer to solid objects belonging to categories organized by shape, biasing children's attention to objects' shapes. A shape bias may facilitate the retrieval of shape information following a spacing interval, facilitating abstraction and categorization by shape. Less attention to features like texture and color, in contrast, may make this information less likely to be retrieved after a spacing interval, hindering abstraction and categorization by such features.

This forgetting account of the data is also consistent with consolidation explanations of the spacing effect (e.g., Landauer, 1969). Consolidation accounts explain the spacing effect as a product of memory consolidation during the intervals between learning events. In the current study, spaced presentation timing resulted in longer intervals between category exemplars, and therefore a longer learning phase, compared with massed presentation timing, allowing for more memory consolidation prior to test. If a shape bias led children to encode objects' shapes more so than objects' textures or colors, consolidation during learning might preferentially strengthen memory for shape, resulting in a spacing effect primarily for shape categories. Nevertheless, a recent study by Vlach and colleagues (2014) found differences in 3-year-olds' shape category learning following two different spacing schedules that were equated in time for consolidation during the learning phase. This finding suggests that consolidation might not be a necessary factor underlying the spacing effect in children's shape category learning.

It is also possible that factors other than memory contributed to children's lack of texture and color category learning in the current study. For instance, research suggests that color categories may be particularly difficult dimensional categories for children to learn (Mervis, Bertrand, & Pani, 1995; Shatz, Behrend, Gelman, & Ebeling, 1996). Although children acquire color categories by 3 or 4 years of age, they nevertheless struggle to learn them (e.g., Backscheider & Shatz, 1993; Sandhofer & Smith, 1999). It is possible, if not likely, that factors in addition to memory make color learning difficult, such that even with optimal presentation timing, young children may struggle to acquire color categories.

Vocabulary structure and category learning

We also analyzed the contents of children's productive vocabularies—shape-side, material-side, and against-the-system nouns—and how these contents related to categorization performance. The size of children's shape-side vocabulary was a significant predictor of shape category performance.

This finding suggests that although all of the children in this study possessed sizable shape-side vocabularies (between 73 and 186 shape-side nouns), each additional shape-side noun that a child learned to produce within this range nevertheless resulted in a small but significantly increased likelihood of learning categories organized by shape. Future research could examine whether, in a population with less extensive shape-side vocabularies, knowing more shape-side nouns would result in even larger increases in shape category performance than those observed here.

In contrast to shape categorization, texture categorization was marginally related to the size of children's against-the-system vocabulary. That is, each additional against-the-system noun that a child had learned to produce resulted in a marginally greater likelihood of learning categories organized by texture. The majority of against-the-system nouns in children's productive vocabularies refer to solid objects categorized by material. Knowing more of these nouns, therefore, may have helped children attend to and categorize the solid objects in the current study based on their textures.

We did not, however, find a significant relation between texture category performance and material-side vocabulary. Material-side nouns refer to *nonsolid* objects categorized by material. Possessing a sizable material-side vocabulary might not have influenced children's performance because the current task involved categorization of solid objects. Future research could examine whether, in a task involving categorization of nonsolids, having a larger material-side vocabulary would relate to categorization performance.

In contrast to texture category performance, knowing more against-the-system nouns did *not* help children attend to and categorize objects based on their colors. Of the 28 against-the-system nouns that children may have known, 25 of them refer to categories of objects organized by similarity in material, whereas only 1 of them (i.e., green beans) refers to a category organized by similarity in both color and material (Samuelson & Smith, 1999). Thus, knowing more against-the-system nouns may train children to better attend to the materials, but not the colors, of solid objects. Thus, a vocabulary structure that directs attention *away* from shape will not necessarily facilitate categorization by non-shape features. Rather, vocabulary structure must also direct attention *toward* a particular feature to facilitate categorization by that feature.

Interestingly, despite the relatively large age range of the participants in this study, we found no evidence to suggest that increased age resulted in better novel category learning. Rather, GEE analyses suggested that vocabulary composition and presentation timing were the strongest predictors of novel category learning. It is possible that with a different age range, perhaps one that included even younger children who might not have yet developed a shape bias, we may have found an effect of age on categorization performance. Nevertheless, the current results suggest that, during the developmental period when children are likely to exhibit a shape bias, children's ability to categorize novel objects by shape and nonshape features is more closely linked with children's individual learning histories, as reflected in the contents of their productive vocabularies, than with chronological age.

Conclusion

The current results suggest that children's attention to and memory for novel object categories is strongly related to their individual word-learning histories. Because the effects of presentation timing are influenced by learners' online memory and attention, children's learning histories affected the types of categories for which spacing facilitated learning. These findings highlight the importance of considering how learners' previous experiences may influence future learning.

The current study focused on the shape bias, a particularly prominent attentional bias learned by young children. However, older populations also exhibit attentional biases that affect perception and memory for objects and events (e.g., Drew, Võ, & Wolfe, 2013; Eitam, Yeshurun, & Hassan, 2013; Mack & Rock, 1998). Future research should examine the extent to which these other types of biases and previous learning experiences may affect the memory and attention foundations of the spacing effect.

Acknowledgments

References

Althaus, N., & Westermann, G. (2016). Labels constructively shape object categories in 10-month-old infants. *Journal of Experimental Child Psychology, 151*, 5–17.

Appleton-Knapp, S. L., Bjork, R. A., & Wickens, T. D. (2005). Examining the spacing effect in advertising: Encoding variability, retrieval processes, and their interaction. *Journal of Consumer Research, 32*, 266–276.

Backscheider, A. G., & Shatz, M. (1993). Children's acquisition of the lexical domain of color. In K. Beals (Ed.). *What we think, what we mean, and how we say it (Papers from the parasession on the correspondence of conceptual, semantic, and grammatical representations)* (Vol. 2, pp. 11–21). Chicago: Chicago Linguistic Society.

Ballinger, G. A. (2004). Using generalized estimating equations for longitudinal data analysis. *Organizational Research Methods, 7*, 127–150.

Berry, W. D. (1993). Understanding regression assumptions. *Quantitative applications in the social sciences* (Vol. 92). Newbury Park, CA: Sage.

Bjork, R. A., & Allen, T. W. (1970). The spacing effect: Consolidation or differential encoding. *Journal of Verbal Learning and Verbal Behavior, 9*, 567–572.

Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes* (Vol. 2, pp. 35–67). Hillsdale, NJ: Lawrence Erlbaum.

Bjork, R. A. (2014). Forgetting as a friend of learning. In D. S. Lindsay, C. M. Kelley, A. P. Yonelinas, & H. L. Roediger (Eds.), *Remembering: Attributions, processes, and control in human memory-Papers in honour of Larry L. Jacoby* (pp. 15–28). New York: Psychology Press.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York: John Wiley.

Castel, A. D., McCabe, D. P., Roediger, H. L., & Heitman, J. L. (2007). The dark side of expertise domain-specific memory errors. *Psychological Science, 18*, 3–5.

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*, 354–380.

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology, 38*, 967–978.

Cuddy, L. J., & Jacoby, L. L. (1982). When forgetting helps memory: An analysis of repetition effects. *Journal of Verbal Learning and Verbal Behavior, 21*, 451–467.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavioral Research Methods, Instruments, & Computers, 28*, 125–127.

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect. *Journal of Applied Psychology, 84*, 795–805.

Drew, T., Võ, M. L. H., & Wolfe, J. M. (2013). The invisible gorilla strikes again: Sustained inattentional blindness in expert observers. *Psychological Science, 24*, 1848–1853.

Ebbinghaus, H. (1913). *Memory* (H. A. Ruger & C. E. Bussenius, Trans.). New York: Teachers College (Original work published 1885).

Eitam, B., Yeshurun, Y., & Hassan, K. (2013). Blinded by irrelevance: Pure irrelevance induced "blindness". *Journal of Experimental Psychology: Human Perception and Performance, 39*, 611–615.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211–245.

Gagné, R. M. (1950). The effect of sequence of presentation of similar items on the learning of paired associates. *Journal of Experimental Psychology, 40*, 61–73.

Gathercole, V. C. M., & Min, H. (1997). Word meaning biases or language-specific effects? Evidence from English, Spanish, and Korean. *First Language, 17*, 31–56.

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training "Greeble" experts: A framework for studying expert object recognition processes. *Vision Research, 38*, 2401–2428.

Geoffrey, N. R., Brooks, L. R., & Allen, S. W. (1989). Recall by expert medical practitioners and novices as a record of processing attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 1166–1174.

Gershkoff-Stowe, L., & Smith, L. B. (2004). Shape and the first hundred nouns. *Child Development, 75*, 1098–1114.

Hahn, E. R., & Cantrell, L. (2012). The shape-bias in Spanish-speaking children and its relationship to vocabulary. *Journal of Child Language, 39*, 443–455.

Hatano, G., & Osawa, K. (1983). Digit memory of grand experts in abacus-derived mental calculation. *Cognition, 15*, 95–110.

Hershler, O., & Hochstein, S. (2009). The importance of being expert: Top-down attentional control in visual search with photographs. *Attention, Perception, & Psychophysics, 71*, 1478–1486.

Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the "enemy of induction"? *Psychological Science, 19*, 585–592.

Landauer, T. K. (1969). Reinforcement as consolidation. *Psychological Review, 76*, 82–96.

Lehmann, A. C., & Ericsson, K. A. (1996). Performance without preparation: Structure and acquisition of expert sight-reading and accompanying performance. *Psychomusicology, 15*, 1–29.

Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika, 73*, 13–22.

Mack, A., & Rock, I. (1998). *Inattentional blindness*. Cambridge, MA: MIT Press.

Mervis, C. B., Bertrand, J., & Pani, J. R. (1995). Transaction of cognitive-linguistic abilities and adult input: A case study of the acquisition of colour terms and colour-based subordinate object categories. *British Journal of Developmental Psychology, 13*, 285–302.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology, 32*, 89–115.

Neisser, U. (Ed.). (1987). *Concepts and conceptual development: Ecological and intellectual factors in categorization.* Cambridge, UK: Cambridge University Press.

Pan, W. (2001). Akaike's information criterion in generalized estimating equations. *Biometrics, 57*, 120–125.

Perry, L. K., Axelsson, E. L., & Horst, J. S. (2015). Learning what to remember: Vocabulary knowledge and children's memory for object names and features. *Infant and Child Development, 25*, 247–258.

Perry, L. K., & Saffran, J. R. (2016). Is a pink cow still a cow? Individual differences in toddlers' vocabulary knowledge and lexical representations. *Cognitive Science.* http://dx.doi.org/10.1111/cogs.12370. Advance online publication.

Perry, L. K., & Samuelson, L. K. (2011). The shape of the vocabulary predicts the shape of the bias. *Frontiers in Psychology, 2.* http://dx.doi.org/10.3389/fpsyg.2011.00345.

Recht, D. R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology, 80*, 16–20.

Samuelson, L. K., & Smith, L. B. (1999). Early noun vocabularies: Do ontology, category structure, and syntax correspond? *Cognition, 73*, 1–33.

Sandhofer, C. M., & Smith, L. B. (1999). Learning color words involves learning a system of mappings. *Developmental Psychology, 35*, 668–679.

Shatz, M., Behrend, D., Gelman, S. A., & Ebeling, K. S. (1996). Colour term knowledge in two-year-olds: Evidence for early competence. *Journal of Child Language, 23*, 177–200.

Smith, L. B. (2005). Action alters shape categories. *Cognitive Science, 29*, 665–679.

Smith, L. B. (2001). How domain-general processes may create domain-specific biases. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 101–131). Cambridge, UK: Cambridge University Press.

Smith, L. B. (2000). Learning how to learn words: An associated crane. In R. M. Golinkoff, K. Hirsh-Pasek, L. Bloom, L. B. Smith, A. L. Woodward, N. Akhtar, M. Tomasello, & G. Hollich (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 51–80). New York: Oxford University Press.

Smith, L. B., & Samuelson, L. (2006). An attentional learning account of the shape bias: Reply to Cimpian and Markman (2005) and Booth, Waxman, and Huang (2005). *Developmental Psychology, 42*, 1339–1343.

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*, 163–168.

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*, 246–254.

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.

Vlach, H. A., Sandhofer, C. M., & Kornell, N. (2008). The spacing effect in children's memory and category induction. *Cognition, 109*, 163–167.

Waxman, S. R., Senghas, A., & Benveniste, S. (1997). A cross-linguistic examination of the noun-category bias: Its existence and specificity in French- and Spanish-speaking preschool-aged children. *Cognitive Psychology, 32*, 183–218.

Zeger, S. L., & Liang, K. Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics, 42*, 121–130.