

Spacing Simultaneously Promotes Multiple Forms of Learning in Children's Science Curriculum

MAXIE GLUCKMAN¹, HALEY A. VLACH^{2*} and CATHERINE M. SANDHOFER¹

¹*Department of Psychology, University of California, Los Angeles, USA*

²*Department of Educational Psychology, University of Wisconsin, Madison, USA*

Summary: The spacing effect refers to the robust finding that long-term memory is promoted when learning events are distributed in time rather than massed in immediate succession. The current study extended research on the spacing effect by examining whether spaced learning schedules can simultaneously promote multiple forms of learning, such as memory and generalization, in the context of an educational intervention. Thirty-six early elementary school-aged children were presented with science lessons on one of three schedules: massed, clumped, and spaced. At a 1-week delayed test, children in the spaced condition demonstrated improvements in both memory and generalization, significantly outperforming children in the other conditions. However, there was no observed relationship between children's memory performance and generalization performance. The current study highlights directions for future research and contributes to a growing body of work demonstrating the benefits of spaced learning for educational curriculum. Copyright © 2014 John Wiley & Sons, Ltd.

A long history of research on human memory has sought to identify the conditions of the learning environment that promote the ability to retain information. A central finding from this work is that the *timing* of learning events may be central in supporting memory. The most robust and highly replicable observed timing phenomenon is often termed the 'spacing effect' (dating back to Ebbinghaus, 1885/1964). The current study sought to extend research on the spacing effect by (i) grounding our investigation in an educational intervention in order to further bridge psychological science and educational practices and (ii) examining if and how spaced learning schedules simultaneously support multiple forms of learning, such as memory and generalization, within the context of an educational intervention.

Recent trends in research on the spacing effect: generalization and education

The spacing effect describes the finding that distributing learning events across time promotes memory to a greater degree than massing learning events in immediate succession (Ebbinghaus, 1885/1964). Hundreds of studies, including reviews (e.g., Delaney, Verkoeijen, & Spirgel, 2010; Dempster, 1988) and meta-analyses (e.g., Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Donovan & Radosevich, 1999), have observed spacing effects in a wide variety of memory tasks. In these studies, learners are typically presented with items of information multiple times, on either a massed schedule or a spaced schedule. After a delay, learners are asked to recall the items of information that had been presented earlier in the experiment. The finding that spaced learning schedules promote memory has been observed across many timescales, from a matter of seconds to yearlong intervals (e.g., Bahrck, Bahrck, Bahrck, & Bahrck, 1993). This research has also demonstrated that spaced learning promotes memory across the lifespan, including in infancy and childhood (Childers & Tomasello, 2002; Rea &

Modigliani, 1987; Rovee-Collier, Evancio, & Earley, 1995; Toppino, 1991; Toppino & DeMesquita, 1984; Vlach, Sandhofer, & Kornell, 2008).

A more recent body of research has extended spaced and interleaved learning schedules to generalization tasks (Birnbaum, Kornell, Bjork, & Bjork, 2013; Kang & Pashler, 2012; Kornell & Bjork, 2008; Kornell, Castel, Eich, & Bjork, 2010; Rohrer, 2012; Vlach, Ankowski, & Sandhofer, 2012; Vlach & Sandhofer, 2012; Vlach et al., 2008; Wahlheim, Dunlosky, & Jacoby, 2011; Zulkipli, McLean, Burt, & Bath, 2012). Generalization tasks differ from memory tasks because these tasks present learners with varied items of information and require learners to abstract across learning events in order to generalize information to new contexts. For example, one study (Vlach et al., 2008) presented children with novel object categories, in which exemplars shared the same shape but varied in color and texture, on a massed schedule and a spaced schedule. At a delayed generalization test, children demonstrated more generalization for categories in which exemplars had been presented on a spaced schedule than categories presented on a massed schedule. Taken together, this recent body of research has demonstrated that spaced learning is a more general learning phenomenon—spaced schedules can promote multiple forms of learning, such as the memory and generalization of learned information.

Another recent trend in research on the spacing effect has been to extend this work to educationally relevant materials and contexts (Bjork, 1994; Carpenter, Pashler, & Cepeda, 2009; Kornell, 2009; Pashler, Rohrer, Cepeda, & Carpenter, 2007; Rohrer, 2009; Seabrook, Brown, & Soly, 2005; Sobel, Cepeda, & Kapler, 2011; Vlach & Sandhofer, 2012). The goal of this research has been to connect laboratory-based psychological research with education to improve the design of educational interventions (see Dempster, 1988, for a review of research directions). This growing body of work has demonstrated that spaced learning schedules contextualized in educational interventions can promote memory (e.g., Kornell, 2009; Sobel et al., 2011) and generalization (Vlach & Sandhofer, 2012).

*Correspondence to: Haley A. Vlach, Department of Educational Psychology, 1025 W. Johnson Street, Madison, WI 53706, USA.
E-mail: hvlach@wisc.edu

Although recent research has demonstrated that spaced learning schedules can promote either memory *or* generalization in the context of an educational intervention, there has yet to be a study demonstrating that spaced learning schedules can simultaneously promote both types of learning. If spaced learning promotes both memory and generalization simultaneously, it is important to understand how changes in one form of learning are related to changes in another form of learning. By one account, generalization is simply an epiphenomenon of memory (e.g., Detterman, 1993). In this case, we would expect to see a strong, positive relationship between memory performance and generalization performance, as both forms of learning stem from the same processes. By another hypothesis, spaced learning schedules support memory for vocabulary and/or facts, which in turn supports generalization. That is, improvements in memory lead to improvements in generalization. In this case, we would expect to observe a positive relationship between memory performance and generalization performance.

However, it may be that the cognitive processes underlying spaced learning contribute to memory and generalization performance in a different manner. Indeed, prior research has hypothesized that spaced learning schedules contribute to memory and generalization in a slightly different way (e.g., Kornell & Bjork, 2008; Vlach et al., 2008, 2012). For example, according to one account (Vlach et al., 2008, 2012), spaced learning promotes memory for relevant features of a category/concept, which are likely to be re-presented to the learner, and deters memory of irrelevant features of a category/concept, which are not likely to be re-presented to the learner. Consequently, when learners generalize at a later point in time, they have stronger memory for relevant information and weaker memory for irrelevant information. In this case, we would not necessarily expect to see a relationship between memory and/or generalization performance; learners would have stronger memory for some information and weaker memory for other information. In order to examine these possibilities, the current study examined memory and generalization in the context of an elementary school science curriculum.

Target educational intervention: elementary school science curriculum

The target domain for this study was an elementary school science curriculum: children's learning of food chains. We chose a food chain curriculum because this curriculum often incorporates multiple forms of learning, such as memory, simple generalization of concepts, and complex generalization of concepts (e.g., Eilam, 2002). In the case of memory, children typically learn new vocabulary terms (e.g., 'biome') and facts (e.g., what a 'carnivore' eats). In the case of simple generalization, children learn that bigger animals typically eat smaller animals and generalize this information to new food chains. This is an example of a simple generalization because children rely on the perceptual features, such as the size of creatures, to generalize information. In the case of complex generalization, children learn the concept of interdependency in food chain curricula. Interdependency describes the concept that creatures in a food chain are dependent upon each other for food and survival. In each

biome, there is an underlying structure among members—if something happens to one creature in the biome, it influences the system as a whole. These structures are often called 'food webs' and have similarities across biomes. This is an example of a complex generalization because children must abstract the underlying relational structure ('food web'), rather than a set of perceptual similarities (as in the case of simple generalization). In sum, a food chain curriculum was chosen because it affords the opportunity to remember new items of information and engage in several levels of generalization (Vlach & Sandhofer, 2012).

Current study

In the current study, school-aged children were presented with lessons about food chains on one of three learning schedules: massed, clumped, or spaced. In the massed condition, participants were presented with four lessons in immediate succession. In the clumped condition, participants were presented with half of the lessons in immediate succession and half of the lessons distributed over time. Children in the spaced condition were presented with four lessons distributed over time. Children in all conditions were given a pre-test prior to instruction and a post-test 1 week following their last lesson. The tests included memory questions (free recall, cued recall memory, and forced choice), simple generalization questions (forced choice), and complex generalization questions (forced choice).

The three learning schedules (massed, clumped, and spaced) allowed for a direct examination of the effects of lesson timing on children's memory, simple generalization of concepts, and complex generalization in concepts. In sum, this study expands upon existing psychological research by determining if spacing simultaneously promotes multiple forms of learning and examining whether these improvements in learning are related to each other. Moreover, we contextualize our examination in elementary school science curriculum, contributing to a growing body of literature demonstrating the implications of spaced learning for educational practices.

METHOD

Participants

The participants were 36 early-elementary school children ($M = 7.12$ years; first and second graders; 16 girls and 20 boys) recruited from the University Laboratory School. Children were randomly assigned to one of three conditions: 12 children were assigned to the massed condition, 12 children to the clumped condition, and 12 children to the spaced condition. An additional nine children were not included in the final group because of school absences that did not allow them to complete all sessions of the study. All children had not received prior formal instruction on food chains in school.

Design

Children were randomly assigned to one of three between-subjects conditions: massed, clumped, or spaced. Children in the massed condition were presented with all four lessons in immediate succession on a Monday. Children in the clumped condition were presented with two lessons in immediate

succession on a Monday and two lessons in immediate succession on the following day, a Tuesday, providing a combination of massing and spacing. Children in the spaced condition were presented with one lesson per day for 4 days. Therefore, children in the spaced condition were presented with one lesson on a Monday, one on a Tuesday, one on a Wednesday, and one on a Thursday. All lessons and tests were given at the same time of day, and all post-tests were presented 1 week after the final lesson.

Materials and procedure

The experimental paradigm and procedure followed that of a previous study on children’s learning of food chains on spaced learning schedules (Vlach & Sandhofer, 2012). The experiment began on a Monday with a pre-test for students in all three conditions. After the pre-test, students received four lessons spaced according to the condition to which they were assigned. During the final phase of the experiment, each student received a post-test 1 week after their last lesson. The pre-tests and post-tests were videotaped in order to record children’s verbal responses to the memory questions.

Lessons

All children received their first lesson immediately following the pre-test. The other three lessons were presented according to the condition (massed, clumped, or spaced) in which children were assigned. Each lesson was then conducted in the context of a particular biome: grasslands, arctic, ocean, swamp, or desert. The ordering of the biomes used during the lessons was randomly assigned. Children were not given a lesson in the biome that was used during the pre-test and post-test.

The lesson began with the experimenter telling children a series of introductory facts that pertained to all food chains (for examples, see Figure 1). Next, children learned about the animals and plants in a particular biome. They were presented with small figurines of the members of the food chain sized appropriately to show that bigger animals eat smaller animals and smaller animals eat plants—the member that is always at the bottom of the food chain. For example, in the arctic biome, children were shown a wolf and told that

‘The wolf eats even smaller animals, the wolf eats the seal. The seal eats even smaller animals, it eats the fish. The fish doesn’t eat animals, the fish eats plants, and the fish eats seaweed’.



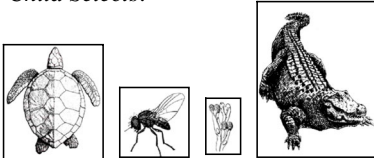
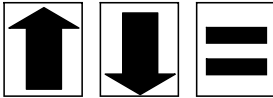

Example Questions from Pre-test & Post-test		
Memory	Simple Generalization	Complex Generalization
Free Recall Question: “What is a biome?”	Question: “What does the Frog eat?”	Scenario: “Let’s say that all the Frogs get captured and taken away by hunters.”
Cued Recall Question: “Bigger animals typically eat _____ animals.”		Question: “What happens to the number of Turtles? Does it go up, down, or stay the same?”
Forced-Choice Question: “What do carnivores eat... animals, plants, or bread?”	<i>Child Selects:</i>	<i>Child Selects:</i>
		
Example Stimuli from Lessons (Tube Set)		
		

Figure 1. Examples of stimuli used during pre-tests/post-tests and lessons. These materials were used for lessons and tests in which the biome was the swamp

Once all figurines in the food chain were presented, they were taken away, and children were shown a set of tubes with pictures of the animals and plants on them.

The tubes varied in size so that the tube of the smaller animal would fit inside the tube of the bigger animal (an example for the swamp biome is depicted in Figure 1). These tubes were used to introduce the concept of interdependency. Children were first told a scenario about a change in the food chain. For example, in the grasslands biome, children were told that a farmer sprayed some poison on the grass. As the experimenter told the children the story, they placed a poison sticker over the grass tube. The experimenter would then say, 'The Cricket comes and eats the Grass. What do you think happens to the Cricket?' The experimenter placed the cricket tube on top of the grass tube so that the grass tube was no longer visible. The experimenter then would lean the tube over and demonstrate that the poison sticker was now inside of the cricket tube. The experimenter would then continue this procedure for all of the creatures of the food chain, to demonstrate that what happened to one creature in the food chain could affect all of the other creatures in that food chain.

Pre-test and post-test

The pre-test and post-test were identical for each participant. The tests consisted of (i) a series of memory questions, (ii) a series of forced-choice simple generalization questions, and (iii) a series of forced-choice complex generalization questions. Examples are shown in Figure 1. Children were asked questions related to a particular biome, which was randomly assigned for each child: grasslands, arctic, desert, ocean, or swamp. The entire test was approximately 5 minutes in length. Children received no instruction or feedback during the test. The ordering of the tests (memory, simple generalization, and complex generalization) was randomly assigned.

Memory test. The memory test consisted of a series of free-recall, cued-recall, and forced-choice questions (see Figure 1 for examples). These questions tested memory for facts that were presented in each lesson. For free-recall questions, children were asked to verbally provide the definition of a word (such as biome or 'food chain'). For the cued-recall questions, children were also required to recall facts (such as 'bigger animals typically eat smaller animals') but were provided with part of the phrasing used by the experimenter during each lesson. For forced-choice questions, children were asked to provide the answers to fact questions (such as 'what is a carnivore?') by selecting from one of three picture options (Figure 1).

Simple generalization test. The simple generalization test consisted of four questions requiring children to make the generalization that bigger animals generally eat smaller animals. The experimenter would show the child a picture of one member of the food chain and then place four pictures of the other members of the food chain in front of the child (Figure 1). The child would be asked to choose what that animal eats. For example, in the swamp biome, the experimenter would show the child a picture of a frog and say 'This is a Frog'. The experimenter would then ask 'Which

of these living things does the Frog eat?' The child would be asked to choose which of the four pictures represented the living thing that the frog eats. This process was then repeated three times for other living things. The ordering of the questions was randomly assigned for each child.

Complex generalization test. The complex generalization test consisted of four questions. In these questions, children were required to generalize the concept of interdependency: A food chain is a dynamic structure in which animals depend on each other for food and survival. To test this generalization, children were told a story about a biome and asked what would happen to the other animals in that biome (Figure 1). For example, in the swamp biome, children were told that all the frogs got captured and taken away by hunters. The experimenter then asked four questions about how the food chain would change on the basis of the given scenario. As an example, children would be asked, 'What do you think happens to the number of Turtles in the swamp? Does it go up, go down, or stay the same?' The experimenter placed three cards on the table, one with an arrow pointing up, one with an arrow pointing down, and one with an equal sign. Children then pointed to the answer they thought was correct. The experimenter then continued with three additional complex generalization questions. The ordering of the questions was randomly assigned for each child.

A week after their last lesson, children received a post-test. Children did not receive instruction in that biome during the lessons. For example, if the child was tested in the swamp biome, they received lessons in the desert, ocean, grasslands, and arctic biomes.

RESULTS

The current study was designed to determine if lesson timing would affect children's memory, simple generalization of concepts, and complex generalization of concepts. We were also interested in whether children's improvements in memory would be related to improvements in generalization. The first step was to determine whether there were differences in performance across the lesson timing conditions. In order to determine if lesson timing affected children's learning, we examined children's pre-test scores and post-test scores, which are summarized in Table 1. Pre-test and post-test scores were calculated using three sub-scores: one for memory questions, one for simple generalization questions, and one for complex generalization questions.

Memory score

The memory score was calculated by separately compiling the sub-scores for the forced-choice questions, cued-recall questions, and free-recall questions. For each question type, we first calculated the pre-test and post-test scores and then calculated a difference score. There were three forced-choice questions, and children received a total of 7 points for each correct answer. Thus, children could receive a score from 0 to 21 points for the forced-choice category. There were two cued-recall and two free-recall questions, which were rated on a scale from 1 (*not correct*) to 7 (*completely correct*)

Table 1. Average pre-test and post-test scores by test (memory, simple generalization, or complex generalization)

Learning schedule	Memory score		Simple generalization score		Complex generalization score	
	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Massed	12.88 (5.24)	18.66 (9.65)*	14.58 (3.61)	18.67 (6.21)	12.17 (5.03)	16.42 (6.15)
Clumped	14.75 (7.38)	21.36 (8.20)*	13.42 (4.68)	18.08 (4.68)*	11.08 (6.59)	16.08 (4.03)*
Spaced	15.77 (4.68)	28.44 (9.96)*	15.08 (4.68)	23.17 (6.56)*	11.08 (8.68)	22.25 (6.06)*

Note: Pre-test scores for the massed, clumped, and spaced conditions did not significantly differ from each other on each test (memory, simple generalization, and complex generalization). For the memory test, the possible score range was 4–49. For the generalization tests, the possible score range was 0–28.

*There was a statistically significant change in score from pre-test to post-test ($p < .05$).

by three independent raters (inter-rater correlations, $r_s > .9$, for all four questions). The score for each question was determined by averaging the scores given by the three raters. Thus, children could receive a score of 1–7 points for each question, making the total possible points for the cued-recall category and free-recall category range between 2 and 14 points. Finally, the composite memory score for the pre-test and post-test were calculated by adding up the sub-scores from each of the three categories of questions (possible score for pre-test and post-test: 4–49 points).

Next, a repeated measures ANOVA was conducted, with lesson timing as the between-subjects variable and performance on the two tests as the within-subjects outcome variable. This analysis revealed a main effect of lesson timing, $F(2, 33) = 3.860$, $p = .045$, a main effect of test, $F(1, 33) = 36.247$, $p < .001$, and an interaction of lesson timing condition and test, $F(2, 33) = 2.042$, $p = .046$. Post-hoc planned comparisons, with Bonferroni corrections, were conducted to examine the interaction between lesson timing and test. These planned comparisons revealed that there was a significant increase in performance from pre-test to post-test for the massed, $t(11) = 3.915$, $p = .002$, clumped, $t(11) = 2.817$, $p = .017$, and spaced conditions, $t(11) = 4.168$, $p = .002$. Thus, there were significant improvements in memory in all three lesson timing conditions. We conducted another set of planned comparisons, with Bonferroni corrections, to examine differences in the amount of change in memory performance across the three conditions. These tests revealed that children in the spaced condition had significantly larger improvements in memory than children in the massed condition, $t(22) = 2.094$, $p = .047$. Taken together, these findings suggest that all lesson timing conditions promoted children's memory for information, but the spaced condition promoted memory to a greater degree than the massed condition.

Simple and complex generalization scores

In addition to improvements in memory across the lesson timing conditions, we were also interested in whether there would be differences in performance on the simple and complex generalization tests. The generalization scores were calculated by compiling pre-test and post-test scores. These calculations were conducted separately for the simple and complex generalization tests. There were four forced-choice questions on each test, and children received a total of 7 points for each correct answer. Thus, children could receive a score from 0 to 28 points for the simple generalization test and from 0 to 28 points for the complex generalization test.

For the simple generalization scores, a repeated measures ANOVA was conducted with lesson timing as the between-subjects variable and performance on the two tests as the within-subjects outcome variable. This analysis revealed a main effect of lesson timing, $F(2, 33) = 4.108$, $p = .026$, a main effect of test, $F(1, 33) = 25.756$, $p < .001$, and a marginally significant interaction of lesson timing condition and test, $F(2, 33) = 2.817$, $p = .074$. Post-hoc planned comparisons, with Bonferroni corrections, were conducted to examine the interaction between lesson timing and test. These planned comparisons revealed that there was a significant increase in performance from pre-test to post-test for the clumped, $t(11) = 3.023$, $p = .012$, and spaced conditions, $t(11) = 4.733$, $p = .001$, but not the massed condition, $t(11) = 1.301$, $p = .220$. Thus, there were significant improvements in simple generalization performance in the clumped and spaced conditions. We conducted another set of planned comparisons, with Bonferroni corrections, to examine differences in the amount of change in simple generalization performance across the three conditions. These tests revealed that children in the spaced condition had significantly larger improvements in simple generalization than children in the massed condition, $t(22) = 2.783$, $p = .023$, and clumped condition, $t(22) = 2.828$, $p = .010$. Taken together, these findings suggest that the spaced condition promoted children's simple generalization performance significantly more than the massed or clumped conditions.

A similar analysis was conducted for the complex generalization scores; a repeated measures ANOVA was conducted with lesson timing as the between-subjects variable and performance on the two tests as the within-subjects outcome variable. This analysis revealed a main effect of lesson timing, $F(2, 33) = 3.860$, $p = .045$, a main effect of test, $F(1, 33) = 25.377$, $p < .001$, and an interaction of lesson timing condition and test, $F(2, 33) = 3.998$, $p = .028$. Post-hoc planned comparisons, with Bonferroni corrections, were conducted to examine the interaction between lesson timing and test. These planned comparisons revealed that there was a significant increase in performance from pre-test to post-test for the clumped, $t(11) = 2.727$, $p = .020$, and spaced conditions, $t(11) = 3.855$, $p = .003$, but not the massed condition, $t(11) = 1.773$, $p = .114$. Thus, there were significant improvements in complex generalization performance in the clumped and spaced conditions. We conducted another set of planned comparisons, with Bonferroni corrections, to examine differences in the amount of change in complex generalization performance across the three conditions. These tests revealed that children in the spaced condition had significantly larger

improvements in complex generalization than children in the massed condition, $t(22)=2.178$, $p=.040$. Taken together, these findings suggest that the spaced condition promoted children's complex generalization performance significantly more than the massed condition.

In sum, children in the spaced condition had significantly higher performance on both the simple and generalization tests, which is the first replication of these findings (Vlach & Sandhofer, 2012).

Memory and generalization scores

Were the changes in memory performance related to changes in generalization performance? In order to examine this possibility, we conducted the same repeated measures ANOVAs described earlier for the simple and complex generalization scores. In these set of analyses, change in memory score (change from pre-test to post-test) was examined as a potential covariate variable. If global memory improvements were promoting generalization, we would expect to observe a significant relationship between the memory scores and generalization scores. However, change in memory score was not a significant covariate for the simple generalization analysis, $F(1, 32)=0.174$, $p=.680$, or the complex generalization analysis, $F(1, 32)=0.891$, $p=.352$. We also conducted correlational analyses with Spearman's ρ and found no significant relationship between improvements in memory and improvements in simple generalization, $r(36)=.018$, $p=.919$, or complex generalization, $r(36)=.069$, $p=.690$.

We were also interested in examining if the same cognitive process that was supporting both simple and complex generalization. In this case, we expected that the two generalization scores should be correlated. The results of an analysis using Spearman's ρ revealed that children's difference scores (changes from pre-test to post-test) on the simple and complex generalization tests were significantly correlated with each other, $r(36)=.416$, $p=.012$. Thus, although children's memory scores were not significantly related to their generalization scores, their two generalization scores were correlated with each other.

Why were the memory scores and generalization scores not significantly correlated? This finding suggests that global improvements in memory may not be responsible for improvements in generalization. Instead, this is the first study to provide evidence that, as previously hypothesized (e.g., Vlach et al., 2008, 2012), the cognitive processes underlying spaced learning are likely contributing to memory and generalization performance in a slightly different manner. We discuss how spaced learning schedules may be uniquely contributing to memory and generalization performance in the Discussion section.

DISCUSSION

The current study was designed to (i) determine if spaced schedules simultaneously promote multiple forms of learning, such as memory and generalization, and (ii) examine whether improvements in different forms of learning are related to each other. The results revealed that children in the spaced condition had significantly larger improvements in memory and generalization performance than children in

the massed and clumped schedules, suggesting that spaced learning schedules can simultaneously support multiple forms of learning. An analysis of whether improvements in memory, simple generalization, and complex generalization performance were related to each other revealed that simple and complex generalization performance were significantly related to each other, but memory performance was not significantly related to generalization performance. These results have implications for our understanding of how spaced schedules promote learning and the application of spaced schedules to educational contexts, which are discussed later.

How does spaced learning promote multiple forms of learning?

How does spaced learning promote memory? Historically, there have been four categories of theories proposed to explain how spaced learning promotes memory: (i) study-phase retrieval theory (e.g., Thios & D'Agostino, 1976), (ii) encoding variability theory (e.g., Glenberg, 1979), (iii) consolidation theories (e.g., Landauer, 1969), and (iv) deficient processing theories (e.g., Hintzman, 1974). To date, the most predominant theory is study-phase retrieval theory (Delaney et al., 2010), which suggests that spaced learning allows time for forgetting between learning events. Forgetting engages learners in more effortful retrieval of prior information during subsequent presentations of that information, which in turn promotes memory. In brief, spaced schedules engage learners in more effortful retrieval of information during the learning period, which supports later retrieval of that information across time.

How does spaced learning promote generalization? Research on spaced learning and generalization is in its infancy, and consequently, this is an open question. By one account, the learning processes that promote memory should promote generalization in the same manner, as generalization is simply an epiphenomenon of memory (e.g., Detterman, 1993). By a similar account, general improvements in memory should support the ability to generalize information. However, the current study does not support this hypothesis—we did not observe a direct relationship between memory performance and generalization performance. Instead, the current study suggests that the cognitive processes underlying spaced learning may be contributing to memory and generalization performance in a different manner.

A frequently proposed explanation for how spaced learning promotes generalization is a variant of study-phase retrieval theory (e.g., Vlach et al., 2008, 2012). By this account, the same cognitive processes, forgetting and effortful retrieval, promote generalization. However, forgetting promotes abstraction and generalization by supporting memory in a specific manner—spaced learning promotes memory for relevant features of a category/concept, by reactivating the information across learning events, and deters memory of irrelevant features of a category/concept, by allowing forgetting to occur between learning events. This differential in memory weights in turn supports generalization because learners are more likely to recollect and generalize on the basis of relevant information rather than irrelevant information.

Does the current study support this account? In this experiment, information that was presented and tested in the memory tasks consisted of new vocabulary terms (e.g., biome) and facts (e.g., what does a carnivore eat) that were related to the simple and complex generalizations. However, these items of information were not the central features of the simple and complex concepts. Thus, although the current study did not directly test this proposed account for how spaced learning promotes generalization, the current results are not inconsistent with this account, and forgetting and retrieval dynamics could be contributing to the observed results. As outlined later, there are additional explanations for why there was no observed relationship between memory and generalization performance. Future research will be needed in order to clarify how spaced learning promotes generalization.

Another possibility for why we did not observe a relationship between memory performance and generalization performance is that there may be individual differences in what forms of learning benefit from spaced learning schedules. To date, there is very little research on individual differences in spaced learning (see Delaney *et al.*, 2010, for a discussion of this issue). However, recent research has indicated that memory capacities might mediate the degree to which spaced learning is beneficial for learning. For example, one recent study (Verkoeijen & Bouwmeester, 2008) suggested that the spacing effect is smaller for college students with an overall lower memory performance level than for students with an overall higher memory performance level. In the case of different forms of learning, such as memory and generalization, it may be that children experience greater benefits of spaced schedules for what they have a greater capacity to learn, on the basis of their prior experience with that information. Although the sample size of the current study is not large enough to observe individual differences, future work should examine the role of individual differences in spaced learning in order to explore these possibilities. Indeed, an understanding of individual differences is both important theoretically and essential for integrating spaced schedules into applied contexts, such as educational interventions.

Conclusion

This study contributes to a growing body of work demonstrating the benefits of spaced learning for educational interventions and contexts (Bjork, 1994; Carpenter *et al.*, 2009; Kornell, 2009; Pashler *et al.*, 2007; Rohrer, 2009; Seabrook *et al.*, 2005; Sobel *et al.*, 2011; Vlach & Sandhofer, 2012). The current experiment expands this research by demonstrating that spaced learning schedules can simultaneously promote multiple forms of learning in educational interventions. Future research should continue to explore how forgetting and effortful retrieval support memory and generalization, and how these learning dynamics can be optimized in the context of educational interventions.

ACKNOWLEDGEMENTS

We thank the undergraduate research assistants of the Language and Cognitive Development Lab for their help with this project. We appreciate the help of the staff, teachers,

and students of the UCLA Lab School. This study was supported by a UCLA Lab School/CONNECT fellowship and grant.

REFERENCES

- Bahrick, H. P., Bahrick, L. E., Bahrick, A. S., & Bahrick, P. E. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science*, 4, 316–321. DOI: 10.1111/j.1467-9280.1993.tb00571.x
- Birnbaum, M., Kornell, N., Bjork, E. L., & Bjork, R. A. (2013). Why interleaving enhances inductive learning: The role of discrimination and retrieval. *Memory & Cognition*, 41, 392–402. DOI: 10.3758/s13421-012-0272-7
- Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. P. Shimura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.
- Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th grade students' retention of U.S. history facts. *Applied Cognitive Psychology*, 23, 760–771. DOI: 10.1002/acp.1507
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132, 354–380. DOI: 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. DOI: 10.1037/0012-1649.38.6.967
- Delaney, P. F., Verkoeijen, P. J., & Spiguel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. *The Psychology of Learning and Motivation: Advances in Research and Theory*, 53, 63–147. DOI: 10.1016/S0079-7421(10)53003-2
- Dempster, F. N. (1988). The spacing effect: A case study in the failure to apply the results of psychological research. *American Psychologist*, 43, 627–634. DOI: 10.1037/0003-066X.43.8.627
- Detterman, D. K. (1993). The case for the prosecution: Transfer as an epiphenomenon. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 1–24). Westport, CT, US: Ablex Publishing.
- Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect. *Journal of Applied Psychology*, 84, 795–805.
- Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology* (H. A. Ruger, C. E. Bussenius & E. R. Hilgard, Trans.). New York: Dover Publications. (Original work published in 1885).
- Eilam, B. (2002). Strata of comprehending ecology: Looking through the prism of feeding relations. *Science Education*, 86, 645–671.
- Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition*, 7, 95–112.
- Hintzman, D. L. (1974). Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology; The Loyola symposium* (pp. 77–97). Potomac, MD: Erlbaum.
- Kang, S. H. K., & Pashler, H. (2012). Learning painting styles: Spacing is advantageous when it promotes discriminative contrast. *Applied Cognitive Psychology*, 26, 97–103. DOI: 10.1002/acp.1801
- Kornell, N. (2009). Optimising learning using flashcards: Spacing is more effective than cramming. *Applied Cognitive Psychology*, 23, 1297–1317. DOI: 10.1002/acp.1537
- Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the “enemy of induction”? *Psychological Science*, 19, 585–592. DOI: 10.1111/j.1467-9280.2008.02127.x
- Kornell, N., Castel, A. D., Eich, T. S., & Bjork, R. A. (2010). Spacing as the friend of both memory and induction in younger and older adults. *Psychology and Aging*, 25, 498–503. DOI: 10.1037/a0017807
- Landauer, T. K. (1969). Reinforcement as consolidation. *Psychological Review*, 76, 82–96. DOI: 10.1037/h0026746
- Pashler, H., Rohrer, D., Cepeda, N. J., & Carpenter, S. K. (2007). Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic Bulletin & Review*, 14, 187–193.
- Rea, C. P., & Modigliani, V. (1987). The spacing effect in 4- to 9-year-old children. *Memory & Cognition*, 15(5), 436–443. DOI: 10.3758/BF03197733

- Rohrer, D. (2009). The effects of spacing and mixing practice problems. *Journal for Research in Mathematics Education*, 40, 4–17.
- Rohrer, D. (2012). Interleaving helps students distinguish among similar concepts. *Educational Psychology Review*, 24, 355–367.
- Rovee-Collier, C., Evancio, S., & Earley, L. A. (1995). The time window hypothesis: Spacing effects. *Infant Behavior and Development*, 18, 69–78. DOI: 10.1016/0163-6383(95)90008-X
- Seabrook, R., Brown, G. D., & Solity, J. E. (2005). Distributed and massed practice: From laboratory to classroom. *Applied Cognitive Psychology*, 19, 107–122.
- Sobel, H. S., Cepeda, N. J., & Kapler, I. V. (2011). Spacing effects in real-world classroom vocabulary learning. *Applied Cognitive Psychology*, 25, 763–767. DOI: 10.1002/acp.1747
- Thios, S. J., & D'Agostino, P. R. (1976). Effects of repetition as a function of study-phase retrieval. *Journal of Verbal Learning & Verbal Behavior*, 15, 529–536. DOI: 10.1016/0022-5371(76)90047-5
- Toppino, T. C. (1991). The spacing effect in young children's free recall: Support for automatic-process explanations. *Memory & Cognition*, 19(2), 159–167. DOI: 10.3758/BF03197112
- Toppino, T. C., & DeMesquita, M. (1984). Effects of spacing repetitions on children's memory. *Journal of Experimental Child Psychology*, 37(3), 637–648. DOI: 10.1016/0022-0965(84)90081-X
- Verkoeijen, P. P. J. L., & Bouwmeester, S. (2008). Modeling bimodality in spacing effect data. *Journal of Memory and Language*, 59, 545–555.
- 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, & Cognition*, 38, 246–254. DOI: 10.1037/a0025260
- Vlach, H. A., & Sandhofer, C. M. (2012). Distributing learning over time: The spacing effect in children's acquisition and generalization of science concepts. *Child Development*, 83, 1137–1144. DOI: 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, 163–167. DOI: 10.1016/j.cognition.2008.07.013
- Wahlheim, C. N., Dunlosky, J., & Jacoby, L. L. (2011). Spacing enhances the learning of natural concepts: An investigation of mechanisms, meta-cognition, and aging. *Memory & Cognition*, 39, 750–763. DOI: 10.3758/s13421-010-0063-y
- Zulkipli, N., McLean, J., Burt, J. S., & Bath, D. (2012). Spacing and induction: Application to exemplars presented as auditory and visual text. *Learning and Instruction*, 22, 215–221. DOI: 10.1016/j.learninstruc.2011.11.002