Title
Increasing the Long-Term Retention of Words Learned from Shared Book Reading

Author
Luna, Michelle Lynn

Publication Date
2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Increasing the Long-Term Retention of Words Learned from Shared Book Reading

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Psychology

by

Michelle Lynn Luna

2022

ABSTRACT OF THE DISSERTATION

Increasing the Long-Term Retention of Words Learned

from Shared Book Reading

by

Michelle Lynn Luna

Doctor of Philosophy in Psychology

University of California, Los Angeles, 2022

Professor Catherine M. Sandhofer, Chair

Shared book reading, reading sessions between adults and young children, leads to better general vocabulary development. However, studies show that preliterate children, although able to remember words from books immediately, do not retain most of the book-specific words they learn. Thus, words learned from books may require additional memory supports to be retained after a delay. The current dissertation integrates memory, word learning, and book reading research to characterize both the variability and presentation timing of words within shared book reading sessions and examines how these factors lead to increased retention of words learned from books. First, this dissertation determines how variability in the questions asked during shared book reading affects the retention of novel words. Specifically, in Study 1, 38 4-year-olds were read books with words that were accompanied by irrelevant questions, the same questions, or different questions each time the word was presented. Next, I examined how two different presentation timing styles affect the retention of novel words. In Study 2, 39 4-year-olds were read books with words presented either sequentially or spaced in time. Children in both studies were tested at three different time delays (i.e., immediate, 5 minutes, and 24 hours) to determine how many words were retained in memory. This dissertation found, broadly, that neither variability nor presentation timing led to better word retention across the three different time delays. However, children in this study remembered more words than previous studies even when tested after 24 hours. Implications of these preliminary findings are discussed. Ultimately, these results inform word learning theories and serve to create materials for design-based interventions aimed to improve the quality of shared book reading sessions between caregivers and children.

The dissertation of Michelle Lynn Luna is approved.

Alison L. Bailey

Idan A. Blank

Melissa Paquette-Smith

James W. Stigler

Catherine M. Sandhofer, Committee Chair

University of California, Los Angeles

2022

Table of Contents

List of Figures

List of Tables

Acknowledgments

As the first person in my family to go to college and graduate school, I can say that these accomplishments would not have been possible without the support from so many people in my life. I would like to thank my family for supporting me throughout this journey including my parents, Antonio and Raquel, and especially my brothers, Pepe and Tony, who helped me move to different apartments in West LA. Thank you to my niece and nephews, Eli, Vivi, and Anthony, for helping me have fun during my program, especially during Halloween, birthday parties, and holidays. Additionally, I wouldn't have been able to get through finishing this degree in a pandemic without the support of my amazing partner, Sthef. I appreciate all the writing sessions we got to do together! Of course, I also can't forget to thank my dissertation support dogs, Mobi & Chewie.

My academic journey wouldn't have been the same without the support of my advisor, Cathy, who has been instrumental in guiding me through graduate school. I am grateful to Alison as well for teaching me about educational perspectives in research. I would like to thank the Language & Cognitive Development Lab and the Language & Literacy Research Group, whose members gave me invaluable feedback throughout the past 6 years. I am grateful to my friends and colleagues in the Developmental area (especially Amalia and Adriana) who helped me along the way with writing groups, accountability, and venting sessions. Thank you to previous LCD lab members, Natsuki and Christina, who helped me with edits and R analyses. I would also like to thank my dissertation committee members for their help in this process and my research assistants, Marianna Martinez, Madeline Netto, Jennifer Park, Mallika Singh, Rhea Varkhedi, and Nicole Wang, for all their help collecting the data for these studies.

VITA

EDUCATION

2017 **University of California, Los Angeles**
M.A., Developmental Psychology
Committee: Catherine M. Sandhofer, Ph.D. (Chair); Elizabeth Bjork, Ph.D.

2015 **University of Notre Dame**
M.A., Romance Languages; Emphasis: French and Francophone Studies

2014 **University of Notre Dame**
B.A., Psychology; French and Francophone Studies, *Cum Laude*
Honors Thesis: *Executive Function in Young Bilinguals: The Effect of Socioeconomic Status on the Bilingual Advantage*
Advisor: Jeanne Day, Ph.D.

SELECTED FELLOWSHIPS & GRANTS

2019 UCLA Graduate Research Mentorship ($36,800)

2018 UCLA Graduate Division Graduate Summer Research Mentorship ($6000)

2017 UCLA Graduate Division Graduate Summer Research Mentorship ($6000)

2016 UCLA Graduate Division Eugene V. Cota-Robles Fellowship ($92,000)

PUBLICATIONS

Shaw, S. T., **Luna, M. L.**, Rodriguez, B., Yeh, J., Villalta, N., & Ramirez, G. (2022) Mathematical Creativity in Elementary School Children: General Patterns and Effects of an Incubation Break. *Frontiers in Education,* 7. https://doi.org/10.3389/feduc.2022.835911

Luna, M. L., & Sandhofer, C. M. (2021). Arbitrary but Predictive Cues Support Shifting Attention to Overlooked Features. *Journal of Memory and Language*, 120. https://doi.org/10.1016/j.jml.2021.104251

Schonberg, C. C., Russell, E. E., & **Luna, M. L.** (2019). 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

SELECTED PRESENTATIONS

***Luna, M. L.,** Ogren, M., & Sandhofer, C. M. (2022, April). *Supporting Children's Word Learning from Storybooks with Emotional Variability*. Poster presented at the 12th Biennial Meeting of the Cognitive Development Society in Madison, Wisconsin.

***Luna, M. L.,** & Sandhofer, C. M. (2021, April). *Preschoolers' Learning of Prepositions from Storybooks with Predictive Cues*. Poster presentation presented at the 2021 Virtual Biennial Meeting of the Society for Research in Child Development.

***Luna, M. L.,** & Sandhofer, C. M. (2021, April). *The Role of Variability in the Questions Asked During Shared Book Reading with Preschoolers*. Poster presentation presented at the 2021 Virtual Biennial Meeting of the Society for Research in Child Development.

***Luna, M. L.,** & Sandhofer, C. M. (2019, October). *Using Predictive Cues to Learn Prepositions from Storybooks.* Poster presented at the 11th Biennial Meeting of the Cognitive Development Society in Louisville, Kentucky.

***Luna, M. L.,** & Sandhofer, C. M. (2019, July). *The Effects of Contextual Cues on the Learning of Prepositions.* Poster presented at the 41st Annual Meeting of the Cognitive Science Society in Montreal, Canada.

***Luna, M. L.,** & Sandhofer, C. M. (2018, July). *Using Correlated Contextual Cues to Shift Attentional Biases in Word Learning.* Poster presented at the Learning Language in Humans and Machines 2018 Conference in Paris, France.

***Luna, M. L.,** & Sandhofer, C. M. (2018, June). *Shifting Attention from Dominant Biases in Word Learning Using Contextual Cues.* Paper presented at the Swiss Graduate School for Cognition, Learning, & Memory in Weggis, Switzerland.

***Luna, M. L.,** & Sandhofer, C. M. (2017, October). *Using Context-Dependency through Redundant Correlated Contextual Cues in Word Learning.* Poster presented at the 10th Biennial Meeting of the Cognitive Development Society in Portland, Oregon.

SELECTED TEACHING EXPERIENCE

2017 – 2022 **Teaching Fellow,** University of California, Los Angeles

Psychological Statistics
Psychology and Education
Research Methods in Psychology
Cognitive Development
Research in Developmental Psychology
Cognitive Psychology
Advanced Psychological Statistics

Introduction

Reading sessions between adults and preliterate children, commonly known as shared book reading, provide children with opportunities to learn unique words (Mol et al., 2008; Montag et al., 2015). These opportunities are important because greater lexical diversity predicts better language (Deckner et al., 2006; Raikes et al., 2006) and cognitive skill development (Kalb & van Ours, 2014; Marchman & Fernald, 2008). Nonetheless, little is known about the mechanisms that underlie vocabulary learning and retention from shared book reading. Books may benefit word learning by exposing children to lower frequency words, repetition of new words, rich conversation surrounding the reading session, and/or other factors that help children increase their general vocabulary levels. However, at the same time, studies on shared book reading suggest that children do not remember most of the book-specific, target words they hear during the reading sessions (Justice, 2002; McLeod & McDade, 2011; Sénéchal & Cornell, 1993). Moreover, cognitive factors found to increase the retention of information, like variability and presentation timing, have not been studied within the context of shared book reading. Perhaps because books contain relatively infrequent words compared to everyday speech, these words may need memory supports to be remembered. In this dissertation, I use a virtual storybook format to examine the roles of variability and presentation timing as memory supports for word retention during shared book reading.

Shared Book Reading

Books are an important source of decontextualized language because they allow parents and children to talk about things that are not present in their immediate environment. Young children hear words within their everyday interactions, but the most diverse and infrequent words are heard when a caregiver reads to them. If children only heard words they encountered in conversation, their vocabularies would likely be limited by their direct experience. For example, the word "caterpillar" may infrequently appear in everyday situations (e.g., a child seeing a caterpillar outside). However, reading a book (e.g., The Very Hungry Caterpillar) may increase the likelihood that a child learns infrequent words, like "caterpillar," because the book provides multiple instances of the word, in addition to presenting the word in novel semantic and grammatical contexts. Hearing words in more unique contextual situations or within different grammatical structures increases the child's likelihood of learning the word (Huttenlocher et al., 1991). Providing children with more opportunities to learn semantically diverse words may also be important because the order of words children learn is predicted by the frequency of those words (Harris et al., 1988; Huttenlocher et al., 1991; Naigles & Hoff-Ginsberg, 1998; Schwartz & Terrell, 1983) where more frequent words are learned earlier. Because of the strong relationship between the frequency of words in language input and children's vocabulary acquisition (Hoff & Naigles, 2002), increasing the frequency of less common words may help build children's vocabularies. Furthermore, as children's lexicons diversify, the newly learned words help children determine the meaning of other new words (Hoff & Naigles, 2002). For example, as children learn the meaning of words like "caterpillar," they may also be able to determine the meaning of other words that appear within the same context, such as "cocoon." The words used in books go beyond commonly available experiences in daily life, both in frequency and diversity.

General Vocabulary versus Target Words

Shared book reading has been robustly linked to higher levels of general vocabulary (Dowdall et al., 2019; Whitehurst et al., 1988), but it is unclear how shared book reading affects targeted word gains (i.e., learning of words specifically presented in a book). Overall, many studies use general vocabulary measures, such as the Peabody Picture Vocabulary Test (PPVT) (Dunn & Dunn, 2007), to measure long-term growth from reading experiences or reading interventions without measuring growth in target words (Arnold et al., 1994; Bojczyk et al., 2016; Dickinson & Smith, 1994; Evans et al., 2000; Haden et al., 1996; Hindman et al., 2008; Kalb & van Ours, 2014; Reese & Cox, 1999; Whitehurst et al., 1988, 1994). Although few studies examine gains in target vocabulary, shared book reading may differentially affect general vocabulary gains versus targeted word gains. For example, one study found that a reading intervention improved general levels of expressive vocabulary and target vocabulary, but this intervention failed to increase general levels of receptive vocabulary (Hargrave & Sénéchal, 2000). Other research found that literacy interventions helped children learn target words but did not improve general expressive and receptive vocabulary levels (Han et al., 2010; Pollard-Durodola et al., 2011). On the other hand, one study found that interventions improved both general vocabulary levels along with target vocabulary (Wasik & Bond, 2001). Furthermore, some studies measured both general and target vocabulary but used the general vocabulary measurement as a covariate or to assign intervention groups (Jimenez & Saylor, 2017; Reese & Cox, 1999; Sénéchal et al., 1995; Vuattoux et al., 2014) and only measured targeted word gains. Thus, because of the emphasis on general vocabulary level over target word learning and methodological inconsistencies between studies, it is unclear how effective books are at teaching children target words specifically found in stories (Jimenez & Saylor, 2017; Justice, 2002; McLeod & McDade, 2011; Sénéchal & Cornell, 1993).

The types of words that are presented as target words vary among studies. For example, some studies focus on noun learning (e.g., Horst et al., 2011; Jimenez & Saylor, 2017), while others incorporate verbs as well (Ard & Beverly, 2004; McLeod & McDade, 2011). Among the studies that examine gains in book-specific target words, there is a focus on learning nouns; specifically, nouns that are object words tend to be easier for children to learn (Landau et al., 1988; McDonough et al., 2011). Because it is unclear how children learn words from books, examining the most easily learned words first might disentangle which shared book reading factors influence word learning. Examining more difficult words that children struggle learning, such as adjectives or verbs, would make it hard to determine whether the factors are not helpful, or the words are too hard to learn at this age. On the other hand, after finding which factors help children learn nouns, then perhaps these findings can be extended to other, more difficult words.

Word Learning and Memory

Studies show different results regarding the number of target words children learn in one session. In laboratory-based novel-word learning studies, preschoolers can learn about six object words in one laboratory session with different memory supports (Goldenberg & Sandhofer, 2013b; Munro et al., 2012; Vlach et al., 2008). In shared book reading paradigms, some studies show that children can only learn about one word after a single reading session (Jimenez & Saylor, 2017; Sénéchal & Cornell, 1993) and about three words with some story repetition (Horst et al., 2011; Justice, 2002; McLeod & McDade, 2011), which suggests that the words learned from books may require additional memory supports (e.g., repetition) to be remembered. Because words that appear in books are less common than words found in everyday speech, learning vocabulary from books may offer additional challenges and require greater memory support compared to learning words from nonbook paradigms. Importantly, these studies all focus on children's learning of novel words (i.e., words that do not exist in English but sound like they could) or extremely low frequency words. By focusing on novel words, word frequency is controlled, and, arguably, words with the lowest frequency are targeted. Using novel words makes these words comparable (and perhaps even more extreme) to the words found in books, which makes it easier to draw conclusions about how words found in books could be supported.

Word learning, in general, is aided by supports for memory. Learning a word involves associating a label with its meaning, integrating examples across time, and retrieving the word when it is needed (Wojcik, 2013). Thus, memory is a fundamental component of word learning. Although children appear to learn words quickly, they also tend to forget them quickly. In laboratory tasks, children show declines in word recall after even short delays. For example, one study taught children the label for a single, novel object and then tested word retention at three time points: immediately after learning, one week after learning, and one month after learning (Vlach & Sandhofer, 2012a). When tested immediately, roughly 65% of children were able to identify the correct object. However, after one week, approximately 35% of children were successful, and after one month, only around 20% of children were able to identify the correct object. In contrast, with memory supports, approximately 80% of children remembered the label immediately, 70% after a week, and 65% after a month. It is possible that these novel words, require memory supports for children to remember them over time because of their low frequency.

Support for Memory: Variability

Because information is not stored in isolation, the context surrounding the information people learn affects the degree to which people remember. For example, if a child is learning the word "cup," the child also encodes where the cup was in space (e.g., on a table) (Samuelson et al., 2017), the color of the cup (Johnson et al., 2011; Perry & Saffran, 2017), and maybe other objects that were near the cup (Horst et al., 2010). Thus, the child is learning a word embedded within a physical context. Previous studies have examined the effects of many types of contextual factors, including external environment (Dulsky, 1935; Godden & Baddeley, 1975; Smith et al., 1978), internal environment (Bower, 1981; Eich & Macaulay, 2000), semantic factors (Howard & Kahana, 2002; Tulving & Thomson, 1973), linguistic diversity (Hills et al., 2010), and chemical state (Overton, 1966; Petersen, 1977). The presence or absence of the encoded contextual factors during recall affects whether information is remembered, where the presence of cues would perhaps make remembering the information more likely.

Additionally, the number of contexts, or contextual variability, that information is associated with during encoding has different effects on memory during recall (Steyvers & Malmberg, 2003). Multiple studies show that children are sensitive to the variability of the visual environment surrounding word learning, including the background color of an object presentation or the person providing the object labels (Goldenberg & Sandhofer, 2013a, 2013b; Vlach & Sandhofer, 2011). When words are presented with too little contextual variability, children's memory becomes context-dependent and they fail to generalize the label in new contexts. However, with too much contextual variability, children fail to learn the word. For example, when 2-year-olds were presented with an object in both repeated contexts and different contexts (e.g., blue, red, blue, orange, blue), they learned the novel labels and generalized to new contexts. By obtaining support to aggregate different examples through the repetition of the same context and support for decontextualization cues through different contexts, children successfully generalized the word in a completely new context (Goldenberg & Sandhofer, 2013b). Older children (4 to 5-year-olds), on the other hand, showed no effects of contextual variability (Vlach & Sandhofer, 2011), likely due to their increased experience with word learning. Altogether, variability is important for learning and memory; however, the amount of contextual variability that benefits learning and memory may vary across development.

Variability in Shared Book Reading

One way variability can be incorporated into shared book reading settings is the variability in the questions asked to children. Shared book reading best promotes learning when the reader provides a highly interactive environment full of questions, gestures, contingent responses, and praise (Dowdall et al., 2019; Hargrave & Sénéchal, 2000; B. A. Walsh & Blewitt, 2006; Whitehurst et al., 1988). Moreover, conversations between adults and children have been found to help increase children's vocabulary when children are provided with an opportunity to practice using words and get feedback on language use (De Temple & Snow, 2003; Deckner et al., 2006; Hindman et al., 2019). Although the type of talk children hear appears to be important for their language development, observational studies primarily demonstrate that teachers and parents ask few questions during shared book reading, and those questions are usually very simple to answer (Anderson et al., 2012; Deshmukh et al., 2019; Dickinson & Smith, 1994; Hargrave & Sénéchal, 2000; Hindman et al., 2012, 2019; R. L. Walsh & Hodge, 2018). Because of the importance of question asking in developing literacy skills (Bailey et al., 2018; Bailey & Moughamian, 2007), it is critical to study question interactions within shared book reading. However, studies have not examined whether greater variability in the types of questions asked (i.e., asking the same or different questions) during repeated readings of the same book affects learning. One study found that repeated book reading with comments (i.e., a restatement of the target word within the context of the story: "The cat sure is *hasting* Leroy") led to more word learning than repeated book reading with questions (e.g., "What is the cat doing to Leroy?"), but this study was underpowered and did not examine whether the same or different questions should be asked in the story repetition (Ard & Beverly, 2004). Further, because the effectiveness of variability might change depending on age (Goldenberg & Sandhofer, 2013b; Steyvers &

Malmberg, 2003; Vlach & Sandhofer, 2011), it is important to examine different levels of variability in the language surrounding shared book reading to determine whether variability is beneficial or not at a certain age and within a shared book reading context.

Support for Memory: Timing

The memory literature robustly shows that the timing of the presentation of information greatly affects how information is remembered. One well-studied presentation style, the spacing effect, has strong effects on memory among adults (Kornell & Bjork, 2008; Toppino et al., 1991), infants (Vander Linde et al., 1985), and animals (Fanselow & Tighe, 1988) across many domains of learning. The spacing effect literature suggests that memory is strongest when information is presented spaced over temporally discontinuous sessions (e.g., seeing a caterpillar picture 3 times for 30 seconds every 3 minutes) instead of presented within a single, massed session (e.g., seeing a caterpillar picture 1 time for 90 seconds) (Cepeda et al., 2006; Underwood, 1961).

There are multiple theories about how spacing information out in time increases retention, including the Encoding Variability Theory and the Study Phase Retrieval Theory. The Encoding Variability Theory suggests that memory is aided when the contextual information available during encoding is present at test (Glenberg, 1979; Melton, 1970). Learning information spaced across time allows for more contextual cues to be associated with that piece of information, which should be helpful later for retrieval. So, if a child sees multiple examples of caterpillars across time, this theory suggests that spacing out those examples would be helpful because each time the child encounters the word caterpillar, it is accompanied by more contextual cues (e.g., caterpillar features, different settings, and varied emotional states). The Study Phase Retrieval Theory suggests that memory traces are activated with each successive presentation of the material (Hintzman et al., 1975; Thios & D′Agostino, 1976; Tzeng et al., 1979). Spaced presentations allow for this reactivation to occur while massed presentations do not because in massing, the first trace is still active when the second one is presented. This theory would explain that seeing two caterpillar examples across time is beneficial because the second example would help reactivate the first example and make that memory trace stronger. Both theories have elements that suggest that spacing, embedded within a storybook context, may help children better remember words.

Studying how people remember involves studying how people forget. Although counterintuitive, forgetting might be a helpful mechanism for remembering. As people forget insignificant details, they are left with the essence of what is important, an abstraction from the specific examples they experience (Vlach & Kalish, 2014). For example, after the first instance of seeing a caterpillar and hearing the word "caterpillar," the passage of time should cause a child to forget specific, unimportant details about the event: that the specific caterpillar was long or spotted. When children receive another presentation after some time has passed, the memory is reactivated, and the important, shared details between presentations get strengthened in memory with each exposure. Thus, after some time, the reactivation of the information could allow for better encoding of that information. Conversely, when information is presented all at once, there is no chance to forget or reactivate memory upon subsequent presentations. Thus, both unimportant, specific details and important, shared features are forgotten at the same rate. For example, hearing the label of a caterpillar multiple times in a single session would result in forgetting both the unimportant details of a caterpillar, such as its size or color, as well as more important details, such as the shape of its body.

Moreover, without memory reactivation, the forgetting curve for both insignificant details and important shared features would be steep. Multiple studies indicate that a spaced out presentation style benefits children's retention of new words (Vlach, 2014; Vlach et al., 2008, 2012; Vlach & Sandhofer, 2012b) and ability to extend the category label to new examples (Vlach et al., 2008, 2012). Thus, manipulating how information is presented might strongly affect how well it is learned and remembered.

Another way spacing might be beneficial is by supporting discrimination between information. Interleaving, a special case of spacing, also presents information after a delay. However, people are presented with other information during the delay instead of having nothing in between presentations (Kornell & Bjork, 2008; Vlach & Johnson, 2013). For example, a child might need to distinguish between a caterpillar and a worm. A child could distinguish between "caterpillar" and "worm" more easily by seeing pictures of both worms and caterpillars to differentiate the features associated with each label. An interleaved presentation would alternate between pictures of each object, while a massed presentation would only show one set of pictures at a time. By only seeing examples of caterpillars and not worms, children might have a harder time establishing the category-defining features of a "caterpillar." By interleaving information, people can compare and contrast different examples in order to discriminate between what they are learning (Kornell & Bjork, 2008). Although learners describe interleaved presentation as unintuitive and less effective, interleaved learning has been shown to be better for memory compared to massed learning, especially for material that is harder to discriminate (Kornell & Bjork, 2008).

Despite an abundance of research on the spacing effect in adults, only a few studies have examined the spacing effect from a developmental perspective or among preschoolers (Childers

& Tomasello, 2002; Rea & Modigliani, 1987; Toppino et al., 1991; Toppino & DiGeorge, 1984; Vlach et al., 2008). The preschool years are of particular importance because of the substantial gains in word learning and cognitive development during this age range. Studies have found contradictory effects of the spacing effect in early childhood. One study, in particular, found no benefit of spacing for preschoolers when remembering pictures but did find a benefit of spacing for first graders (Toppino & DiGeorge, 1984). Other studies find a beneficial effect of spacing when remembering both words and pictures, even among preschoolers (Rea & Modigliani, 1987; Toppino et al., 1991). In addition to the effects of memory in general, studies have also looked at the spacing effect during word learning specifically. These studies find that children learn more words when the objects and labels are presented in a spaced fashion across multiple sessions rather than all together in one session (Childers & Tomasello, 2002; Vlach et al., 2008). However, even though spacing might be beneficial for memory, material that is too far spaced out might not be remembered as well (Cepeda et al., 2006; Childers & Tomasello, 2002; Vlach, 2014). Thus, studying how presentation timing affects memory while learning words may provide insight into what best supports word learning and memory for preschoolers.

Timing in Shared Book Reading

The timing with which instances are presented is given little consideration in developmental research or designing educational materials. Word learning research, for example, almost exclusively relies on massed presentations of words. In a typical word learning study, participants are presented with multiple instances of novel categories, one after the other, with minimal time delay in between presentations. This presentation timing contrasts strikingly with real-world experiences in which individual instances of words appear spaced out in time. For example, children hear the word “cup” spread out across the day and over multiple days before they understand or produce the word. One exception to this spaced presentation of words may be during storybook reading, especially if children want to read the same book multiple times successively. Books are probably also read over several days, spread out in time. Although spacing has been shown to be beneficial to memory and even word learning, it has not been explicitly studied within a shared book reading setting. In classrooms, some studies showed that using spacing while teaching benefitted student learning of educational concepts (Gluckman et al., 2014; Vlach & Sandhofer, 2012b). Thus, presentation timing might be an effective tool to increase learning in educational settings. Spacing could be examined using storybooks in two ways: within and across book reading sessions. Within books, words can be manipulated to occur either all at once or interleaved across different pages. Across book reading sessions, books can be read over multiple days. Ultimately, incorporating a spaced presentation of words into storybooks can help determine whether the presentation timing of words matters specifically for words learned in shared book reading sessions.

Retention

An important factor when considering remembering, learning, and forgetting is time. Generally, studies that examine the role of time in forgetting find that forgetting happens steeply and quickly if memory is not supported in some way (Ebbinghaus, 1964; Fawcett & Hulbert, 2020; Loftus, 1985; Rubin & Wenzel, 1996; Shiffrin & Atkinson, 1969; Storm, 2011; White, 2001). Although forgetting is intuitively thought of as a memory failure, studies on human cognition show that forgetting is useful across many domains, including abstraction, generalization, creative problem-solving, and general learning of information (Fawcett & Hulbert, 2020; Storm, 2011; Vlach, 2014; Vlach & Sandhofer, 2012a). Forgetting allows people to forget the small, insignificant details of the information they are learning to improve memory and learning in the future. Because forgetting happens so quickly, the amount of information retained can change drastically between an immediate test and a test after a delay. Although forgetting can benefit learning and memory, testing people immediately after seeing the information may not accurately measure how well information is retained in memory. Additionally, studies show that supporting learning during encoding may flatten the forgetting curve and maintain a higher level of memory for learned information (Loftus, 1985; Vander Linde et al., 1985; Vlach & Sandhofer, 2012a; White, 2001). Thus, it is important to add delays in testing to determine how effective the learning was and how long the information stays in memory.

Because it is well documented that memory for learned information quickly drops after learning, it is important to add delays when testing information. Some word learning studies tested whether children learned words immediately after learning (or with very short delays) without determining whether children retained the words they learned over time (Ard & Beverly, 2004; Blewitt & Langan, 2016; Goldenberg & Sandhofer, 2013b; Vlach & Sandhofer, 2011). Other studies that looked at retention intervals of different lengths found that children quickly forgot words, sometimes depending on how well the words were encoded (Booth, 2015; Horst et al., 2011; Munro et al., 2012; Vlach et al., 2012; Vlach & Sandhofer, 2012a; Wojcik, 2017). Testing children after a delay can also change the effectiveness of the experimental manipulation (Booth, 2015; Horst et al., 2011; Vlach et al., 2012). For example, when children were immediately tested on how many words they learned after either a simultaneous or spaced presentation, they provided more correct responses with a simultaneous presentation. However, after a 15-minute delay, children who saw simultaneously presented objects dropped to chance levels, whereas children who saw objects spaced out in time correctly remembered the labels at above chance levels (Vlach et al., 2012). Other word learning studies found that even after short delays from one to five minutes, children dropped from knowing an average of almost six words to only knowing an average of about half a word (Horst & Samuelson, 2008; Munro et al., 2012). Studies can better determine what helps children learn and retain words over time by incorporating delays into testing.

Retention in Shared Book Reading

The shared book reading literature has primarily focused on how book reading affects general vocabulary levels instead of whether children learn story-specific words from books. Of the few studies that examine the learning of story-specific words (Jimenez & Saylor, 2017; Justice, 2002; McLeod & McDade, 2011; Robbins & Ehri, 1994; Sénéchal & Cornell, 1993), there are even fewer that examine whether these words are retained after periods of time (Horst et al., 2011; McLeod & McDade, 2011; Sénéchal & Cornell, 1993). Word retention across time might be one way shared book reading improves overall vocabulary levels. However, it is impossible to determine whether children remember the words they have learned without adding time delays between when children learn the words and when they are tested. Additionally, examining the effect of time may reveal a change in the efficacy of different reading approaches. For example, one study taught children new words either across three different stories or within one story read three times (Horst et al., 2011). At the immediate test, children who learned the words hearing the same book read three times knew more words than children who heard three different stories, even though both groups were above chance. However, after a few days, only children who learned the words within the same story scored above chance in remembering the words they learned. Thus, by testing vocabulary after a delay, the researchers determined that reading the same book was more beneficial than reading different books despite initial learning in both conditions (Horst et al., 2011). Because information must be bolstered with memory supports for retention, it is important to determine how much support is needed for words to be retained by testing children's knowledge after a delay.

Overview of Studies

The studies in this dissertation uniquely combine findings from the memory, word learning, and shared book reading literature to examine which supports help children's memory of words learned in a book reading setting. Given the importance of early vocabulary development (De Temple & Snow, 2003; Sénéchal et al., 1998) and the effectiveness of early vocabulary interventions on vocabulary gains of at-risk children (Justice et al., 2005; Sénéchal et al., 1995), it is important to examine the effects of shared book reading with preschoolers who are just starting to gain literacy knowledge.

The current studies have one main goal: to characterize both the type and presentation timing of language within shared book reading that might support word learning. Because we know that variability is an important memory support, Study 1 characterizes the type of language involved in the questions asked during shared book reading. Although interactive reading is particularly effective for building children's *general* vocabulary, we do not know if it improves children's memory of the *story-specific* words learned from books. Studies suggest that books should be read multiple times to increase children's likelihood of remembering the words they hear. However, it is unclear whether the same questions should be asked every time the book is read or if children might benefit from variability in the questions they hear.

Additionally, the presentation timing of information is important such that the retention of information is strongest when information is presented spaced out in time instead of presented within a single, massed session. Study 2 determines how presentation timing affects words learned from books. Because the spacing and massing effects have not been examined in storybooks with novel words, it is unclear which would be better for word learning and retention in books.

Study 1

The goal of Study 1 was to determine how the variability of the types of questions asked in shared book reading sessions affect the learning and retention of novel, story-specific words. Specifically, I examined the effects of asking irrelevant questions, the same questions, or different questions during a shared book reading session. Because the amount of variability that is helpful changes with development, comparing three levels of variability will help determine what variability is best for preschool children learning words from books. Additionally, I tested children at three delay intervals to determine which type of questions led to longer word retention. I hypothesized that asking the same questions would help children remember more words if tested immediately after learning but asking different questions would help children retain more words if tested after a delay.

Method

Participants

A total of 38 4-year-olds (M_{age} = 4.53 years, SD_{age} = 0.37 years, 21 males, 17 females) were included in the preliminary sample. The sample size was calculated using G*Power (Faul et al., 2007). Effects of other related studies (Blewitt et al., 2009; Horst et al., 2011) are typically quite large. Detecting a large effect size (f = 0.57) required a total of 43 participants to achieve 80% power. However, because the methods in this study involve new procedures (e.g., on Zoom) for which there are no prior effect sizes available, a sample size using a more conservative Cohen's f of 0.33 (medium/large effect) indicated that detecting a medium/large effect size (f =

0.33) required a total of 114 children (38 per condition). Because the study is currently underpowered[1], all results will be preliminary and should be interpreted with caution.

The 38 participants spoke English at least 70% of the time, thus ensuring children had enough English comprehension to understand the story in the present experiment. In total, 31 participants came from an English-only household, and 36 participants were read books in English at least 90% of the time at home. On average, parents reported reading about 5 hours per week (*SD* = 3.5 hrs, range = 0.5 – 20 hrs) to their children. Parents reported that their children produced an average of 113 words (*SD* = 30) out of the 212 words on the Developmental Vocabulary Assessment for Parents (DVAP) list. Currently there are no vocabulary norms for the DVAP.

Parents reported that their children were from the following racial/ethnic groups: White (*N* = 15), Multiracial (*N* = 13), Asian (*N* = 3), Hispanic/Latinx (*N* = 3), chose not to answer (*N* = 4). Further, 25 participants had at least one parent with a graduate degree, 11 participants had at least one parent with a bachelor's degree, and one participant had at least one parent with some college education. Thirty-four of the participants had previously attended preschool. In the sample, 37 participants lived in the US and one participant lived in Japan, but the parent reported the child as a native English speaker. An additional one child was excluded from the final sample due to technical issues.

Participants were recruited through lists of birth records provided by the county and through social media sites. Parents received informed consent forms explaining the study. Only

[1] The studies in this dissertation were run during an active pandemic, and thus I had constraints for data collection. I could only run participants remotely because our lab was not allowed to conduct research in person. The procedure had to be modified and piloted so that 4-year-olds could sit through the task in an online setting, which set back data collection. It was also difficult to recruit participants as parents had to schedule sessions when they were out of work and when the kids were home from school or when they had access to a tablet or laptop. Data collection is ongoing, but this dissertation serves as a preliminary analysis of the data collected so far.

children with parent permission participated, as required by the UCLA IRB. Children were recruited and tested online and received a $5 gift card as compensation for participating.

Design and Procedure

This study was a 3 (Variability) x 3 (Test Delay) mixed design. Variability (same, different, and irrelevant questions) was a within-subjects factor controlling for individual differences in word learning. Test Delay (0 min./immediate, 5 min./short delay, and 24 hr./long delay) was a between-subjects factor to eliminate any possible testing effects (i.e., the effect of improving information recall simply by testing the information multiple times).

The study was conducted on Zoom. Before the Zoom session, parents filled out a language and demographic questionnaire and the DVAP. The experimenter read the virtual storybook during the Zoom session, lasting approximately 3 minutes. Children were then tested in one of three test delays: 0 minutes, 5 minutes, or 24 hours. Children in the 0-minute delay condition were tested immediately after the conclusion of the virtual storybook. Children in the 5-minute delay condition played an iSpy game during the five-minute delay interval. In the iSpy game, children were asked to find a picture described by the experimenter from an array of 20+ familiar images (e.g., apple, tree, ball, cat, etc.). Children in the 24-hour delay condition logged off of zoom immediately after the conclusion of the virtual storybook and were tested 24 hours later the following day.

Materials

Parent Surveys

Parents completed a language and demographic questionnaire. This questionnaire included questions regarding the child's language environment, parents' reading behaviors, and socioeconomic status.

Additionally, parents completed the DVAP (Libertus et al., 2015). Children's productive vocabulary (i.e., words they produce) was assessed using the DVAP, a checklist of 212 words that is a validated measure of children's productive vocabulary with scores that highly correlate with the Communicative Development Inventories (CDI) (coefficient of .79) and the PPVT (coefficient of .69). However, the DVAP has the advantage of testing a broader age range than the CDI, which is restricted to children aged 16-30 months, and is less time-consuming than the PPVT, which is a direct vocabulary assessment requiring individual administration. Using a language measure allowed us to control for prior language knowledge and get a clearer effect of our experimental manipulations.

Novel Words and Pictures

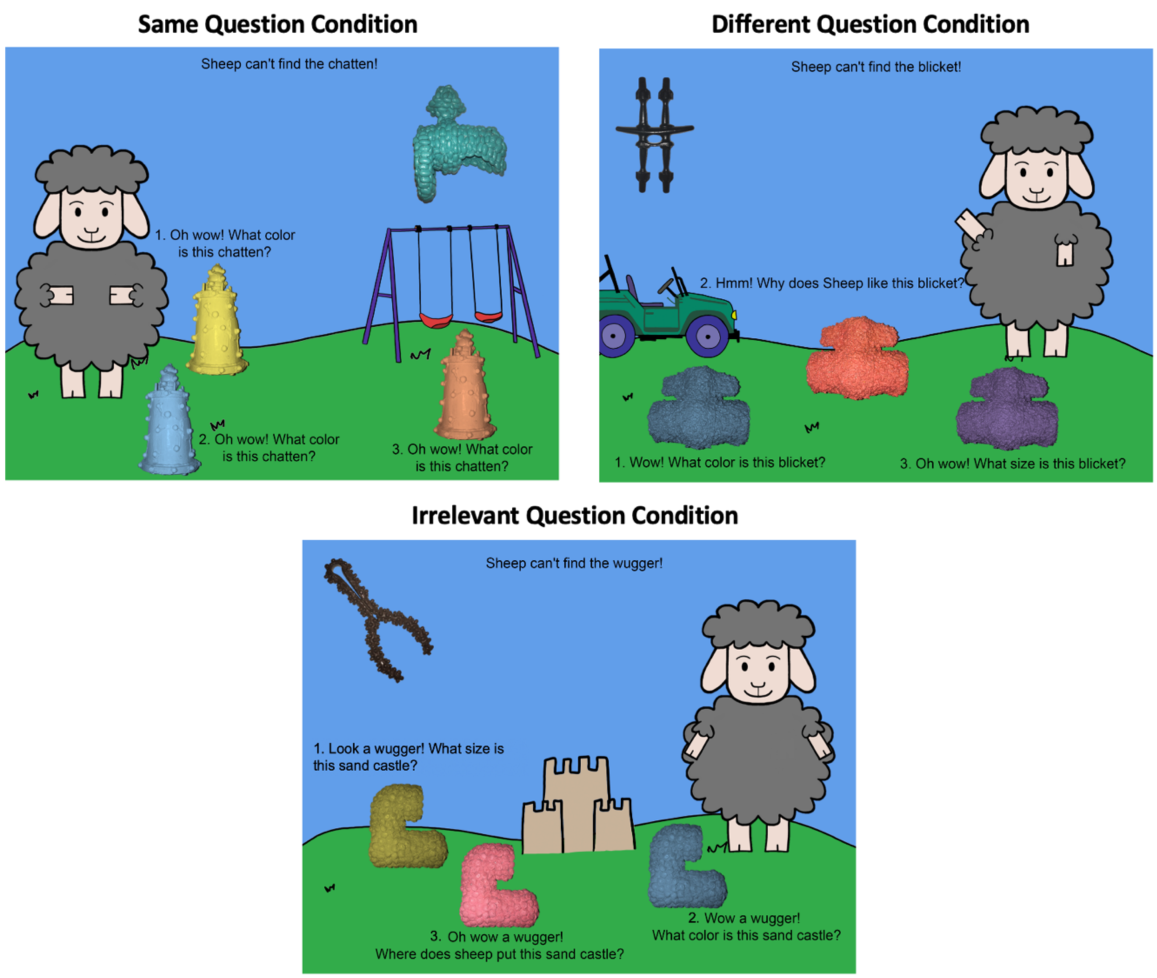
Novel words and object categories were used to control for children's prior knowledge. The novel object categories were unfamiliar shapes (Figure 1). Within each novel category, each exemplar was the same shape but differed in color so that children could generalize the category. The novel categories were paired with novel labels that sounded like possible English words. The selected words were: blicket, wugger, chatten, toma, koba, gazzer, modi, kita, and rido. Nine novel object categories were paired with one of the nine novel labels to create the target object words in the study. An additional nine objects were presented in the story without a label and served as options in the test trials. The word-object pairs were counterbalanced across orders. Additionally, nine drawings of familiar pictures were used in the story (i.e., a car, sandcastle, swings, bed, scooter, blocks, Legos, boat, and slide).

Virtual Storybook

A virtual storybook was created in Photoshop to present the stimuli to children. Children were told that the main character (i.e., Sheep) was looking for the novel objects. Each page of the virtual storybook consisted of the main character, an unlabeled object, a familiar object, and three exemplars of a target object category. Three versions of the book were created to counterbalance all the word-label pairings and the question-block order.

Figure 1

Example Book Pages for each Condition in Study 1



Question Variability Conditions

There were three within-subject question variability conditions: same, different, and irrelevant questions (Figure 1). Three word-object pairs were presented within each condition. The questions were piloted with 3-5-year-olds, and only the questions that elicited responses were used. The final questions chosen were: what color is the X, why does sheep like this X, what size is this X, where does sheep put this X. In the same question condition, word-object pairs were accompanied by the same questions each time the word was presented. For example, for the novel word-object pair “chatten,” the question “What color is this chatten?” was repeated for each of the three “chatten” exemplars. In the different question condition, word-object pairs were accompanied by different questions each time the word was presented. For example, for the novel word-object pair “blicket,” children were asked the questions, “What color is this blicket?”, “Why does Sheep like this blicket?” and “What size is this blicket?”. Finally, in the irrelevant questions condition, the object was labeled as a statement accompanied by an irrelevant question about a familiar object on the page (e.g., “Look, a wugger! What size is this sandcastle?”, “Look, a wugger! Where does Sheep put this sandcastle?”, and “Look, a wugger! What color is this sandcastle?”). Labeling these words outside of a question served as a control to compare the effects of question-asking over and above simply labeling the object while controlling for label frequency. Children’s responses to the questions asked were not recorded.

Test Trials

Before the test began, children participated in a practice test trial. Pictures of four familiar animals were placed in the four corners of the screen. There were small roman numeral numbers next to each picture for the parent to confirm the child's choice. Roman numerals were selected as they are less familiar to children and therefore less likely to distract 4-year-olds than familiar numbers or letters. Parents were asked to state the roman numeral that matched their child’s pointing choice on the screen. Children were then asked to point to each picture as they heard it labeled (e.g., “point to the cat”). This practice test trial ensured that children knew how to point at the screen and helped parents practice reporting the number corresponding to their child's selection. After the practice trial, children immediately advanced to the test trials.

The test trials were identical for all conditions. There were nine forced-choice receptive generalization trials, one for each novel word. In each trial, children saw four objects on their screen (one on each corner). The experimenter asked the child to point to an object (e.g., “point to the blicket”). The four choices within each test trial were the target object, a familiar/labeled object (an object that appeared in the book with another label), a familiar/unlabeled object (an object that appeared in the book but was never labeled), and an unfamiliar/unlabeled object (a completely novel object never seen before). The target object differed in color from the presentations in the book but was the same shape. The order of the test trials was identical to the order of the word presentation in the book, allowing for a maximum possible testing delay for each object and guaranteed that the last word learned was not the first word tested.

Results

The primary goal of Study 1 was to determine whether the type of question variability affected word learning and retention, tested at three different time delays. Because of the mixed design, a two-way mixed ANOVA was conducted with both question type (within) and test delay (between) as the predictor variables and the number of correct answers on test trials as the outcome variable. Because each question type (i.e., same, different, and irrelevant questions) had three objects each, the possible number of correct responses for each question type was three.

There was a total of 13 children in the immediate test condition, 14 in the 5-minute delay condition, and 11 in the 24-hour delay condition. Across all test delays, children scored an average of 1.34 (SD = 1.02) out of 3 words correct when accompanied by the same question, 1.37 (SD = 0.79) with different questions, and 1.45 (SD = 0.83) with irrelevant questions. Across all question types, children scored approximately 4.38 (*SD* = 1.66) out of 8 word in the 0-minute delay condition, 4.36 (*SD* = 1.55) in the 5-minute delay condition, and 3.64 (*SD* = 2.16) in the 24-hour delay condition. Overall, across all levels of the independent variables, 16% of scores were 0s, 40% were 1s, 33% were 2s, and 11% were 3s out of 3 words, suggesting a moderate level of difficulty. Table 1 shows the mean and standard deviation for the nine subgroups at the intersection of both question type and test delay. Children were not significantly different on vocabulary levels across the three test delays (F (2, 32) = 0.94, p = .401) or based on their sex (F (1, 32) = 0.023, p = .636).

Table 1

Means and Standard Deviations of Children's Target Vocabulary Score (out of 3)

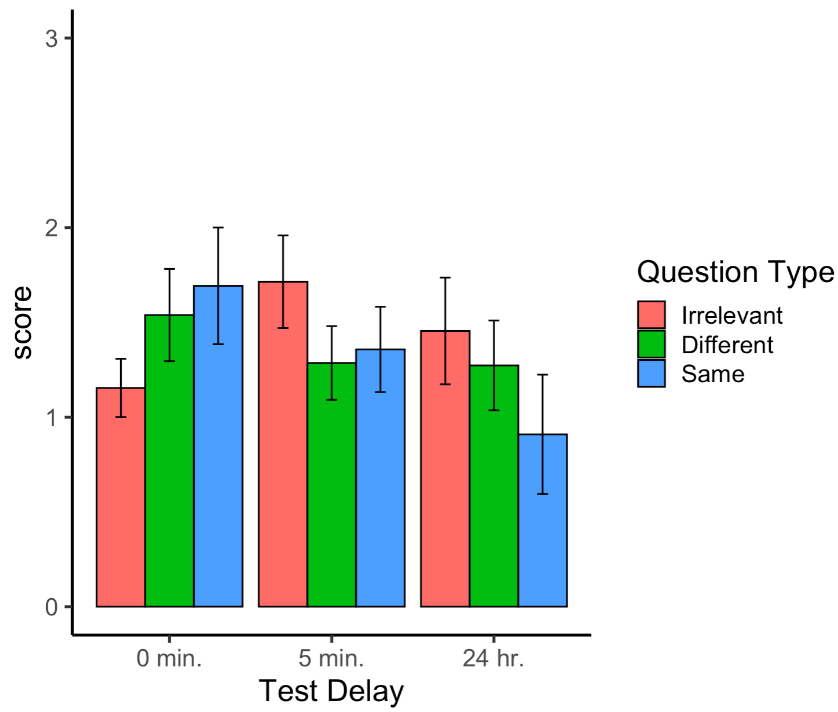
Question Type	0 min. Delay	5 min. Delay	24 hr. Delay
Irrelevant Questions	1.15 (0.56)	1.69 (0.95)	1.46 (0.93)
Different Questions	1.54 (0.88)	1.31 (0.75)	1.27 (0.79)
Same Questions	1.69 (1.11)	1.38 (0.87)	0.91 (1.04)

To determine the effect of question type and test delay, I conducted a mixed ANOVA. I first checked assumptions and found that, due to the small sample size, there was a violation of normality for six subgroups (all except Same/0min., Same/5min., & Irrelevant/24hr.). The homogeneity of variance assumption was met for all three question types (all p's > .32). Finally, the homogeneity of covariance assumption was also met (p = .81). Results (Figure 2) show that there was no significant difference between the question types (F (2,70) = 0.80, p = .797), the test delays (F (2,35) = 0.67, p = .520), or their interaction (F (4,70) = 1.99, p = .106). To determine whether age or vocabulary level affected children's performance, I ran a multilevel model with random intercepts, where question type conditions were nested within children. The

0-minute test delay and the irrelevant question condition were used as comparison groups. The equation for this model was defined as $score \sim Age + Vocabulary + QuestionType * TestDelay + (1|ID)$. This model shows that both age (t (30) = -1.57, p = .13) and vocabulary (t (30) = 0.17, p = .87) were not significant (Table 2). However, this model did reveal a significant interaction between the same question condition and the 24-hour condition. This interaction suggests that children scored higher in the same question condition compared to the irrelevant condition when tested immediately, whereas children scored higher in the irrelevant question condition compared to the same question condition when tested after 24 hours (t (70) = -2.37, p = .020).

Figure 2

Study 1 Results



Note. This figure shows the average target vocabulary score for each condition. Error bars represent the standard error of the mean.

Table 2

Multilevel Model Results for Study 1

Predictor	*β* (*SE*)	*t*	*p*
Intercept	1.03 (0.43)	2.38	.021
Age	-0.40 (0.28)	-1.45	.157
Vocabulary	0.001 (0.003)	0.26	.800
QuestionType Different	0.38 (0.31)	1.24	.218
QuestionType Same	0.54 (0.31)	1.74	.086
TestDelay 5min	0.64 (0.34)	1.88	.064
TestDelay 24hr	0.27 (0.36)	0.75	.457
QuestionType Different*TestDelay 5min	-0.81 (0.43)	-1.89	.062
QuestionType Same*TestDelay 5min	-0.90 (0.43)	-2.09	.041
QuestionType Different*TestDelay 24hr	-0.57 (0.46)	-1.24	.219
QuestionType Same*TestDelay 24hr	-1.08 (0.46)	-2.37	.020

Because neither of the independent variables was significant, I collapsed scores across all levels of the independent variables to determine whether, overall, children were scoring above chance. Chance was defined as scoring 25% because the test trials had four options to choose from. On average, children across all conditions scored higher than expected by chance (M = 46%, t (37) = 6.67, p < .001).

I also examined children's selections when they did not select the correct target object. The incorrect choices were either the familiar/labeled object (an object that appeared in the book with a different label), the familiar/unlabeled object (an object that appeared in the book but was never labeled), or the unfamiliar/unlabeled object (a completely novel object). Overall, across all conditions, children chose the correct object 46% of the time, the incorrect familiar/labeled object 32% of the time, the familiar/unlabeled object 11% of the time, and the unfamiliar/unlabeled object 11% of the time. A repeated-measures ANOVA showed that these responses were significantly different from each other (F (2.37, 87.83) = 32.48, p < .001). All pairwise comparisons except familiar/unlabeled vs. unfamiliar/unlabeled were significantly different using a Bonferroni correction. These results suggests that when children were incorrect, they were more likely to select another object that had been labeled within the story.

Discussion

Study 1 found that, overall, the type of questions asked did not affect children's memory of the new words they learned. Originally, I had intended for the irrelevant question condition to act as a control for the benefit of question asking in general. However, the results suggest that perhaps asking irrelevant questions is just as good as asking the same or different questions about a target object. Although it was unexpected that irrelevant questions would be helpful, this finding makes sense according to the dialogic reading literature that encourages readers to create an interactive reading environment (e.g., Whitehurst et al., 1988). Perhaps the important thing is to incorporate a lot of questions, no matter what kind, during shared book reading. Additionally, there was no difference in words remembered whether children were tested immediately, after 5 minutes, or after 24 hours. Initially, I had expected children to remember far fewer words if tested after 24 hours because of the quick memory declines shown in the literature (Vlach & Sandhofer, 2012a). However, finding no difference between the test delays shows that children are retaining the words they do learn during this quick shared book reading session. Finally, there was a significant interaction such that asking the same questions was better for children who tested immediately but asking irrelevant questions was better for children who tested after 24 hours. Perhaps irrelevant questions strengthen memory by creating an association between the irrelevant object and the target object, making that object more recallable after a delay.

Study 2

The goal of Study 2 was to examine how designing a storybook with two different presentation timing styles affects the learning and retention of novel, story-specific words.

Specifically, I compared pages with words presented sequentially (massed) to pages with words presented spaced in time (interleaved). To space out the words, each interleaved word was presented with the massed words in between. I tested children at three delay intervals to determine which presentation timing in the book design led to longer word retention. I hypothesized that presenting words all at once (massed) would produce higher recall when tested immediately; however, presenting words in an interleaved fashion would produce higher recall when tested after a delay.

Method

Participants

A total of 39 additional 4-year-olds (M_{age} = 4.56 years, SD_{age} = 0.29 years, 20 males, 19 females) were included in the preliminary sample for Study 2. Due to the similarity of stimuli used in both Study 1 and Study 2, a separate sample had to be collected for Study 2. The sample size was calculated using G*Power (Faul et al., 2007). Effects of other related studies (Vlach et al., 2008) are typically quite large. Detecting a large effect size (f = 1.43) required a total of 11 participants to achieve 80% power. However, because the methods in this study involve new procedures (e.g., on Zoom) for which there are no prior effect sizes available, a sample size using a more conservative Cohen's f of 0.33 (medium/large effect) indicated that detecting a medium/large effect size (f = 0.33) required a total of 90 children (30 per condition). Study 2 is currently powered to detect an effect size of f = 1.43 but is underpowered to detect the effect size of f = 0.33 selected for this study. Therefore, all results will be preliminary and should be interpreted with caution.

The 39 participants spoke English at least 70% of the time, thus ensuring children had enough English comprehension to understand the story in the present experiment. In total, 31 participants came from an English-only household, and 36 participants were read books in English at least 90% of the time at home. On average, parents reported reading 4 hours per week to their children. Parents reported that their children produced an average of 103 words (*SD* = 19) out of the 212 from the DVAP list. Currently there are no vocabulary norms for the DVAP.

Of the sample of 39 children, parents reported that their children were from the following racial/ethnic groups: White (*N* = 19), Multiracial (*N* = 13), Hispanic/Latinx (*N* = 3), Asian (*N* = 2), Black/African-American (*N* = 1), chose not to answer (*N* = 1). Further, 22 participants had at least one parent with a graduate degree, 15 participants had at least one parent with a bachelor's degree, two had at least one parent with some college, and one chose not to answer. Thirty-five of the participants had previously attended preschool, and 30 of the participants lived in the US, one lived in Australia, one lived in Canada, and one lived in the UK. An additional one child was excluded from the final sample due parental interference.

Participants were recruited through lists of birth records provided by the county and through social media sites. Parents received parent permission forms explaining the study. Only children with parent permission participated, as required by the UCLA IRB. Children were recruited and tested online and received a $5 gift card as compensation for participating.

Design and Procedure

This study was a 2 (Presentation Timing) x 3 (Test Delay) mixed design. Presentation Timing (massed and spaced) was a within-subjects factor to control for individual differences in word learning. Test Delay (0 min./immediate, 5 min./short delay, and 24 hr./long delay) was a between-subjects factor to eliminate any possible testing effects (i.e., the effect of improving information recall simply by testing the information multiple times).

Study 2 followed the same general procedure as Study 1. The study was conducted on Zoom. Before the Zoom session, parents first filled out a language and demographic questionnaire and the DVAP. During the Zoom study, the experimenter read the virtual storybook lasting approximately 3 minutes. Children were then tested in one of three test delay conditions: 0 minutes, 5 minutes, or 24 hours. Children in the 0-minute delay condition were tested immediately. Children in the 5-minute delay condition played an iSpy game during the five-minute delay interval. Children in the 24-hour delay condition logged off after the conclusion of the virtual storybook and were tested 24 hours later the following day.

Materials

Parent Surveys

The same surveys from Study 1, the language and demographic questionnaire and the DVAP, were used in Study 2.

Novel Words and Pictures

As in Study 1, novel words and object categories were used to control for children's prior knowledge. The novel object categories were unfamiliar shapes (for an example, see Figure 3). Within each novel category, each exemplar was the same shape but differed in color. The novel categories were paired with novel labels that sounded like possible English words. The selected words were: blicket, teaver, modi, koba, wugger, gazzer, toma, and sibu. Eight novel object categories were paired with one of the eight novel labels to create the target objects in the study. An additional eight objects were presented without a label and served as options in the test trials. The word-object pairs were counterbalanced across orders. Additionally, 26 drawings of familiar pictures were used (e.g., a car, sandcastle, swings, bed, etc.) in the story.

Virtual Storybook

A virtual storybook was created to present the stimuli to children. Children were told that the main character (i.e., Sheep) encountered new objects in space. Each page of the virtual storybook consisted of the main character, an unlabeled object, a familiar object, and one exemplar of a target object category (Figure 3).

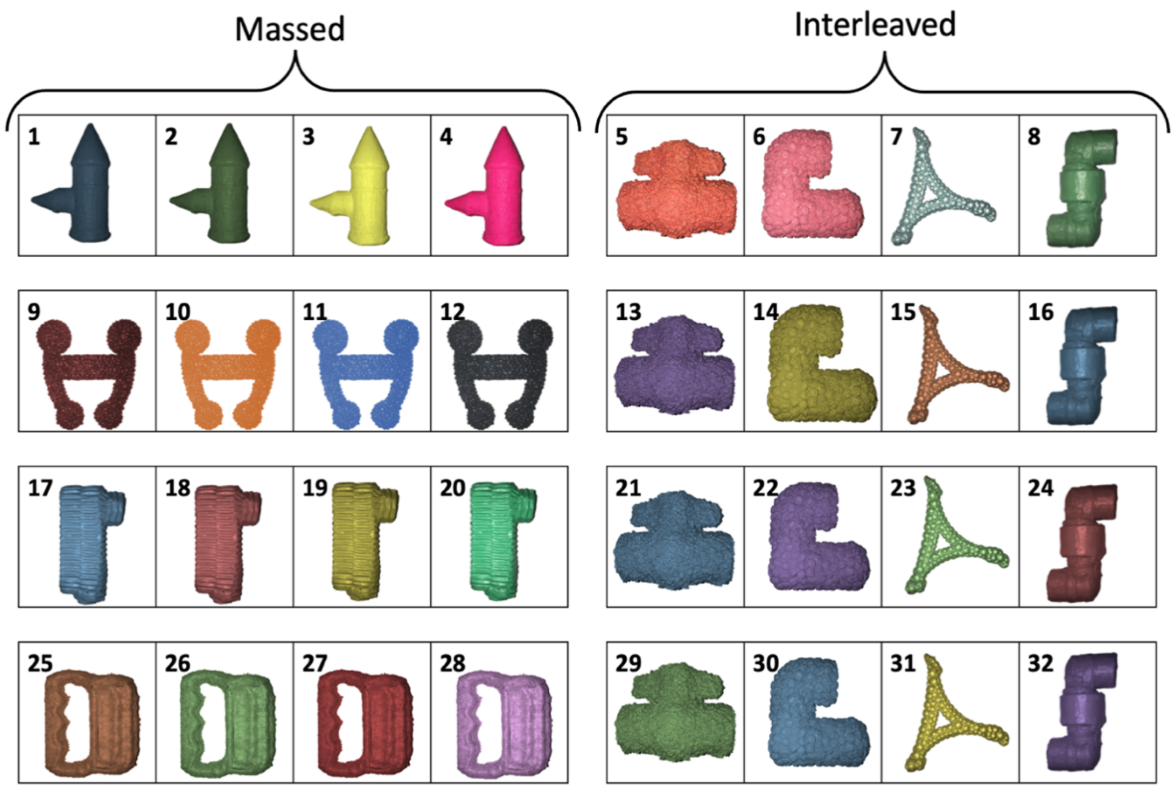
Figure 3

Example Page of the Storybook Used in Study 2

The book in Study 2 contained a total of eight novel words presented four times each. To manipulate the spacing of the words, half of the words were presented sequentially massed in time, and the other half was presented spaced in time with interleaving. Figure 4 shows the sequence of the object presentations. One object was shown four times to create a massed presentation on pages 1-4. In the following four pages, four different objects appeared as part of an interleaved presentation. Across the 32 pages, all objects appeared a total of 4 times, either presented sequentially (object on page 1 is massed on pages 1-4) or interleaved (object on page 5 is spaced on pages 5, 13, 21, 29). Three counterbalanced book versions were created to assign object-label pairings and presentation timing randomly.

Figure 4

Overview of the Object Presentation Order in Study 2



Test Trials

The test followed the same design and procedure as Study 1. The test trials were the same for all children regardless of their condition. There were eight generalization test trials, one for each word that was learned. The test trials presented the massed words first, followed by the spaced words, allowing for a maximum delay for the words spaced out in time.

Results

The primary goal of Study 2 was to determine whether presentation timing affected word learning when tested at three different time delays. Because of the mixed design, a two-way mixed ANOVA was conducted with both presentation (within) and test delay (between) as the predictor variables and the number of correct answers on test trials as the outcome variable. Because each presentation timing (i.e., massed, spaced) had four objects each, the possible number of correct responses for each presentation timing was four.

There was a total of 13 children in the immediate test condition, 13 in the 5-minute delay condition, and 13 in the 24-hour delay condition. Across all different test delays, children scored an average of 1.90 (*SD* = 1.05) out of 4 words correct when presented massed in time and 1.82 (*SD* = 1.07) when presented spaced in time. Across all different presentation timings, children scored 3.23 (*SD* = 1.30) out of 8 words in the 0-minute delay condition, 3.62 (*SD* = 1.80) in the 5-minute delay condition, and 4.31 (*SD* = 1.60) in the 24-hour delay condition. Overall, across all levels of the independent variables, 9% of scores were 0s, 30% were 1s, 34% were 2s, 21% were 3s, and 6% were 4s out of 4 words, suggesting a moderate level of difficulty. Table 3 shows the mean and standard deviation for the six subgroups at the intersection of both presentation timing and test delay. Children were not significantly different on vocabulary levels across the three test delays (F (2, 33) = 0.82, p = .451) or based on their sex (F (1, 33) = 0.10, p = .758).

Table 3

Means and Standard Deviations of Children's Target Vocabulary Score (out of 4)

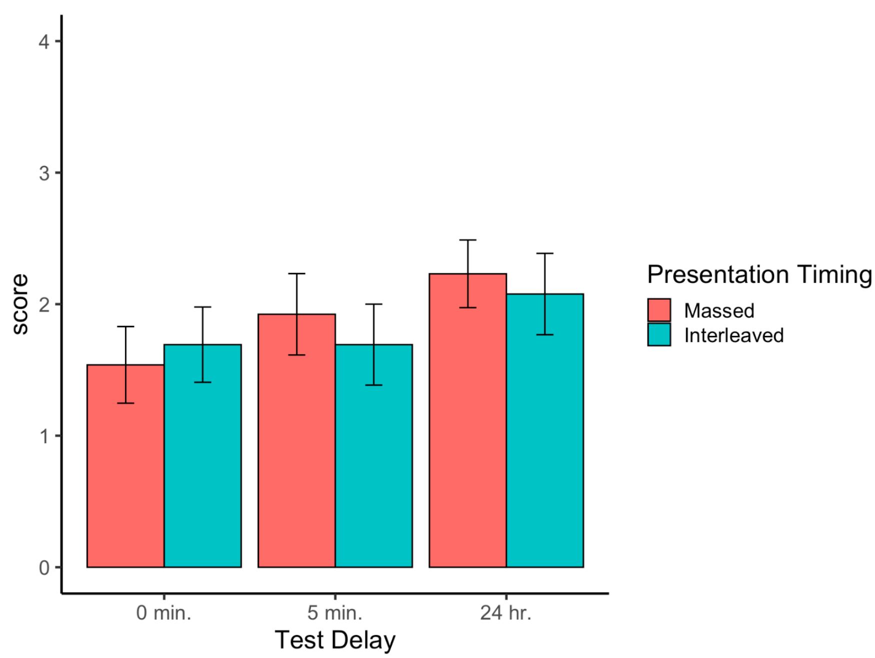
Question Type	0 min. Delay	5 min. Delay	24 hr. Delay
Massed	1.54 (1.05)	1.92 (1.12)	2.23 (0.93)
Spaced	1.69 (1.03)	1.69 (1.11)	2.08 (1.12)

To determine the effect of presentation timing and test delay, I conducted a mixed ANOVA. I first checked assumptions and found that there were no violations of normality (all p's > .08). The homogeneity of variance assumption was met for all presentation timings (all p's

> .82). Finally, the homogeneity of covariance assumption was also met (p = .90). Results (Figure 5) show that there was no significant difference between the presentation timing (F (1,36) = 0.12, p = .736), the test delays (F (2,36) = 1.55, p = .227), or their interaction (F (2,36) = 0.27, p = .765). To determine whether age or vocabulary level was significant on children's performance, I ran a multilevel model with random intercepts, where presentation timing conditions were nested within children. The 0-minute test delay and the massed presentation condition were used as comparison groups. The equation for this model was defined as $score \sim Age + Vocabulary + PresentationTiming * TestDelay + (1|ID)$. This model shows that both age (t (34) = 0.45, p = .659) and vocabulary (t (34) = 0.26, p = .798) were not significant (Table 4).

Figure 5

Study 2 Results



Note. This figure shows the average target vocabulary score for each condition. Error bars represent the standard error of the mean.

Table 4

Multilevel Model Results for Study 2

Predictor	β (*SE*)	*t*	*p*
Intercept	1.36 (0.74)	1.84	.074
Age	0.20 (0.45)	0.45	.659
Vocabulary	0.002 (0.01)	0.26	.798
Presentation Spaced	0.15 (0.39)	0.39	.697
TestDelay 5min	0.38 (0.42)	0.88	.380
TestDelay 24hr	0.68 (0.43)	1.58	.118
Presentation Spaced*TestDelay 5min	-0.38 (0.55)	-0.70	.492
Presentation Spaced*TestDelay 24hr	-0.31 (0.55)	-0.56	.582

Because neither of the independent variables was significant, I collapsed scores across all levels of the independent variables to determine whether, overall, children were scoring above chance. Chance was defined as scoring 25% because the test had four options to choose from. On average, children across all conditions scored higher than expected by chance (M = 46%, t (38) = 6.68, $p < .001$).

I also examined which items children selected when they made an incorrect selection. The incorrect choices were either the familiar/labeled object (an object that appeared in the book with a different label), the familiar/unlabeled object (an object that appeared in the book but was never labeled), or the unfamiliar/unlabeled object (a completely novel object). Overall, across all conditions, children chose the correct object 46% of the time, the incorrect familiar/labeled object 21% of the time, the familiar/unlabeled object 18% of the time, and the unfamiliar/unlabeled object 15% of the time. A repeated-measures ANOVA showed that these responses were significantly different from each other (F (2.38, 90.62) = 25.47, $p < .001$). Pairwise comparisons showed that the correct choice response was significantly different than the other three categories using a Bonferroni correction. Again, this suggests that when children were incorrect, they were more likely to select another object that had been labeled within the story.

Discussion

Study 2 found that, overall, the presentation timing did not affect children's memory of the new words they learned during shared book reading. Although a benefit for spacing information out in time has been found in prior work with children (e.g., Vlach, 2014), it seems as though preschoolers do not always benefit from learning information spaced in time (Benitez et al., 2020; Vlach & DeBrock, 2017). Perhaps word learning from shared book reading at four years old is another context in which spacing is not beneficial. Additionally, there was no difference in words remembered whether children were tested immediately, after 5 minutes, or after 24 hours. Like in Study 1, children in Study 2 are retaining the words they learned during a quick three-minute reading session even when tested after 24 hours.

General Discussion

Little is known about the mechanisms that produce vocabulary learning and retention from shared book reading. Taken together, the goal of this dissertation was to determine how variability and spacing support the memory of learned words from storybooks. In Study 1, I found no significant main effect of variability in the questions asked during the shared book reading session. I also found no significant main effect of testing delay, suggesting that children remember the same number of words across the three time delays. Interestingly, the multilevel model did reveal a significant interaction between the same question condition and the 24-hour condition where children remembered more words if they were asked the same questions compared to irrelevant questions when tested immediately, but children remembered more words if they were asked irrelevant questions compared to the same questions when tested after 24 hours. Because the study is underpowered, this interaction should be interpreted with caution. In Study 2, I found no significant main effect of the way words are presented, either massed or interleaved, and no significant main effect of testing delay. Additionally, in both studies, neither age nor vocabulary level was a significant predictor of performance. It is important to note that these results are only preliminary as the study design is currently underpowered and data collection is ongoing.

Although correlational studies have linked book reading to improved language, academic, and cognitive outcomes, recent meta-analyses demonstrate that the effects of shared book reading on language may not be as strong as previously reported (Dowdall et al., 2019; Noble et al., 2019). Additionally, the effects of shared book reading on language outcomes disappear when active control groups are used instead of "business as usual" control groups (i.e., having teachers or parents read how they normally do) (Noble et al., 2019). These weak effect sizes found in recent meta-analyses are most likely occurring due to the highly variable methodology among book reading studies, as well as a lack of active control groups in the literature. There is a lot of variability among methodological choices, such as the number of reading repetitions, number of words taught, type of vocabulary test used, length of delay between learning and test, and type of control group. Studies suggest introducing longer interventions (e.g., 6-12 months instead of 6-8 weeks) to examine how interventions realistically affect language outcomes (e.g., Noble et al., 2019). However, without first determining which shared book reading factors lead to better word learning, these longer, time-intensive interventions might not be as effective as they could be. Thus, randomized controlled trials with active control groups are needed to determine what features of shared book reading help word learning. This dissertation begin to fill some of the gaps by determining how reading sessions should be structured in terms of question variability and word presentation style.

Because of the Covid-19 pandemic, these studies had to be modified to be conducted online. One consequence of doing so was a reduction in the number of trials, and thus I had to reduce the number of target words because children were unable to sit through the longer task on Zoom. For example, Study 1 was originally designed to have 15 target words instead of 9. Having 15 target words would allow a child to score one of six possible values (0-5) per question type. This change resulted in children scoring one of four possible values (0-3). Although both dependent variables are treated as ordinal due to lack of continuity, with a high number of categories, ordinal numbers can be treated as quasi-interval, numeric variables. Because Study 1 only has four categories and Study 2 has five categories, once additional data is collected, these studies will be analyzed using a cumulative logit model that treats the dependent variable as an ordered category instead of a continuous number. This analytic strategy will allow data analysis using a more powerful statistical method that better fits the structure of the data.

If we examine child performance without considering the independent variables, children in both studies remembered about four of the nine (Study 1) or four of the eight (Study 2) target words (46% in both Study 1 and Study 2). Children in both studies remembered more words than would be predicted by chance, and they scored higher than in previous shared book reading studies. Prior studies show that children, on average, learn about one word (17%) after reading a storybook once (Jimenez & Saylor, 2017) and three words (33%) after reading a storybook twice (Justice, 2002). Because test delay was not significant, this finding suggests that children remember the same number of words no matter if they are tested immediately, after 5 minutes, or after 24 hours. Prior work shows that even short delays in testing (1-5 minutes) decrease the number of words that preschoolers remember from a mean of six words to a mean of about half a word (Horst et al., 2011; Horst & Samuelson, 2008; Munro et al., 2012). Children in the current studies remember more words than in prior studies, and they are retaining these words across a longer time delay, possibly indicating that their memory for the target words is being supported in some way during these short highly interactive 3-minute reading sessions.

So far, these preliminary analyses show no statistically significant differences in performance when considering question variability, word presentation timing, and test delay across both Study 1 and Study 2. There are a few statistical possibilities as to why there currently is not a significant effect. First, as mentioned previously, the studies are currently underpowered (data collection is about 36% complete). Once I reach the expected power, the results could change. Second, if the results remain non-significant, it could be that the effects of these variables need to be examined separately as between-subject variables instead of within-subjects variables. Manipulating both variability and presentation timing between-subjects would allow us to include more test trials within each level of the independent variables, allowing for better discrimination of the effect over the noise of the data.

Nevertheless, it is also possible that there are no significant differences and perhaps variability and spacing are not helpful in this shared book reading context. One possibility for nonsignificant effects in this study may be due to methodological differences between the current and prior studies. Prior work that shows high levels of word retention (70% after about a week) incorporates repetition, such that only children who hear the same story multiple times can remember the words they learn after a week-long delay (Horst et al., 2011). In the current studies, the book was only read once, and it lasted approximately three minutes. Perhaps repeated reading sessions are needed for variability and spacing to show effects on learning.

Furthermore, because of the pandemic, the current studies had to be moved to an online format, and it is possible that this may have affected young children's word learning, especially if children were used to in-person shared book reading sessions. Indeed, studies have reported deficits in young children's learning via media (e.g., Barr, 2013). However, it is unlikely that the online format alone is responsible for the present findings. Most studies of media-based learning find that learning deficits decrease as children get older (Myers et al., 2017) or by adding some learning supports like repetition (Barr, 2013), attention-direction cues (Neuman et al., 2019), and social contingency through video chat (Barr, 2019; Myers et al., 2017). Additionally, recent studies find that effects of language interventions and cognitive assessments found in in-person studies replicate online (Bambha & Casasola, 2021; Lapidow et al., 2021). Given that the current studies were done via video chat, and we hovered the mouse over the object while labeling it, children should perform similarly as they would in person. However, without directly testing children in person, it is difficult to determine whether the present findings reflect deficits from the learning environment.

Finally, both Study 1 and Study 2 yielded surprising preliminary results. In Study 1, I did not see an effect of question type, which was unexpected, especially for the irrelevant question condition. Although it may seem counterintuitive that asking a question about an irrelevant object would be helpful, this is in line with the dialogic reading literature that values a highly interactive reading environment (e.g., Whitehurst et al., 1988). It is possible that the types of questions asked during shared book reading are not as important as having extratextual talk during the reading session (Ard & Beverly, 2004; Hindman et al., 2019; Zucker et al., 2010). This finding is aligned with other work suggesting that simply labeling an object leads to more learning than asking questions (Justice, 2002; Sénéchal & Cornell, 1993). In Study 2, I did not find an effect of spacing, which is also surprising but supported in the preschool literature. Prior work has found that the spacing effect may not always be helpful for preschoolers (Toppino & DiGeorge, 1984), especially when learning categories based on features like texture or color (Slone & Sandhofer, 2017). Other studies find that young children often perform equally well both in massed and interleaved conditions (Benitez et al., 2020; Vlach & DeBrock, 2017). If this finding holds, then shared book reading could be another context in which children learn equally well no matter how the information is presented, perhaps because the information is embedded in a rich, interactive learning environment.

These two studies integrate shared book reading, memory, and word learning research to improve child-reader interactions and the design of children's books to help children learn and retain words. Study 1 experimentally isolated and compared the effect of variability on the questions asked during shared book reading. Study 2 used a robust memory presentation style (i.e., spacing) to determine whether its effects extended to shared book reading. By examining the factors that best help children learn and retain words, research could be particularly useful for children from low socioeconomic backgrounds and linguistically diverse environments. Because some evidence suggests that the effects of shared book reading are smaller for children from disadvantaged backgrounds (Manz et al., 2010; Mol et al., 2008), interventions must be introduced earlier and structured to create more robust vocabulary learning. By understanding the mechanisms that produce vocabulary learning and retention from shared book reading, future interventions can improve both the quality of book design and reader-child interactions to help children's vocabulary development.

References

Anderson, A., Anderson, J., Lynch, J., Shapiro, J., & Eun Kim, J. (2012). Extra-textual talk in shared book reading: A focus on questioning. *Early Child Development and Care*, *182*(9), 1139–1154. https://doi.org/10.1080/03004430.2011.602189

Ard, L. M., & Beverly, B. L. (2004). Preschool word learning during joint book reading: Effect of adult questions and comments. *Communication Disorders Quarterly*, *26*(1), 17–28. https://doi.org/10.1177/15257401040260010101

Arnold, D. H., Lonigan, C. J., Whitehurst, G. J., & Epstein, J. N. (1994). Accelerating language development through picture book reading: Replication and extension to a videotape training format. *Journal of Educational Psychology*, *86*(2), 235–243. https://doi.org/10.1037/0022-0663.86.2.235

Bailey, A. L., & Moughamian, A. C. (2007). Telling stories their way: Narrative scaffolding with emergent readers and readers. *Narrative Inquiry*, *17*(2), 203–229. https://doi.org/10.1075/ni.17.2.04bai

Bailey, A. L., Moughamian, A., Kelly, K., Mccabe, A., & Huang, B. (2018). Leap-frog to literacy: Maternal narrative supports differentially relate to child oral language and later reading outcomes. *Early Child Development and Care*, *190*, 1–14. https://doi.org/10.1080/03004430.2018.1521807

Bambha, V. P., & Casasola, M. (2021). From Lab to Zoom: Adapting Training Study Methodologies to Remote Conditions. *Frontiers in Psychology*, *12*, 694728. https://doi.org/10.3389/fpsyg.2021.694728

Barr, R. (2013). Memory Constraints on Infant Learning From Picture Books, Television, and Touchscreens. *Child Development Perspectives*, *7*(4), 205–210. https://doi.org/10.1111/cdep.12041

Barr, R. (2019). Growing Up in the Digital Age: Early Learning and Family Media Ecology. *Current Directions in Psychological Science*, *28*(4), 341–346. https://doi.org/10.1177/0963721419838245

Benitez, V. L., Zettersten, M., & Wojcik, E. (2020). The temporal structure of naming events differentially affects children's and adults' cross-situational word learning. *Journal of Experimental Child Psychology*, *200*, 104961. https://doi.org/10.1016/j.jecp.2020.104961

Blewitt, P., & Langan, R. (2016). Learning words during shared book reading: The role of extratextual talk designed to increase child engagement. *Journal of Experimental Child Psychology*, *150*, 404–410. https://doi.org/10.1016/j.jecp.2016.06.009

Blewitt, P., Rump, K. M., Shealy, S. E., & Cook, S. A. (2009). Shared book reading: When and how questions affect young children's word learning. *Journal of Educational Psychology*, *101*(2), 294–304. https://doi.org/10.1037/a0013844

Bojczyk, K. E., Davis, A. E., & Rana, V. (2016). Mother–child interaction quality in shared book reading: Relation to child vocabulary and readiness to read. *Early Childhood Research Quarterly*, *36*, 404–414. https://doi.org/10.1016/j.ecresq.2016.01.006

Booth, A. E. (2015). Effects of causal information on the early word learning: Efficiency and longevity. *Cognitive Development*, *33*, 99–107. https://doi.org/10.1016/j.cogdev.2014.05.001

Bower, G. H. (1981). Mood and memory. *American Psychologist*, *36*(2), 129–148. https://doi.org/10.1037/0003-066X.36.2.129

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. https://doi.org/10.1037/0033-2909.132.3.354

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology*, *38*(6), 967–978. https://doi.org/10.1037/0012-1649.38.6.967

De Temple, J., & Snow, C. E. (2003). Learning words from books. In A. van Kleeck, S. A. Stahl, & E. B. Bauer (Eds.), *On reading books to children: Parents and teachers* (pp. 16–36). Lawrence Erlbaum Associates, Inc.

Deckner, D. F., Adamson, L. B., & Bakeman, R. (2006). Child and maternal contributions to shared reading: Effects on language and literacy development. *Journal of Applied Developmental Psychology*, *27*(1), 31–41. https://doi.org/10.1016/j.appdev.2005.12.001

Deshmukh, R. S., Zucker, T. A., Tambyraja, S. R., Pentimonti, J. M., Bowles, R. P., & Justice, L. M. (2019). Teachers' use of questions during shared book reading: Relations to child responses. *Early Childhood Research Quarterly*, *49*, 59–68. https://doi.org/10.1016/j.ecresq.2019.05.006

Dickinson, D. K., & Smith, M. W. (1994). Long-term effects of preschool teachers' book readings on low-income children's vocabulary and story comprehension. *Reading Research Quarterly*, *29*(2), 104. https://doi.org/10.2307/747807

Dowdall, N., Melendez-Torres, G. J., Murray, L., Gardner, F., Hartford, L., & Cooper, P. J. (2019). Shared picture book reading interventions for child language development: A systematic review and meta-analysis. *Child Development*. https://doi.org/10.1111/cdev.13225

Dulsky, S. G. (1935). The effect of a change of background on recall and relearning. *Journal of Experimental Psychology*, *18*(6), 725–740. https://doi.org/10.1037/h0058066

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test (PPVT-IV). Ed Fourth*. Pearson Assessments Minneapolis, MN:

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussenius, Trans.). Dover Publications. (Original work published 1885).

Eich, E., & Macaulay, D. (2000). Are real moods required to reveal mood-congruent and mood-dependent memory? *Psychological Science*, *11*(3), 244–248. https://doi.org/10.1111/1467-9280.00249

Evans, M. A., Shaw, D., & Bell, M. (2000). Home literacy activities and their influence on early literacy skills. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *54*(2), 65–75. https://doi.org/10.1037/h0087330

Fanselow, M. S., & Tighe, T. J. (1988). Contextual conditioning with massed versus distributed unconditional stimuli in the absence of explicit conditional stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, *14*(2), 187–199. https://doi.org/10.1037/0097-7403.14.2.187

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

Fawcett, J. M., & Hulbert, J. C. (2020). The many faces of forgetting: Toward a constructive view of forgetting in everyday life. *Journal of Applied Research in Memory and Cognition*, *9*(1), 1–18. https://doi.org/10.1016/j.jarmac.2019.11.002

Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition*, *7*(2), 95–112. https://doi.org/10.3758/BF03197590

Gluckman, M., Vlach, H. A., & Sandhofer, C. M. (2014). Spacing simultaneously promotes multiple forms of learning in children's science curriculum: Spacing, memory, and generalization. *Applied Cognitive Psychology*, *28*(2), 266–273. https://doi.org/10.1002/acp.2997

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, *66*(3), 325–331. https://doi.org/10.1111/j.2044-8295.1975.tb01468.x

Goldenberg, E. R., & Sandhofer, C. M. (2013a). Who is she? Changes in the person context affect categorization. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00745

Goldenberg, E. R., & Sandhofer, C. M. (2013b). 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

Haden, C. A., Reese, E., & Fivush, R. (1996). Mothers' extratextual comments during storybook reading: Stylistic differences over time and across texts. *Discourse Processes*, *21*(2), 135–169. https://doi.org/10.1080/01638539609544953

Han, M., Moore, N., Vukelich, C., & Buell, M. (2010). Does Play Make a Difference? How Play Intervention Affects the Vocabulary Learning of At-Risk Preschoolers. *American Journal of Play*, *3*(1), 82–105.

Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly*, *15*(1), 75–90. https://doi.org/10.1016/S0885-2006(99)00038-1

Harris, M., Barrett, M., Jones, D., & Brookes, S. (1988). Linguistic input and early word meaning. *Journal of Child Language*, *15*(1), 77–94. https://doi.org/10.1017/S030500090001206X

Hills, T. T., Maouene, J., Riordan, B., & Smith, L. B. (2010). The associative structure of language: Contextual diversity in early word learning. *Journal of Memory and Language*, *63*(3), 259–273. https://doi.org/10.1016/j.jml.2010.06.002

Hindman, A. H., Connor, C. M., Jewkes, A. M., & Morrison, F. J. (2008). Untangling the effects of shared book reading: Multiple factors and their associations with preschool literacy outcomes. *Early Childhood Research Quarterly*, *23*(3), 330–350. https://doi.org/10.1016/j.ecresq.2008.01.005

Hindman, A. H., Wasik, B. A., & Bradley, D. E. (2019). How classroom conversations unfold: Exploring teacher–child exchanges during shared book reading. *Early Education and Development*, *30*(4), 478–495. https://doi.org/10.1080/10409289.2018.1556009

Hindman, A. H., Wasik, B. A., & Erhart, A. C. (2012). Shared book reading and head start preschoolers' vocabulary learning: The role of book-related discussion and curricular connections. *Early Education & Development*, *23*(4), 451–474. https://doi.org/10.1080/10409289.2010.537250

Hintzman, D. L., Summers, J. J., & Block, R. A. (1975). Spacing judgments as an index of study-phase retrieval. *Journal of Experimental Psychology: Human Learning and Memory*, *1*(1), 31–40. https://doi.org/10.1037/0278-7393.1.1.31

Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development*, *73*(2), 418–433. https://doi.org/10.1111/1467-8624.00415

Horst, J. S., Parsons, K. L., & Bryan, N. M. (2011). Get the story straight: Contextual repetition promotes word learning from storybooks. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00017

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., Scott, E. J., & Pollard, J. A. (2010). The role of competition in word learning via referent selection. *Developmental Science*, *13*(5), 706–713. https://doi.org/10.1111/j.1467-7687.2009.00926.x

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*(3), 269–299. https://doi.org/10.1006/jmps.2001.1388

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236–248. https://doi.org/10.1037/0012-1649.27.2.236

Jimenez, S. R., & Saylor, M. M. (2017). Preschoolers' word learning and story comprehension during shared book reading. *Cognitive Development*, *44*, 57–68. https://doi.org/10.1016/j.cogdev.2017.08.011

Johnson, E. K., McQueen, J. M., & Huettig, F. (2011). Toddlers' Language-Mediated Visual Search: They Need not have the Words for It. *Quarterly Journal of Experimental Psychology*, *64*(9), 1672–1682. https://doi.org/10.1080/17470218.2011.594165

Justice, L. M. (2002). Word exposure conditions and preschoolers' novel word learning during shared storybook reading. *Reading Psychology*, *23*(2), 87–106. https://doi.org/10.1080/027027102760351016

Justice, L. M., Meier, J., & Walpole, S. (2005). Learning new words from storybooks: An efficacy study with at-risk kindergartners. *Language, Speech, and Hearing Services in Schools*, *36*(1), 17–32. https://doi.org/10.1044/0161-1461(2005/003)

Kalb, G., & van Ours, J. C. (2014). Reading to young children: A head-start in life? *Economics of Education Review*, *40*, 1–24. https://doi.org/10.1016/j.econedurev.2014.01.002

Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the "enemy of induction"? *Psychological Science*, *19*(6), 585–592. https://doi.org/10.1111/j.1467-9280.2008.02127.x

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*(3), 299–321. https://doi.org/10.1016/0885-2014(88)90014-7

Lapidow, E., Tandon, T., Goddu, M., & Walker, C. M. (2021). A Tale of Three Platforms: Investigating Preschoolers' Second-Order Inferences Using In-Person, Zoom, and Lookit Methodologies. *Frontiers in Psychology*, *12*. https://www.frontiersin.org/article/10.3389/fpsyg.2021.731404

Libertus, M. E., Odic, D., Feigenson, L., & Halberda, J. (2015). A Developmental Vocabulary Assessment for Parents (DVAP): Validating Parental Report of Vocabulary Size in 2- to

7-Year-Old Children. *Journal of Cognition and Development*, *16*(3), 442–454. https://doi.org/10.1080/15248372.2013.835312

Loftus, G. R. (1985). Evaluating forgetting curves. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(2), 397–406. https://doi.org/10.1037/0278-7393.11.2.397

Manz, P. H., Hughes, C., Barnabas, E., Bracaliello, C., & Ginsburg-Block, M. (2010). A descriptive review and meta-analysis of family-based emergent literacy interventions: To what extent is the research applicable to low-income, ethnic-minority or linguistically-diverse young children? *Early Childhood Research Quarterly*, *25*(4), 409–431. https://doi.org/10.1016/j.ecresq.2010.03.002

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*(3), F9–F16. https://doi.org/10.1111/j.1467-7687.2008.00671.x

McDonough, C., Song, L., Hirsh-Pasek, K., Golinkoff, R. M., & Lannon, R. (2011). An image is worth a thousand words: Why nouns tend to dominate verbs in early word learning: Imageability and early word learning. *Developmental Science*, *14*(2), 181–189. https://doi.org/10.1111/j.1467-7687.2010.00968.x

McLeod, A. N., & McDade, H. L. (2011). Preschoolers' incidental learning of novel Words during storybook reading. *Communication Disorders Quarterly*, *32*(4), 256–266. https://doi.org/10.1177/1525740109354777

Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior*, *9*(5), 596–606.

Mol, S. E., Bus, A. G., de Jong, M. T., & Smeets, D. J. H. (2008). Added value of dialogic parent–child book readings: A meta-analysis. *Early Education and Development*, *19*(1), 7–26. https://doi.org/10.1080/10409280701838603

Montag, J. L., Jones, M. N., & Smith, L. B. (2015). The words children hear: Picture books and the statistics for language learning. *Psychological Science*, *26*(9), 1489–1496. https://doi.org/10.1177/0956797615594361

Munro, N., Baker, E., McGregor, K., Docking, K., & Arculi, J. (2012). Why word learning is not fast. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00041

Myers, L. J., LeWitt, R. B., Gallo, R. E., & Maselli, N. M. (2017). Baby FaceTime: Can toddlers learn from online video chat? *Developmental Science*, *20*(4), e12430. https://doi.org/10.1111/desc.12430

Naigles, L. R., & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs? Effects of input frequency and structure on children's early verb use. *Journal of Child Language*, *25*(1), 95–120. https://doi.org/10.1017/S0305000997003358

Neuman, S. B., Wong, K. M., Flynn, R., & Kaefer, T. (2019). Learning vocabulary from educational media: The role of pedagogical supports for low-income preschoolers. *Journal of Educational Psychology*, *111*(1), 32–44. https://doi.org/10.1037/edu0000278

Noble, C., Sala, G., Peter, M., Lingwood, J., Rowland, C., Gobet, F., & Pine, J. (2019). The impact of shared book reading on children's language skills: A meta-analysis. *Educational Research Review*, *28*, 100290. https://doi.org/10.1016/j.edurev.2019.100290

Overton, D. A. (1966). State-dependent learning produced by depressant and atropine-like drugs. *Psychopharmacologia*, *10*(1), 6–31. https://doi.org/10.1007/BF00401896

Perry, L. K., & Saffran, J. R. (2017). Is a Pink Cow Still a Cow? Individual Differences in Toddlers' Vocabulary Knowledge and Lexical Representations. *Cognitive Science*, *41*(4), 1090–1105. https://doi.org/10.1111/cogs.12370

Petersen, R. C. (1977). Retrieval failures in alcohol state-dependent learning. *Psychopharmacology*, *55*(2), 141–146. https://doi.org/10.1007/BF01457849

Pollard-Durodola, S. D., Gonzalez, J. E., Simmons, D. C., Kwok, O., Taylor, A. B., Davis, M. J., Kim, M., & Simmons, L. (2011). The Effects of an Intensive Shared Book-Reading Intervention for Preschool Children at Risk for Vocabulary Delay. *Exceptional Children*, *77*(2), 161–183. https://doi.org/10.1177/001440291107700202

Raikes, H., Alexander Pan, B., Luze, G., Tamis-LeMonda, C. S., Brooks-Gunn, J., Constantine, J., Banks Tarullo, L., Abigail Raikes, H., & Rodriguez, E. T. (2006). Mother-child bookreading in low-income families: Correlates and outcomes during the first three years of life. *Child Development*, *77*(4), 924–953. https://doi.org/10.1111/j.1467-8624.2006.00911.x

Rea, C. P., & Modigliani, V. (1987). The spacing effect in 4- to 9-year-old children. *Memory & Cognition*, *15*(5), 436–443. https://doi.org/10.3758/BF03197733

Reese, E., & Cox, A. (1999). Quality of adult book reading affects children's emergent literacy. *Developmental Psychology*, *35*(1), 20–28. https://doi.org/10.1037/0012-1649.35.1.20

Robbins, C., & Ehri, L. C. (1994). Reading storybooks to kindergartners helps them learn new vocabulary words. *Journal of Educational Psychology*, *86*(1), 54–64. https://doi.org/10.1037/0022-0663.86.1.54

Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review*, *103*(4), 734–760. https://doi.org/10.1037/0033-295X.103.4.734

Samuelson, L. K., Kucker, S. C., & Spencer, J. P. (2017). Moving Word Learning to a Novel Space: A Dynamic Systems View of Referent Selection and Retention. *Cognitive Science*, *41*(S1), 52–72. https://doi.org/10.1111/cogs.12369

Schwartz, R. G., & Terrell, B. Y. (1983). The role of input frequency in lexical acquisition. *Journal of Child Language*, *10*(1), 57–64. https://doi.org/10.1017/S0305000900005134

Sénéchal, M., & Cornell, E. H. (1993). Vocabulary acquisition through shared reading experiences. *Reading Research Quarterly*, *28*(4), 360. https://doi.org/10.2307/747933

Sénéchal, M., Lefevre, J.-A., Thomas, E. M., & Daley, K. E. (1998). Differential effects of home literacy experiences on the development of oral and written language. *Reading Research Quarterly*, *33*(1), 96–116. https://doi.org/10.1598/RRQ.33.1.5

Sénéchal, M., Thomas, E., & Monker, J.-A. (1995). Individual differences in 4-year-old children's acquisition of vocabulary during storybook reading. *Journal of Educational Psychology*, *87*(2), 218–229. https://doi.org/10.1037/0022-0663.87.2.218

Shiffrin, R. M., & Atkinson, R. C. (1969). Storage and retrieval processes in long-term memory. *Psychological Review*, *76*(2), 179–193. https://doi.org/10.1037/h0027277

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

Smith, S. M., Glenberg, A., & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition*, *6*(4), 342–353. https://doi.org/10.3758/BF03197465

Steyvers, M., & Malmberg, K. J. (2003). The effect of normative context variability on recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(5), 760–766. https://doi.org/10.1037/0278-7393.29.5.760

Storm, B. C. (2011). The benefit of forgetting in thinking and remembering. *Current Directions in Psychological Science*, *20*(5), 291–295. https://doi.org/10.1177/0963721411418469

Thios, S. J., & D'Agostino, P. R. (1976). Effects of repetition as a function of study-phase retrieval. *Journal of Verbal Learning and Verbal Behavior*, *15*(5), 529–536. https://doi.org/10.1016/0022-5371(76)90047-5

Toppino, T. C., & DiGeorge, W. (1984). The spacing effect in free recall emerges with development. *Memory & Cognition*, *12*(2), 118–122. https://doi.org/10.3758/BF03198425

Toppino, T. C., Kasserman, J. E., & Mracek, W. A. (1991). The effect of spacing repetitions on the recognition memory of young children and adults. *Journal of Experimental Child Psychology*, *51*(1), 123–138. https://doi.org/10.1016/0022-0965(91)90079-8

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*(5), 352–373. https://doi.org/10.1037/h0020071

Tzeng, O. J., Lee, A. T., & Wetzel, C. D. (1979). Temporal coding in verbal information processing. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(1), 52–64. https://doi.org/10.1037/0278-7393.5.1.52

Underwood, B. J. (1961). Ten years of massed practice on distributed practice. *Psychological Review*, *68*(4), 229–247. https://doi.org/10.1037/h0047516

Vander Linde, E., Morrongiello, B. A., & Rovee-Collier, C. (1985). Determinants of retention in 8-week-old infants. *Developmental Psychology*, *21*(4), 601–613. https://doi.org/10.1037/0012-1649.21.4.601

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., 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., & DeBrock, C. A. (2017). Remember dax? Relations between children's cross-situational word learning, memory, and language abilities. *Journal of Memory and Language*, *93*, 217–230. https://doi.org/10.1016/j.jml.2016.10.001

Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition*, *127*(3), 375–382. https://doi.org/10.1016/j.cognition.2013.02.015

Vlach, H. A., & Kalish, C. W. (2014). Temporal dynamics of categorization: Forgetting as the basis of abstraction and generalization. *Frontiers in Psychology*, *5*. https://www.frontiersin.org/article/10.3389/fpsyg.2014.01021

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. (2012a). 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

Vlach, H. A., & Sandhofer, C. M. (2012b). Distributing learning over time: The spacing effect in children's acquisition and generalization of science concepts: spacing and generalization. *Child Development*, *83*(4), 1137–1144. https://doi.org/10.1111/j.1467-8624.2012.01781.x

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

Vuattoux, D., Japel, C., Dion, E., & Dupéré, V. (2014). Targeting the Specific Vocabulary Needs of At-Risk Preschoolers: A Randomized Study of the Effectiveness of an Educator-Implemented Intervention. *Prevention Science*, *15*(2), 156–164. https://doi.org/10.1007/s11121-013-0379-5

Walsh, B. A., & Blewitt, P. (2006). The effect of questioning style during storybook reading on novel vocabulary acquisition of preschoolers. *Early Childhood Education Journal*, *33*(4), 273–278. https://doi.org/10.1007/s10643-005-0052-0

Walsh, R. L., & Hodge, K. A. (2018). Are we asking the right questions? An analysis of research on the effect of teachers' questioning on children's language during shared book reading with young children. *Journal of Early Childhood Literacy*, *18*(2), 264–294. https://doi.org/10.1177/1468798416659124

Wasik, B. A., & Bond, M. A. (2001). Beyond the pages of a book: Interactive book reading and language development in preschool classrooms. *Journal of Educational Psychology*, *93*(2), 243–250. https://doi.org/10.1037/0022-0663.93.2.243

White, K. G. (2001). Forgetting functions. *Animal Learning & Behavior*, *29*(3), 193–207. https://doi.org/10.3758/BF03192887

Whitehurst, G. J., David S. Arnold, Epstein, J. N., Angell, A. L., Smith, M., & Fischel, J. E. (1994). A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology*, *30*(5), 679.

Whitehurst, G. J., Falco, F. L., Lonigan, C. J., Fischel, J. E., DeBaryshe, B. D., Valdez-Menchaca, M. C., & Caulfield, M. (1988). Accelerating language development through picture book reading. *Developmental Psychology*, *24*(4), 552–559. https://doi.org/10.1037/0012-1649.24.4.552

Wojcik, E. H. (2013). Remembering new words: Integrating early memory development into word learning. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00151

Wojcik, E. H. (2017). 2.5-year-olds' retention and generalization of novel words across short and long delays. *Language Learning and Development*, *13*(3), 300–316. https://doi.org/10.1080/15475441.2016.1275644

Zucker, T. A., Justice, L. M., Piasta, S. B., & Kaderavek, J. N. (2010). Preschool teachers' literal and inferential questions and children's responses during whole-class shared reading. *Early Childhood Research Quarterly*, *25*(1), 65–83. https://doi.org/10.1016/j.ecresq.2009.07.001