Journal of Experimental Child Psychology 183 (2019) 115–133

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

To mass or space? Young children do not possess adults' incorrect biases about spaced learning

Haley A. Vlach *, Catherine A. Bredemann, Carla Kraft

Department of Educational Psychology, University of Wisconsin–Madison, Madison, WI 53706, USA

A R T I C L E I N F O

Article history:
Received 16 March 2018
Revised 7 February 2019
Available online 11 March 2019

Keywords:
Spaced learning
Massed bias
Memory development
Metamemory
Metacognition
Cognitive development

A B S T R A C T

The spacing effect is the robust finding that learners have stronger long-term memory for information presented on a spaced schedule, in which learning events are distributed across time, rather than a massed schedule, in which learning events are presented in immediate succession. Despite the fact that the spacing effect is highly replicable across tasks and timescales, most adults do not believe that spaced learning promotes memory. Instead, there is a persistent "massed bias"; adults believe that massed learning promotes memory to a greater degree than spaced learning. The developmental origins of the massed bias have yet to be studied; thus, the goals of the current research were to (a) identify a developmental period in which we do not observe a massed bias and (b) determine whether metamemory is related to the onset of the massed bias. The results revealed that children (aged 2–10 years; $N = 109$) do not have a persistent massed bias, and the number of massed endorsements increased across the early elementary school years. Children's age predicted a massed bias, but individual differences in children's metamemory abilities were not related to bias development when controlling for age. Taken together, this work suggests that researchers will need to reconceptualize dual-process theoretical accounts of metamemory and spaced learning to explain why and how children develop a massed bias.

* Corresponding author.
E-mail address: hvlach@wisc.edu (H.A. Vlach).

https://doi.org/10.1016/j.jecp.2019.02.003

Introduction

A long history of research on memory has sought to identify the conditions of the learning environment that promote the retention and retrieval of knowledge across time (dating back to Ebbinghaus, 1885, 1964). A key finding from this body of work is that the *timing of learning events* contributes to learners' ability to remember information. However, learners are not always cognizant of what learning conditions are best, and they can even develop misconceptions as to what processes do and do not benefit their own learning. To date, we lack an understanding of when and how learners come to believe that certain schedules are more or less advantageous for memory. To address this gap in the extant literature, the current research examined whether there are differences in children's and adults' conceptions about how the timing of learning affects memory. In particular, we focused our investigation on learners' conceptions of the spacing effect.

The *spacing effect* is the most widely studied learning phenomenon in human memory (Ebbinghaus, 1885, 1964; for reviews and meta-analyses, see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Delaney, Verkoeijen, & Spirgel, 2010; Toppino & Gerbier, 2014). The spacing effect refers to the finding that long-term memory is enhanced when learning events are distributed in time (i.e., spaced learning) rather than presented in immediate succession (i.e., massed learning). In a typical study (e.g., Bjork & Allen, 1970; Toppino, Kasserman, & Mracek, 1991), participants are presented with a list of items in which each item is presented multiple times. Half of the items are distributed in time (i.e., on a spaced schedule) and half of the items are presented in immediate succession (i.e., on a massed schedule). After a delay, participants' memory for the massed and spaced items is tested. Across studies, participants consistently demonstrate stronger memory for items presented on a spaced schedule than for items presented on a massed schedule.

The spacing effect is highly replicable across timescales, learning tasks, and populations; there are more than 1000 published experiments that have observed a spacing effect in human memory, making the spacing effect the most highly replicated learning phenomenon in psychological science (Cepeda et al., 2006; Delaney et al., 2010). We also observe spacing effects across developmental periods, including during early childhood (e.g., Rovee-Collier, Evancio, & Earley, 1995; Vlach, Sandhofer, & Kornell, 2008), adulthood (e.g., Cepeda et al., 2006; Yan, Bjork, & Bjork, 2016), and older adulthood (e.g., Benjamin & Craik, 2001; Kornell, Castel, Eich, & Bjork, 2010). Indeed, researchers have argued that the mechanisms underlying the spacing effect are developmentally invariant; for instance, there is evidence that study-phase retrieval theory can account for spacing effects during infancy (e.g., time-window hypothesis; Rovee-Collier et al., 1995), childhood (e.g., forgetting-as-abstraction theory; Vlach, 2014), and adulthood (Johnston & Uhl, 1976; Thios & D'Agostino, 1976). According to study-phase retrieval theory (Johnston & Uhl, 1976; Thios & D'Agostino, 1976), spacing effects emerge because the time between events provides learners with the opportunity to forget information; this forgetting causes learners to experience greater difficulty when retrieving prior information during learning. As a result, learners engage in more cognitive effort in retrieving information, solidifying the memory trace, and slowing the future forgetting rate of that information. In sum, although infants and children forget at a faster rate than adults (e.g., Bauer, 2015; Rovee-Collier et al., 1995), the pattern of forgetting follows a similar pattern/mathematical function; thus, the mechanism underlying spacing effects (i.e., forgetting) has been argued to be the same across developmental periods.

It is important to note that although most learners are likely to benefit from a spaced schedule, within a given task there are also likely to be some learners who benefit from a massed schedule or demonstrate equal performance on massed and spaced schedules. For instance, Verkoeijen and Bouwmeester (2008) found that the degree to which adults benefitted from a spaced schedule was moderated by their memory abilities. Thus, for a small group of learners, there could be a benefit of a massed schedule or no differences in performance between massed and spaced schedules.

Over the past few decades, there has been a special interest in connecting the spacing effect with other domains of research such as metamemory. Metamemory is the understanding of the general processes of human memory, including the understanding of both one's own memory and other people's memory (for reviews, see Cavanaugh & Perlmutter, 1982; Schneider & Pressley, 1997). Metamemory has historically been characterized as having two overlapping and interactive components, procedural metacognitive knowledge and declarative metacognitive knowledge, which develop between early childhood and adulthood (Flavell & Wellman, 1977; Schneider, 2010). Procedural metacognitive knowledge refers to implicit behavioral knowledge of when memory is necessary, whereas declarative metacognitive knowledge refers to explicit factual knowledge about the variables that affect memory performance. Researchers have been particularly interested in declarative metacognitive knowledge and the spacing effect such as whether children and adults think that timing affects memory performance and whether children and adults realize that spaced learning promotes memory to a greater degree than massed learning. The long-term goal of this work has been to understand how individual learners come to identify the conditions of the learning environment that promote their own memory.

Most research on the spacing effect and metamemory has been conducted with adult learners. In these studies, adults are asked to predict which learning schedule (massed, spaced, or neither schedule) will optimize memory by making judgments of learning (JOLs). For example, Kornell and Bjork (2008) presented adults with paintings to learn on one of two schedules: massed or spaced. Participants were taught the terms *massed* and *spaced* and then were presented with a JOL; participants were asked, "Which do you think helped you learn more, massed or spaced?" They were given three response options: *massed, spaced,* or *about the same.* The results demonstrated that despite the fact that most participants had higher performance on the spaced schedule, the majority of participants believed that they had higher performance on the massed schedule. Many studies have replicated the same general finding; adults incorrectly believe that massed learning promotes memory to a greater degree than spaced learning (e.g., Baddeley & Longman, 1978; Kornell et al., 2010; McCabe, 2011; Simon & Bjork, 2001; Zechmeister & Shaughnessy, 1980). In sum, there is a persistent *massed bias* during adulthood.

How do misconceptions about massed and spaced learning develop? According to dual-process theories of adults' metamemory, metacognitive judgments are a result of both mnemonic-based (or experience-based) and theory-based processes (e.g., Kelley & Jacoby, 1996; Koriat & Bjork, 2006; Koriat, Bjork, Sheffer, & Bar, 2004; Matvey, Dunlosky, & Guttentag, 2001). Experienced-based JOLs arise from cues gathered during learners' online processing of to-be-remembered items. These cues, such as encoding fluency (i.e., did it feel easy or difficult to learn these items?), are used for experience-based JOLs. In the context of massed and spaced learning, researchers have argued that adults use processing fluency as the basis for a massed bias (e.g., Simon & Bjork, 2001; Zechmeister & Shaughnessy, 1980). That is, because spaced schedules cause learners to experience more retrieval difficulty during learning than massed schedules, learners assume that spaced learning is less advantageous for memory (Simon & Bjork, 2001).

In contrast, theory-based JOLs are inferences about memory performance based on theories of how memory works. Theories of memory are generated through a variety of information stored in long-term memory such as an individual's past experiences, perceived memory competency, and beliefs about factors that can influence memory. By this account, a learner's massed bias would stem from previous successes in using massed learning (e.g., rehearsing a phone number in immediate succession and successfully remembering it later), perceived memory competence (e.g., I have a bad memory; I won't remember anything unless someone says it to me over and over again right now), and/or believing a rule about massed learning (e.g., massed learning is superior to spaced learning). Indeed, previous research has suggested that adults have a massed bias before even starting an experiment or memory task (McCabe, 2011), suggesting that they are using theories of how memory works to guide their JOLs.

Although we know that the massed bias is present during adulthood, we do not know when or why the massed bias develops. One reason for this is that there are a limited number of studies examining children's metamemory and the spacing effect. Two studies have examined elementary school-aged children's JOLs of massed and spaced learning (Son, 2005, 2010). The results of this work revealed that older children can have the same incorrect biases as adults, but these studies do not pinpoint the period during development when children do not have a massed bias. By identifying the period when children do not demonstrate a massed bias, we can begin to understand when and how the massed bias develops. The current experiments addressed this gap in the extant literature.

The current research

This work examined whether preschool-aged children (Experiments 1 and 2), early elementary school-aged children (Experiment 2), and adults (Experiment 1) predict that massed learning will promote memory to a greater degree than spaced learning. We chose these periods of development because (a) these age groups were younger than has been examined in previous research (Son, 2005, 2010) and (b) children have a basic understanding of memory and can make metacognitive judgments (e.g., Gordon & Flavell, 1977; Johnson & Wellman, 1980). Prior research has shown that general metacognitive abilities and metamemory begin to develop early during the preschool years (e.g., Gordon & Flavell, 1977; Johnson & Wellman, 1980; Lyon & Flavell, 1994; Miscione, Marvin, O'Brien, & Greenberg, 1978). For instance, by 3 years of age, children have an understanding of words such as "remember" and "forget" (Johnson & Wellman, 1980; Lyon & Flavell, 1994).

Our hypothesis was that young children would not have a massed bias. Preschool-aged children have a basic understanding of memory and have only begun to think about how their own memory works. As a result, it is unlikely that they have developed strong misconceptions about how memory works such as a massed bias. That is, based on dual-process theories of metamemory (e.g., Koriat & Bjork, 2006; Koriat et al., 2004), preschool-aged children should not demonstrate a massed bias because they have no theories, or limited theories, of memory and are just beginning to understand their own subjective experiences of remembering information. However, as children's understanding of their own memory develops and they generate theories of memory, they may begin to demonstrate a massed bias during elementary school. That is, if children's general understanding of memory is driving the origin of the massed bias, we would expect to observe a relationship between children's metamemory abilities and the degree to which children endorse massed schedules as optimal for memory. The alternative hypothesis is that young children will have a massed bias. This would suggest that the origins of the massed bias start early in development, before the onset of basic metamemory abilities.

The current experiments tested these hypotheses by examining preschool-aged and early elementary school-aged children's JOLs about massed and spaced learning. In Experiment 1, preschool-aged children and adults were presented with a memory task and a series of JOLs about massed and spaced learning. This afforded the opportunity to establish that adults demonstrate a massed bias in our task and to examine whether preschool-aged children have a massed bias that mirrors the massed bias observed during adulthood. In Experiment 2, preschool-aged and elementary school-aged children were presented with the same procedure as in Experiment 1 and were also presented with a series of metamemory tasks to examine whether there would be a relationship between children's biases and metamemory abilities across development. Taken together, these experiments were designed to determine the ontology of the massed bias and, in turn, to inform theories of metamemory and provide guidance to educators as to when they should expect to see the massed bias arise during schooling.

Experiment 1

Method

Participants

The participants were 33 preschool-aged children (M_{age} = 4.5 years, range = 41–62 months; 24 girls) and 59 college students (M_{age} = 20.5 years; 54 women). Effect sizes were gathered from recently published studies on the spacing effect with this age group, which had consistently large effect sizes (ds > 1.0; e.g., Vlach, Ankowski, & Sandhofer, 2012; Vlach et al., 2008). To be conservative in determining a sample size, we used a smaller effect size in the large effect category: d = .80. A power analysis for a two-tailed t test, with α = .05 revealed that we would need at least 26 participants to have 80% power to observe an effect. Thus, we decided to collect data for 3 months or until we reached 26 participants (if we did not reach this number within the 3-month period).

Children were recruited from local daycare centers and received a storybook for their participation in the study. Children were monolingual English speakers and from predominantly middle- to upper-socioeconomic status (SES) families. Adult participants were recruited from the department's participant pool and received course credit for their participation. An additional 4 children participated in the experiment but were not included in the final sample because of their inability to follow directions during the experiment (e.g., always choosing answer choice "C" at test).

Design

Participants were presented with six tasks: an introduction to massed and spaced learning task, three JOLs, memory task training, and a memory task. The presentation ordering of the tasks can be seen in Fig. 1.

Apparatus and stimuli

Participants were presented with all tasks on an iPad. The pictures used during the introduction to massed and spaced learning task, memory task training, and memory task were pictures of common objects (Fig. 1). The pictures were taken from a standardized set of pictures normed for name agreement, image agreement, familiarity, and visual complexity for children and adults (Snodgrass & Vanderwart, 1980). Each of the pictures was presented in only one task of the experiment.

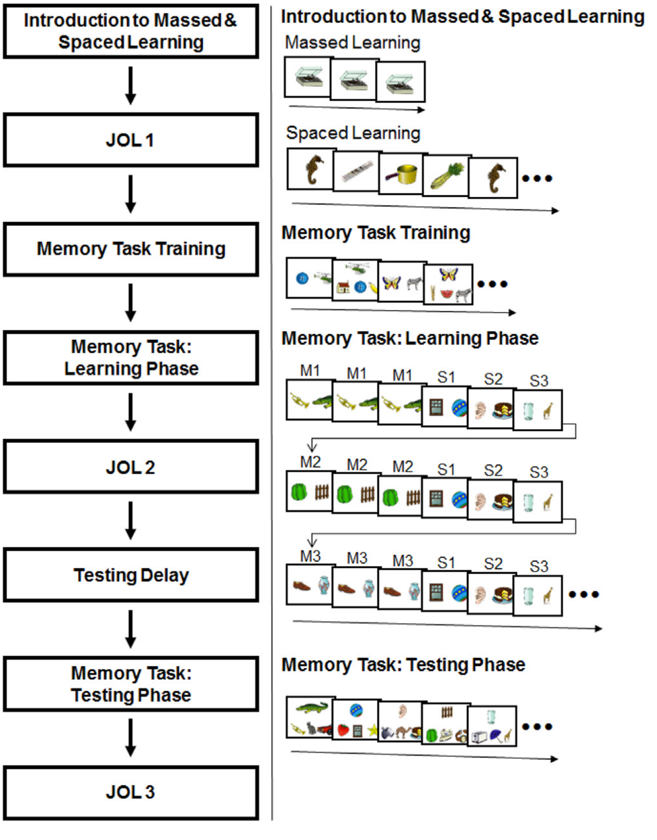

Fig. 1. Left side: Order of tasks presented in both Experiments 1 and 2. In Experiment 2, children were also presented with a massed and spaced learning training task after the introduction to massed and spaced learning. Right side: Examples of stimuli used during the experiments. M, massed item; S, spaced item.

Procedure

The ordering of the tasks can be seen in Fig. 1. After completing each task, the experimenter would proceed to the next task until all tasks were completed. The duration of the entire experiment was about 10 min.

Introduction to massed and spaced learning task. Previous research has taught participants the terms "massed" and "spaced" because adults and older adults are highly unfamiliar with these terms (e.g., Kornell & Bjork, 2008; Kornell et al., 2010). Given that adults are unlikely to know the terms "massed" and "spaced," we suspected that preschool-aged children would also be unlikely to know these words. Thus, the introduction to massed and spaced learning task taught participants the terms "massed learning" and "spaced learning." The experimenter started explaining massed learning to children and adults by saying, "Massed learning is when you see a picture over and over again, right in a row." To give a visualization of massed learning, the experimenter then presented a picture on the iPad screen for 3 s and said, "Look at this picture!" (Fig. 1). The picture then disappeared for 250 ms to give an appearance of the screen changing, and the same picture appeared again for another 3 s. On the second presentation the experimenter said, "Look how you see it again, right in a row." The picture then disappeared for 250 ms to give an appearance of the screen changing, and the same picture appeared again for another 3 s. On the third presentation the experimenter said, "Look how you see it again, right in a row." The picture was then removed from the screen and the experimenter said, "What you saw is called massed learning."

The experimenter explained spaced learning by saying, "Spaced learning is when you see a picture, then it goes away for a while, then you see the picture again." To give a visualization of spaced learning, the experimenter presented a picture on the iPad screen for 3 s and said, "Look at this picture!" (Fig. 1). Next, three different pictures were presented in immediate succession, for 3 s each, and then the first picture reappeared for 3 s. On the second presentation, the experimenter said, "Look. It went away but here it is again." Three more pictures were presented in immediate succession, for 3 s each, and then the first picture reappeared for 3 s. On the third presentation, the experimenter said, "Look. It went away but here it is again." The picture was then removed from the screen and the experimenter said, "What you saw is called spaced learning." In sum, children and adults were provided with the same instructions; the entire introduction to massed and spaced learning task took less than 1 min.

Judgments of learning. The JOLs assessed participants' perception and predictions of how massed and spaced learning would affect their memory. At the beginning of each of the three JOLs (Fig. 1), the experimenter provided a verbal reminder of the definition of massed and spaced learning using the same language used in the introduction to massed and spaced learning task. For JOL 1, which came after the introduction to massed and spaced learning task but before the memory task training, the experimenter asked, "Which one do you think will help you learn the best? Massed, spaced, or will they be about the same?" For JOL 2, immediately after the learning phase of the memory task, the experimenter asked, "Which pictures will you remember the best in three minutes? Massed, spaced, or will they be about the same?" For JOL 3, after the memory test, the experimenter asked, "Which one do you think helped you learn the best? Massed, spaced, or were they about the same?" These prompts were designed to be developmentally appropriate for preschool-aged children. The experimenter recorded participants' responses to the three JOLs on a piece of paper.

Memory task training. The memory task training familiarized participants with the protocol for the memory task. There were 10 training trials, each of which consisted of a learning slide and a test slide. The experimenter began the task by saying, "I am going to show you some pictures, so please look at the screen." The experimenter then pulled up the first learning slide, which consisted of two randomly selected pictures from the stimuli set (Snodgrass & Vanderwart, 1980). The experimenter then said, "Look at these pictures." After the two pictures had been presented for 3 s, the screen changed to a test slide (Fig. 1). The experimenter then said, "Can you please point to the picture that was next to this one?" The experimenter simultaneously pointed to the object toward the top of the screen. The test choices included the target object (i.e., the object shown on the previous slide) and two new objects from the stimuli set (Snodgrass & Vanderwart, 1980). The experimenter recorded participants' responses to the test prompt and then continued to the next training trial. Participants were presented with 5–10 training trials. All participants were able to meet the criterion of at least 5 correct test trials in a row.

Memory task. The memory task was designed to test participants' long-term memory for information presented on massed and spaced schedules. The paradigm was a memory task that is developmentally appropriate for preschool-aged children: cued-recognition memory for lists of pictures (e.g., Howe, 1995). The task had two phases: a learning phase and a test phase (Fig. 1).

During the learning phase, participants were presented with 36 learning trials, each presented for 3 s (Fig. 1). Twelve distinct learning trials were created by randomly selecting two pictures from the stimuli set (Snodgrass & Vanderwart, 1980). Each of the 12 learning trials was presented three times during the learning phase on either a massed or spaced schedule. In the massed schedule, a learning trial was presented three times in immediate succession. In the spaced schedule, each of the 3 learning trial presentations was separated by 5 other learning trials (i.e., 15 s). To ensure equal intervals of time between each of the spaced learning trials, the learning trials were presented in massed and spaced blocks (Fig. 1), with each block consisting of 3 massed learning trials or 3 spaced learning trials. After the last learning trial had been presented to participants, the 3-min test delay began. At the beginning of the 3-min test delay, participants were asked JOL 2. For the remaining portion of the 3-min test delay, child participants played with Play-Doh and adult participants played Angry Birds.

After the 3-min delay, the test phase began. Participants were presented with 24 test trials. Memory for each of the 12 learning trials was tested twice during the test phase. Each of the 12 learning trials was tested once in the same order as presented during the learning phase. Each of the 12 learning trials was then tested again, but the second set of test items tested the inverse pairings from the first 12 test trials (i.e., switched the sides of the screen in which items were presented). That is, the object that was on top of the screen during the first test trial became the correct test choice during the second test trial, and the correct test choice used during the first trial was the object placed toward the top of the screen during the second trial. There were no ordering effects observed between these two sets of test trials ($p > .10$). The test trials had the same protocol as used in the memory training task. The experimenter said, "Can you please point to the picture that was next to this one?" The experimenter simultaneously pointed to the object toward the top of the screen. The test trials included three choices: the target object (i.e., the object shown during the learning phase) and two new objects from the stimuli set (Snodgrass & Vanderwart, 1980). The experimenter recorded participants' responses to the test prompt and then continued to the next test trial. After participants had completed all the test trials, participants were asked JOL 3 and the experiment concluded.

Results and discussion

A central goal of this experiment was to determine whether preschool-aged children have a massed bias and whether adults have a massed bias in this particular task. The descriptive statistics, by age group, are presented in Table 1. As can be seen in this table, all participants were able to successfully complete the memory task training, but there appeared to be differences between children's and adults' performance on the memory task and JOLs. Thus, the analyses examined children's and adults' performance separately on these measures.

Memory task

We hypothesized that children would have higher performance on the spaced schedule of the memory task. To determine whether there was a spacing effect, we first calculated the number of correct responses on the massed and spaced items (Table 1). We then conducted paired-samples t tests, which revealed that children had significantly higher performance on a spaced schedule, $t(32) = 2.860$, $p = .007$, $d = 0.51$, but there was no significant difference between performance on the massed and spaced schedule for adults ($p > .10$). The memory task was designed to be developmentally appropriate for preschool-aged children, and (as to be expected) adults appeared to have very high performance on the memory task. Two one-sample t tests for massed and spaced items confirmed that there was not a significant difference from ceiling performance (12 of 12 correct) ($ps > .10$.)

Table 1
Descriptive statistics for tasks in Experiment 1 by age group.

	Children	Adults
Mean age (years)	4.5	20.5
Successful completion of memory task training (%)	100	100
Memory performance for massed items		
Mean (*SD*)	8.82 (3.066)	11.98 (0.430)
Memory performance for spaced items		
Mean (*SD*)	9.97 (2.589)	11.81 (0.572)
JOL 1: Before learning (%)		
Massed choices	6.1	50.8
Spaced choices	30.3	28.8
Same choices	63.6	20.3
Accurate choices	42.4	22.0
JOL 2: After learning (%)		
Massed choices	15.2	42.4
Spaced choices	21.2	37.3
Same choices	63.6	20.3
Accurate choices	42.4	18.6
JOL 3: After test (%)		
Massed choices	15.2	18.6
Spaced choices	24.2	45.8
Same choices	60.6	35.6
Accurate choices	33.3	37.3

Note. Data are broken down by age group: preschool-aged children ($n = 33$) and adults ($n = 59$). The highest possible score for massed and spaced items was 12.

Participants were also categorized into subgroups based on their final test performance. Specifically, participants were categorized as having higher performance on a massed schedule (*n*s = 3 children and 7 adults), higher performance on a spaced schedule (*n*s = 15 children and 15 adults), or the same performance on both schedules (*n*s = 15 children and 51 adults). These subcategories were used to assess the accuracy of participants' metacognitive judgments in relation to their own memory performance, as described in the section below.

Judgments of learning

We started by examining the adults' JOL choices to establish that this was a task in which they believe massed presentation would be advantageous for learning. A one-way chi-square analysis examining adults' JOL choices relative to chance performance (i.e., comparison with expected frequencies) revealed that adults had a higher number of massed choices at JOL 1, $\chi^2(2) = 8.78$, $p = .012$, Cramer's $V = .385$, and a marginally higher number of massed choices at JOL 2, $\chi^2(2) = 4.712$, $p = .095$, Cramer's $V = .282$. These findings show that in the context of the current task, adult participants were biased to think that a massed schedule would promote stronger memory.

To determine whether children had the same biases as adults, the next set of analyses examined children's JOL choices. A one-way chi-square analysis examining children's choices relative to chance performance revealed that children had a higher number of same choices at JOL 1, $\chi^2(2) = 16.55$, $p < .001$, Cramer's $V = .708$, at JOL 2, $\chi^2(2) = 13.82$, $p < .001$, Cramer's $V = .647$, and at JOL 3, $\chi^2(2) = 11.46$, $p = .003$, Cramer's $V = .589$. These results suggest that children had little or no bias for massed or spaced learning. Thus, the preschool children in this study did not share the same biases as adults.

The last set of analyses examined whether children and adults changed their responses to the JOLs over the course of the experiment. A chi-square analysis of the three choice categories (i.e., massed, spaced, and same) at each JOL revealed no significant differences in children's JOL choices during the experiment ($p > .10$). Indeed, most children (73.9%) were consistent in their JOLs, either never changing their JOL or changing their JOL only once. The same analysis with adults revealed that there were changes in adults' JOL choices during the experiment, $\chi^2(4) = 14.691$, $p = .005$, Cramer's $V = .204$.

In particular, adults demonstrated a significant decrease in massed choices, $\chi^2(1) = 8.820$, $p = .0122$, Cramer's $V = .387$, but no significant changes in the number of spaced and same choices ($ps > .10$), across the three JOLs. Thus, adults, but not children, slightly changed their pattern of metacognitive judgments during the experiment.

JOLs and memory performance

We compared participants' JOL choices with their performance on the memory task. As described above, participants were classified as having higher performance on a massed schedule, higher performance on a spaced schedule, or the same performance on both schedules based on their final test performance in the memory task. Participants' memory performance was then directly compared with their predictions on JOLs 1–3. If participants' JOL choice matched their final memory test performance, they were considered as having an accurate JOL. The percentage of participants correctly predicting their performance on the final test at JOLs 1–3 can be seen in Table 1 ("accurate choices"). Neither children nor adults were particularly accurate with their JOLs, consistent with previous research on metamemory and the spacing effect in adult learners (e.g., Kornell & Bjork, 2008).

The results of Experiment 1 revealed differences between preschool-aged children and adults. Children were not biased to think that massed or spaced schedules would be more beneficial for memory. However, adults demonstrated a massed bias. In sum, this study is the first to identify a developmental period in which children do not demonstrate a massed bias, the preschool years, and also confirms a robust massed bias during adulthood, including in this task.

When and why does the onset of the massed bias occur? Given that previous research has found a massed bias with older elementary school-aged children (e.g., Son, 2010), it is likely that children develop a massed bias during the preschool and early elementary school years. Thus, we examined these age groups in Experiment 2. During the onset of metacognition, when children begin thinking about how memory works, they may generate the hypothesis that massed learning is better than spaced learning. That is, the development of metamemory may drive early beliefs about massed and spaced learning. Indeed, metacognitive theories of the massed bias have proposed that the bias is the result of theories of how memory works (e.g., Hartwig & Dunlosky, 2012; Koriat & Bjork, 2006; Koriat et al., 2004). In this case, we would expect to see individual differences in the onset of the massed bias during the preschool and early elementary school years. In particular, preschool-aged children with limited metamemory abilities would have no massed or spaced bias, whereas elementary school-aged children with strong metamemory abilities would have a hypothesis about how timing affects learning such as a massed bias. This possibility was examined in Experiment 2.

Experiment 2 also addressed a potential issue in the design of Experiment 1. One explanation for why children had a lower incidence of massed or spaced choices during JOLs is that they did not understand the terms "massed" and "spaced." That is, it could be that the introduction to massed and spaced learning task was ineffective at teaching children to judge the difference between the two learning schedules. Thus, in Experiment 2 we included a massed and spaced training task to assess whether children could reliably tell the difference between massed and spaced schedules. Taken together with performance on the metamemory task, we would be able to determine whether children can tell the difference between massed and spaced schedules and understand general questions about metamemory.

Experiment 2

In this experiment, we examined children's metacognitive judgments of massed and spaced learning in relation to their own memory performance and metamemory abilities. We predicted that children's metamemory abilities would be related to the degree to which they picked massed learning as the most effective learning schedule for memory. In particular, we hypothesized that children with limited metamemory abilities would have no bias, whereas children with more developed metamemory abilities would have a massed bias. To test these hypotheses, we replicated Experiment 1 and included three new tasks: a massed and spaced training task and two metamemory tasks.

Method

Participants

The participants were 76 preschool- and elementary school-aged children (M_{age} = 59.00 months, SD = 21.26, range = 28–129 months, 2.5–10 years; 32 girls). To be conservative in determining a sample size, we used an effect size of d = .80 in Experiment 1. Because of the wider age range in this experiment, we used an even more conservative effect size: d = .60. A power analysis for a two-tailed t test, with α = .05, revealed that we would need at least 45 participants to have 80% power to observe an effect. Thus, we decided to collect data for 6 months or until we reached 45 participants (if we did not reach this number within the 6-month period).

Children were recruited from local daycare centers and elementary schools and received a storybook for their participation in the study. Parents received $10 for bringing children to the lab to participate in the study. Children were monolingual English speakers and from predominantly middle- to upper-SES families. An additional 13 children participated in the experiment but were not included in the final sample because of their inability to follow directions or to understand the task during the experiment (e.g., always choosing answer choice "C" at test).

Design

Children were presented with nine tasks: an introduction to massed and spaced learning task, a training task for massed and spaced learning, three JOLs, memory task training, a memory task, and two metamemory tasks. The two metamemory tasks, presented together, either preceded or followed the remaining tasks; order was counterbalanced across participants. The training task for massed and spaced learning occurred immediately after the introduction to massed and spaced learning but before JOL 1 (see Fig. 1).

Apparatus and stimuli

Children were presented with most tasks on an iPad. One of the metamemory tasks was presented using cards. As in Experiment 1, the pictures used during the introduction to massed and spaced learning task, massed and spaced training task, memory task training, and memory task were pictures of common objects (Fig. 1). The pictures were taken from Snodgrass and Vanderwart (1980), a standardized set of pictures normed for name agreement, image agreement, familiarity, and visual complexity. Each of the pictures was presented in only one task of the experiment.

Procedure

The ordering of the tasks can be seen in Fig. 1. After completing each task, the experimenter would proceed to the next task until all tasks were completed. The duration of the entire experiment was about 20 min.

Introduction to massed and spaced learning task. This was the same as in Experiment 1.

Massed and spaced learning training task. This task assessed children's understanding of the introduction to massed and spaced learning. In this task, children were presented with 2 sample trials of massed learning and 2 sample trials of spaced learning, for a total of 4 trials. The experimenter provided a prompt to refer to the previous introduction, prompting: "Now, let's see if we can tell the difference between massed and spaced learning. We're going to see some pictures, and I want you to tell me which type of learning we're using with these pictures." After each of the 4 trials, children were asked to identify whether they had just observed an example of massed learning or spaced learning. There was only 1 child who was not able to respond correctly on any of the training trials; thus, this child was removed from the sample for failure to understand the task.

Memory task training. This was the same as in Experiment 1.

Memory task. This was the same as in Experiment 1.

Judgments of learning. This was the same as in Experiment 1.

Metamemory task 1: Memory moderators. The first metamemory task was adapted from a frequently used metamemory assessment by Wellman (1977). This task assessed children's knowledge of memory-relevant variables. For instance, the task measured children's understanding of what sorts of individual, task, and circumstantial variables might support or impair one's memory for previously learned information. Stimuli consisted of two-dimensional colored line drawings on 24 laminated 4×6-inch cards that the experimenter laid on the table before the participants (Fig. 2). Children were presented with 12 pairs of cards, with each pair comparing two scenes side by side. Each pair depicted two scenes of children doing a similar activity, with only one variable differing between the two scenes. There were 10 pairs that tested individual, task, and circumstantial moderators of memory and 2 pairs that tested irrelevant factors (e.g., the color of one's shirt). Examples of moderators included how much time on task children had available to study information, whether children had a friend to help them remember information, and the number of distractors in the environment.

Children were presented with one pair of cards at a time and were instructed to point to whichever card depicted the child with the easier task. If they thought that both children would experience the same amount of ease in remembering information, they were instructed to point in the middle of the two cards. For instance, the experimenter would say, "We're going to look at some pictures of different boys and girls trying to do different activities. I want you to show me with your pointer finger which kid will have an easier time remembering than the other kid. If you think they will be about the same, point in the middle, between the two pictures." The experimenter then read from a script containing short narratives explaining and comparing the two scenes. Once children selected one of the two

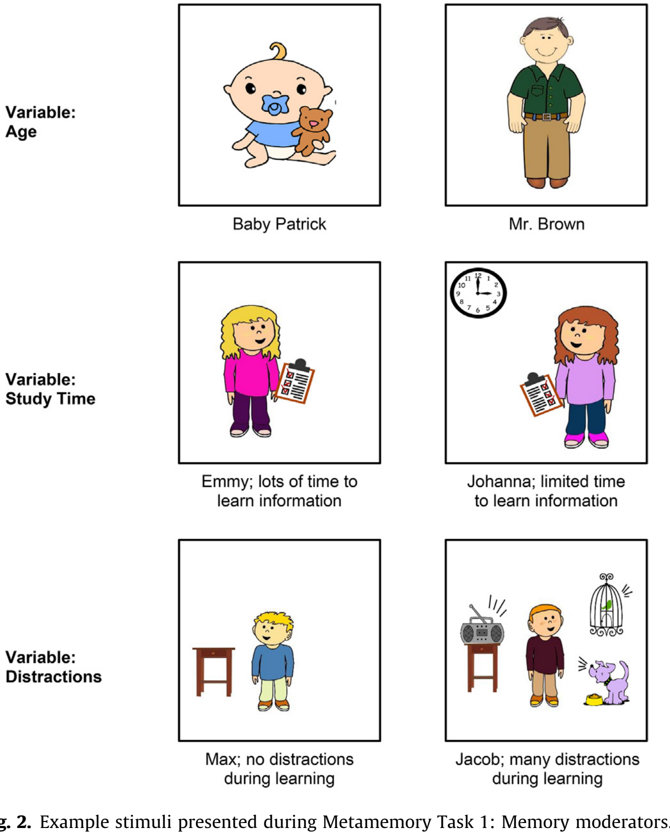

Fig. 2. Example stimuli presented during Metamemory Task 1: Memory moderators.

scenes or indicated that there would be no difference, the pair of cards in front of them was removed and the subsequent pair was presented. Responses were recorded on a piece of paper by the experimenter.

Metamemory task 2: Remember and forget. The second metamemory task was adapted from a frequently used metamemory assessment by Wellman and Johnson (1979). This task assessed children's knowledge of the words "remember" and "forget"; thus, it measured children's understanding of these two cognitive states. Stimuli consisted of two short animated presentations, each about 40 s long (Fig. 3), narrated by the experimenter. Both animations portrayed a scenario in which a child stores his or her coat in a certain place. In one scenario, the child looks for his coat in the correct location and finds it, resembling remembering. In the other scenario, the child looks for her coat in a wrong location, resembling forgetting. After being presented with the first animated scenario, participants were asked to identify what had just occurred using the two terms being assessed. The experimenter would ask "Did Toby *remember* where his coat was? Did Toby *forget* where his coat was?" The same two questions were posed after the child was presented with the second scenario, for a total of four questions in the task. Responses were recorded by the experimenter on a piece of paper.

Results and discussion

A central goal of this experiment was to replicate Experiment 1 and determine whether there is a relation between children's JOLs and their metamemory abilities. The descriptive statistics for each task are presented in Table 2.

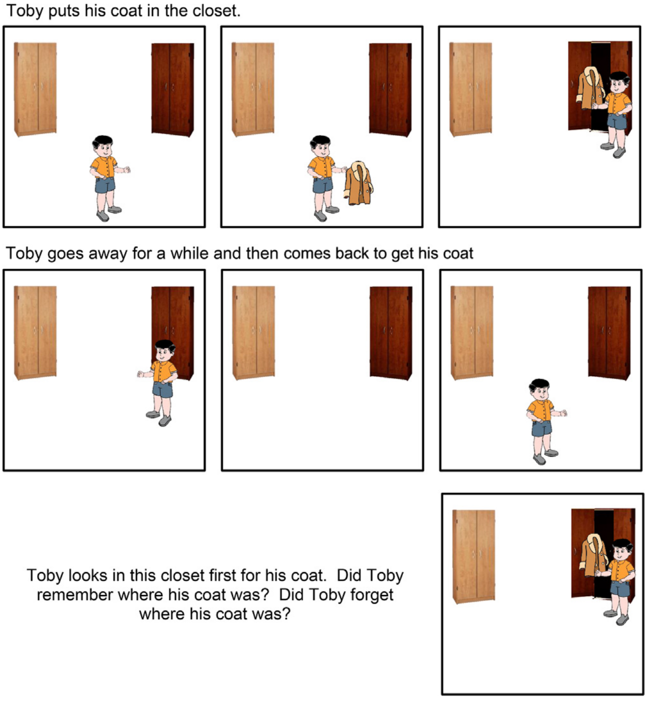

Fig. 3. Example of animated stimuli presented during Metamemory Task 2: Remember and forget. The storyline shown is the story in which the child character correctly remembers information.

Table 2
Descriptive statistics for tasks in Experiment 2.

	Children
Mean age (months)	58.75
Performance on massed and spaced learning training task	
Mean (*SD*)	2.79 (1.024)
Successful completion of memory task training (%)	100
Memory performance for massed items	
Mean (*SD*)	7.13 (3.549)
Memory performance for spaced items	
Mean (*SD*)	7.74 (3.399)
JOL 1: Before learning (%)	
Massed choices	22.4
Spaced choices	15.8
Same choices	61.8
Accurate choices	34.2
JOL 2: After learning (%)	
Massed choices	13.2
Spaced choices	18.4
Same choices	68.4
Accurate choices	34.2
JOL 3: After test (%)	
Massed choices	22.1
Spaced choices	17.1
Same choices	61.8
Accurate choices	36.8
Metamemory Task 1: Memory moderators	
Mean (*SD*)	7.13 (2.685)
Metamemory Task 2: Remember and forget	
Mean (*SD*)	3.61 (0.732)

Note. $N = 76$. The highest possible score for massed and spaced items and Metamemory Task 1 was 12. The highest possible score for the massed and spaced learning training task and Metamemory Task 2 was 4.

Massed and spaced learning training task

Experiment 2 included a new task, the massed and spaced learning training task, to determine whether children could reliably identify a massed and spaced schedule after the introduction to massed and spaced learning. A one-sample t test was conducted in comparison with chance performance (2 correct responses) and revealed that performance was significantly above chance, $t(75) = 6.722$, $p < .001$, $d = 0.77$. More than 50% of children performed above chance (3 or 4 correct answers), and 35.5% of children had a perfect score (4 correct answers). Thus, children were able to successfully identify massed and spaced learning after the first 4 trials of the introduction to massed and spaced learning. This finding suggests that children's lack of a massed bias in Experiment 1 was unlikely to be because children did not understand the task.

Memory task

We hypothesized that children would have higher performance on the spaced schedule of the memory task. To determine whether there was a spacing effect, we first calculated the number of correct responses on the massed and spaced items (Table 2). We then conducted a paired-samples t test, which revealed that children had significantly higher performance on a spaced schedule, $t(75) = 2.999$, $p = .004$, $d = 0.34$. Children were also categorized into subgroups based on their final test performance. Specifically, children were categorized as having higher performance on a massed schedule ($n = 23$ children), higher performance on a spaced schedule ($n = 40$ children), or the same performance on both schedules ($n = 13$ children). These subcategories were used to assess the accuracy of children's metacognitive judgments in relation to their own memory performance, as described in the section below.

Judgments of learning

We hypothesized that we would replicate Experiment 1 and find that children do not have a massed bias when making JOLs. We started by examining the JOL 1 choices, which were made before the memory task. A one-way chi-square analysis directly comparing children's choices on JOL 1 with chance performance revealed that children had a higher number of same choices, $\chi^2(2) = 28.29$, $p < .001$, Cramer's $V = .431$. The same set of analyses was conducted for JOL 2, which occurred after the learning phase, and JOL 3, which occurred after the test phase of the memory task. These analyses revealed that children had a higher number of same choices for JOL 2, $\chi^2(2) = 42.42$, $p < .001$, Cramer's $V = .528$, and JOL 3, $\chi^2(2) = 27.98$, $p < .001$, Cramer's $V = .429$. The last set of analyses examined whether children changed their responses to the JOLs over the course of the experiment. A chi-square analysis of the three choice categories (i.e., massed, spaced, and same) at each JOL revealed no significant differences in children's JOL choices during the experiment ($p > .10$). Indeed, most children (84.2%) were consistent in their JOLs, either never changing their JOL or changing their JOL only once. This pattern of performance replicates Experiment 1.

JOLs and memory performance

To test the replication of Experiment 1, we examined the degree to which children were able to accurately predict their own memory performance. Thus, we compared children's JOL choices with their performance on the memory task. As described above, participants were classified as having higher performance on a massed schedule, higher performance on a spaced schedule, or the same performance on both schedules based on their final test performance during the memory task. Participants' memory performance was then directly compared with their predictions on JOLs 1–3. If children's JOL choice matched their final memory test performance, children were considered as having an accurate JOL. The percentage of children correctly predicting their performance on the final test at JOLs 1–3 can be seen in Table 2 ("accurate choices"). A chi-square analysis between Experiments 1 and 2 revealed no significant differences in accuracy across the two experiments ($p > .10$). Thus, the pattern of performance in Experiment 2 replicated that in Experiment 1.

Jols and metamemory performance

The central goal of Experiment 2 was to examine children's JOLs in relation to their metamemory abilities. We first examined whether there was sufficient variance in the metamemory tasks to capture individual and developmental differences. That is, we examined whether performance on the two metamemory tasks was significantly above chance performance and below ceiling performance using a series of one-sample *t* tests. For Metamemory Task 1, children's performance was significantly above chance performance (4 correct responses), $t(75) = 10.168$, $p < .001$, $d = 1.16$, and below ceiling performance, $t(75) = 15.807$, $p < .001$, $d = 1.81$. For Metamemory Task 2, children's performance was significantly above chance performance (2 correct responses), $t(75) = 19.125$, $p < .001$, $d = 2.19$, and below ceiling performance, $t(75) = 4.703$, $p < .001$, $d = 0.54$. Thus, both tasks were able to capture variance in children's metamemory abilities. Moreover, performance on the two tasks was correlated, $r(76) = .441$, $p < .001$. To create a global measure of metamemory abilities, children's scores on the two tasks were combined ($M = 10.74$ [of 16], $SD = 3.078$, range = 3–16).

To examine whether there were relations among children's age, metamemory abilities, and JOL biases, we ran a series of Pearson's *r* correlations among children's age, metamemory abilities, and total number of massed and spaced choices during the experiment. These statistics are presented in Table 3, including partial correlations controlling for age. The correlations revealed a significant relation between children's age and the number of massed choices, suggesting that as children get older they are more likely to demonstrate a massed bias. A scatterplot of this finding can be seen in Fig. 4. We then ran a regression analysis with the number of massed choices as the outcome variable and age, performance on the massed and spaced training task, and performance on the metamemory tasks as the three predictor variables. The results of the regression revealed that age was the only significant

Table 3
Pearson's *r* correlations of measures in Experiment 2.

	1	2	3	4	5
1. Age (months)	1				
2. Metamemory abilities (raw score)	.66*	1	.19	−.04	−.13
3. Massed/Spaced training (raw score)	.53*	.47*	1	.06	−.15
4. Number of massed choices	.58*	.36*	.35*	1	−.18
5. Number of spaced choices	.03	−.07	−.11	−.13	1

Note. $N = 76$. Metamemory abilities was the summed total of the number of correct responses on the two metamemory tasks. Bottom left correlations are simple correlations, and top right correlations are partial correlations (controlling for age).
* $p < .05$.

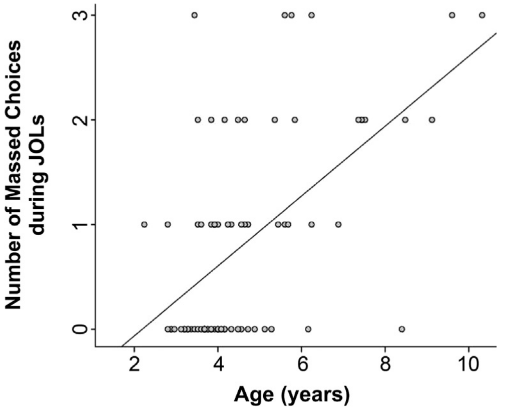

Fig. 4. Scatterplot of children's ages by the number of massed choices during JOLs in Experiment 2. There were three JOLs during the experiment. The angled line represents the best-fit line for the data.

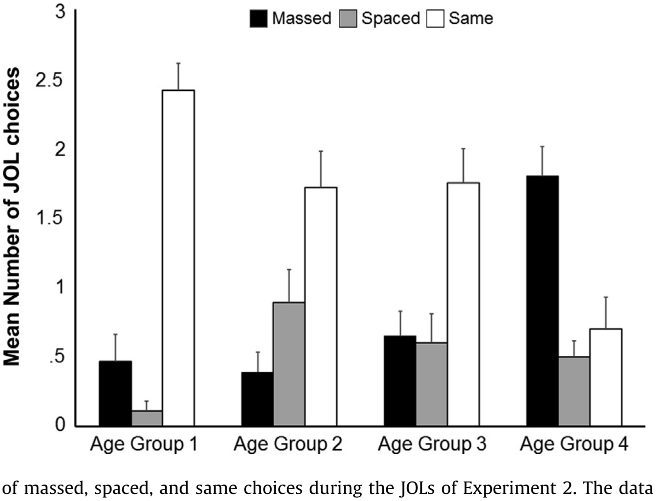

Fig. 5. Mean numbers of massed, spaced, and same choices during the JOLs of Experiment 2. The data are grouped into four equivalently sized age groups: Age Group 1 (28–45 months, 2- and 3-year-olds; $n = 19$), Age Group 2 (46–50 months, 4-year-olds; $n = 18$), Age Group 3 (51–67 months, 4- and 5-year-olds; $n = 20$), and Age Group 4 (68–129 months, 6- to 10-year-olds; $n = 19$). The three younger age groups show a same bias, whereas the oldest age group shows a massed bias. Error bars represent 1 standard error of the mean.

predictor of the number of massed choices ($\beta = .579$; model: $R = .580$, $R^2 = .336$), $F(3, 72) = 12.168$, $p < .001$.

We continued to examine the relation between children's age and the number of massed choices by examining the data discontinuously; this afforded us the opportunity to determine a narrower age range in which we observe the onset of the massed bias. We divided the sample into four equivalently sized age groups based on natural breaks in the age data: Age Group 1 (28–45 months, 2- and 3-year-olds; $n = 19$), Age Group 2 (46–50 months, 4-year-olds; $n = 18$), Age Group 3 (51–67 months, 4- and 5-year-olds; $n = 20$), and Age Group 4 (68–129 months, 6- to 10-year-olds; $n = 19$). Descriptive statistics for the JOLs by age group are presented in Fig. 5. A multivariate analysis of variance (ANOVA), with number of massed, spaced, and same choices as the outcome variables and age group as fixed factor, revealed significant main effects of age group, Wilks' lambda = .562, $F(6, 142) = 7.902$, $p < .001$. Bonferroni post hoc tests revealed a significantly larger number of same choices for the three youngest age groups and a significantly larger number of massed choices in the oldest age group (ps < .05). The oldest age group was then split in half by age; results of an independent-samples t test revealed that there was no significant difference in the number of massed choices in the younger group of Age Group 4 ($n = 10$; $M = 1.7$ massed choices, $SD = 1.06$) and the older group of Age Group 4 ($n = 9$; $M = 2.0$ massed choices, $SD = 0.87$), $p > .10$.

Taken together, the results of Experiment 2 reveal two important findings. First, the general pattern of performance observed in children who participated in Experiment 2 replicated the pattern of performance observed in Experiment 1. Second, children's age (and not metamemory abilities) was the strongest predictor of children's likelihood of choosing massed schedules as optimal. These results have several theoretical and applied implications, which are outlined in the General Discussion below.

General discussion

Many studies have demonstrated that adults consistently believe that massed learning is more advantageous for memory than spaced learning (e.g., Baddeley & Longman, 1978; Kornell et al., 2010; McCabe, 2011; Simon & Bjork, 2001). Experiment 1 provided a new demonstration of this bias; despite showing ceiling performance on a memory task designed for preschoolers, when they should have said that the learning schedule did not matter, adults still persistently believed that massed learning led to better memory. In contrast, Experiments 1 and 2 revealed that most preschool-aged children believed that massed and spaced learning would not differentially affect their memory.

When do children acquire a massed bias? The degree to which children chose massed schedules as optimal was related to children's age and a decreasing number of same choices in Experiment 2. Moreover, the youngest three age groups in Experiment 2 demonstrated a same bias, whereas the oldest group demonstrated a massed bias (Fig. 5). Thus, across elementary school, children are likely to shift from thinking that there is no difference between massed and spaced schedules to having a massed bias. That is, they are unlikely to go through a period when they think that spaced schedules are more advantageous. Indeed, previous research has shown that children in later grades of elementary school have a massed bias rather than a spaced bias (Son, 2005, 2010).

There are a number of ways in which children may come to believe that massed learning is more advantageous for memory. Given that children's metamemory abilities were not a strong predictor of the massed bias in Experiment 2, we do not believe that children's general understanding of memory is the main factor that leads to the massed bias. Instead, we predict that children learn a specific rule about massed versus spaced learning. One mechanism might be children's observation of adults' informal memory strategies. That is, children may observe adults in their environment using massed learning. When adults are trying to remember information, such as a phone number, they often verbally repeat the information to themselves in immediate succession. Children may see adults rehearsing information in one moment in time, rather than across moments in time, and implicitly learn from these observations that massed practice is optimal for memory.

Another way in which children may come to believe that massed learning is more advantageous than spaced learning is through direct instruction from adults. When children reach kindergarten, they may take their first formal tests such as handwriting, addition, and spelling tests. When preparing for these first tests, parents and teachers may instruct children to practice or study with massed learning (e.g., immediate repetition, cramming). Indeed, as demonstrated in the current study and in previous research (e.g., Baddeley & Longman, 1978; Kornell et al., 2010; McCabe, 2011; Simon & Bjork, 2001), adults are likely to believe that massed learning is more advantageous for learning. Moreover, research has shown that elementary school teachers' use of mnemonic language and strategies is correlated with children's memory strategy use (Grammer, Coffman, & Ornstein, 2013). Thus, it could be that parents and elementary school teachers are explicitly telling kindergartners and first graders that massed learning is more advantageous than spaced learning, leading children to develop a massed bias. In both the informal and formal experiences, children might not be developing a theory of how massing promotes memory akin to dual-process theories (Koriat & Bjork, 2006; Koriat et al., 2004) but instead simply encoding a rule based on adults' experiences and beliefs. Because there are several experiences that could be contributing to the development of the massed bias, future research should examine the informal and formal experiences that contribute to children's metacognitive conceptions of the spacing effect.

The current findings have important implications for theories of metamemory. As outlined earlier, dual-process theories of metamemory (Kelley & Jacoby, 1996; Koriat & Bjork, 2006; Koriat et al., 2004; Matvey et al., 2001) often assume that learners' JOLs are both experienced based and theory based. We argue that within the current task, adults were using a theory-based only system for their JOLs. At the beginning of the experiment, before experiencing the memory task, adults had a strong massed bias. After being presented with the memory task, they still had a massed bias. Thus, adults did not consider the many cues (e.g., easy encoding, memory task designed for preschool-aged children) that should have led them to not have a massed bias. Indeed, adults had ceiling performance at the final test and should have realized that the learning schedule did not influence their performance. In sum, this study suggests that there may be cases in which adult learners' JOLs are completely theory based rather than a combination of theory and experience based.

The current experiments suggest that children may shift from having no theory or limited theories for JOLs to using a more developed theory-based system for JOLs. Preschool-aged children had no bias, suggesting that neither their theories nor their experience was sufficient to sway them to have a bias. Elementary school-aged children, in contrast, did have a massed bias. However, this massed bias was unlikely the result of their experiences during the experiment. As children gained more memory performance evidence (e.g., learning and posttest performance) during both experiments, they did not change their JOL responses. Moreover, we argue that even older elementary school-aged children who possess a massed bias (Son, 2010) are highly unlikely to have the cognitive resources to simultaneously weigh a multitude of experience-based variables, such as study time, presentation timing, and retention interval, when making JOLs. Thus, accurately assessing whether massed or spaced learning is more advantageous for one's learning is unlikely to be guided by experience-based JOLs during childhood. We predict that children may simply be applying a rule (e.g., "My parent/teacher uses massed learning, so massed is better"). That is, we predict that a one-process account, in which theories begin as simple rules (i.e., rules without an underlying conceptual framework), may be a better metacognitive theory for characterizing the development of the massed bias than dual-process accounts proposed for adult learners.

The current findings also have several implications for applied settings, such as for educational practices. Unfortunately, most students are unaware that spaced learning can promote learning and memory (Karpicke, Butler, & Roediger, 2009). Even when college students receive instruction on the spacing effect and other memory strategies, they do not readily adopt them (for a discussion, see Kornell & Bjork, 2007). Thus, an open question in research on metamemory is when and how students should be taught about memory strategies so that they (a) believe these strategies to be helpful for their own learning and (b) adopt these strategies during learning. One variable that may affect the degree to which learners adopt memory strategies is the developmental period in which students first learn about memory strategies. The current work identified a developmental period in which learners may be likely to change their conceptions about massed and spaced learning: early childhood. Indeed, the optimal time to teach memory strategies might not be during high school or college. Instead, we may want to teach preschool- and early elementary school-aged children, who have yet to develop misconceptions about memory, about optimal ways of learning.

In sum, the current research identifies the point during the lifespan when we are likely to develop misconceptions about the spacing effect: the early elementary school years. This work suggests that to understand how children generate a massed bias, we will likely need to create new theoretical accounts of the massed bias such as one-process theoretical accounts. Future research should continue to examine preschool-aged children's understanding of the spacing effect to identify other factors that contribute to the development of the massed bias. Indeed, this work will allow us to have a complete theoretical account of the origins of memory misconceptions and solutions for integrating memory strategies into educational contexts.

Acknowledgments

This work was supported in part by the Wisconsin Alumni Research Foundation and the Wisconsin Center for Education Research. Many thanks go to the undergraduate research assistants of the University of Wisconsin–Madison Learning, Cognition, & Development Lab for their help with this project.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2019.02.003.

References

Baddeley, A. D., & Longman, D. J. A. (1978). The influence of length and frequency of training session on the rate of learning to type. *Ergonomics, 21*, 627–635.
Bauer, P. J. (2015). Development of episodic and autobiographical memory: The importance of remembering forgetting. *Developmental Review, 38*, 146–166.
Benjamin, A. S., & Craik, F. I. (2001). Parallel effects of aging and time pressure on memory for source: Evidence from the spacing effect. *Memory & Cognition, 29*, 691–697.
Bjork, R. A., & Allen, T. W. (1970). The spacing effect: Consolidation or differential encoding? *Journal of Verbal Learning and Verbal Behavior, 9*, 567–572.
Cavanaugh, J. C., & Perlmutter, M. (1982). Metamemory: A critical examination. *Child Development, 53*, 11–28.
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.
Delaney, P. F., Verkoeijen, P. J., & Spirgel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. *Psychology of Learning and Motivation, 53*, 63–147.
Ebbinghaus, H. (1964). Memory: A contribution to experimental psychology (H. A. Ruger, C. E. Bussenius, & E. R. Hilgard, Trans.). New York: Dover. (Original work published 1885).
Flavell, J. H., & Wellman, H. M. (1977). Metamemory. In R. Kail & W. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 3–31). Hillsdale, NJ: Lawrence Erlbaum.
Gordon, F. R., & Flavell, J. H. (1977). The development of intuitions about cognitive cueing. *Child Development, 48*, 1027–1033.
Grammer, J., Coffman, J. L., & Ornstein, P. (2013). The effect of teachers' memory-relevant language on children's strategy use and knowledge. *Child Development, 84*, 1989–2002.
Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement? *Psychonomic Bulletin & Review, 19*, 126–134.
Howe, M. L. (1995). Interference effects in young children's long-term retention. *Developmental Psychology, 31*, 579–596.
Johnson, C. N., & Wellman, H. M. (1980). Children's developing understanding of mental verbs: Remember, know, and guess. *Child Development, 51*, 1095–1102.
Johnston, W. A., & Uhl, C. N. (1976). The contributions of encoding effort and variability to the spacing effect on free recall. *Journal of Experimental Psychology: Human Learning and Memory, 2*, 153–160.
Karpicke, J. D., Butler, A. C., & Roediger, H. L. III., (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory, 17*, 471–479.
Kelley, C. M., & Jacoby, L. L. (1996). Adult egocentrism: Subjective experience versus analytic bases for judgment. *Journal of Memory and Language, 35*, 157–175.
Koriat, A., & Bjork, R. A. (2006). Mending metacognitive illusions: A comparison of mnemonic-based and theory-based procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 1133–1145.
Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General, 133*, 643–656.
Kornell, N., & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review, 14*, 219–224.
Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the "enemy of induction"? *Psychological Science, 19*, 585–592.
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.

Lyon, T. D., & Flavell, J. H. (1994). Young children's understanding of "remember" and "forget". *Child Development, 65*, 1357–1371.
Matvey, G., Dunlosky, J., & Guttentag, R. (2001). Fluency of retrieval at study affects judgments of learning (JOLs): An analytic or nonanalytic basis for JOLs? *Memory & Cognition, 29*, 222–233.
McCabe, J. (2011). Metacognitive awareness of learning strategies in undergraduates. *Memory & Cognition, 39*, 462–476.
Miscione, J. L., Marvin, R. S., O'Brien, R. G., & Greenberg, M. T. (1978). A developmental study of preschool children's understanding of the words "know" and "guess". *Child Development, 49*, 1107–1113.
Rovee-Collier, C., Evancio, S., & Earley, L. A. (1995). The time window hypothesis: Spacing effects. *Infant Behavior and Development, 18*, 69–78.
Schneider, W. (2010). Metacognition and memory development in childhood and adolescence. In H. Salatas Waters & W. Schneider (Eds.), *Metacognition, strategy use, and instruction* (pp. 54–81). New York: Guilford.
Schneider, W., & Pressley, M. (1997). *Memory development between 2 and 20*. Mahwah, NJ: Lawrence Erlbaum.
Simon, D. A., & Bjork, R. A. (2001). Metacognition in motor learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 907–912.
Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 174–215.
Son, L. K. (2005). Metacognitive control: Children's short-term versus long-term study strategies. *Journal of General Psychology, 132*, 347–364.
Son, L. K. (2010). Metacognitive control and the spacing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 255–262.
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*, 529–536.
Toppino, T. C., & Gerbier, E. (2014). About practice: Repetition, spacing, and abstraction. In B. H. Ross (Ed.). *The psychology of learning and motivation* (Vol. 60, pp. 113–189). San Diego: Academic Press.
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*, 123–138.
Verkoeijen, P. P. J. L., & Bouwmeester, S. (2008). Using latent class modeling to detect bimodality in spacing effect data. *Journal of Memory and Language, 59*, 545–555.
Vlach, H. A. (2014). The spacing effect in children's generalization of knowledge: Allowing children time to forget promotes their ability to learn. *Child Development Perspectives, 8*, 163–168.
Vlach, H. A., Ankowski, A. A., & Sandhofer, C. M. (2012). At the same time or apart in time? The role of presentation timing and retrieval dynamics in generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 246–254.
Vlach, H. A., Sandhofer, C. M., & Kornell, N. (2008). The spacing effect in young children's memory and category induction. *Cognition, 109*, 163–167.
Wellman, H. M. (1977). Preschoolers' understanding of memory-relevant variables. *Child Development, 48*, 1720–1723.
Wellman, H. M., & Johnson, C. N. (1979). Understanding of mental processes: A developmental study of "remember" and "forget". *Child Development, 50*, 79–88.
Yan, V. X., Bjork, E. L., & Bjork, R. A. (2016). On the difficulty of mending metacognitive illusions: A priori theories, fluency effects, and misattributions of the interleaving benefit. *Journal of Experimental Psychology: General, 145*, 918–933.
Zechmeister, E. B., & Shaughnessy, J. J. (1980). When you know that you know and when you think that you know but you don't. *Bulletin of the Psychonomic Society, 15*(1), 41–44.