



Spacing and Interleaving Effects Require Distinct Theoretical Bases: a Systematic Review Testing the Cognitive Load and Discriminative-Contrast Hypotheses

Ouhao Chen¹ · Fred Paas^{2,3} · John Sweller⁴

Accepted: 4 March 2021/Published online: 26 March 2021
© The Author(s) 2021

Abstract

Spaced and interleaved practices have been identified as effective learning strategies which sometimes are conflated as a single strategy and at other times treated as distinct. Learning sessions in which studying information or practicing problems are spaced in time with rest-from-deliberate-learning periods between sessions generally result in better learning outcomes than massed practice without rest-from-deliberate-learning periods. Interleaved practice also consists of spaced sessions, but by interleaving topics rather than having rest-from-deliberate-learning periods. Interleaving is usually contrasted with blocking in which each learning topic is taught in a single block that provides an example of massed practice. The general finding that interleaved practice is more effective for learning than blocked practice is sometimes attributed to spacing. In the current paper, the presence of rest-from-deliberate-learning periods is used to distinguish between spaced and interleaved practice. We suggest that spaced practice is a cognitive load effect that can be explained by working memory resource depletion during cognitive effort with recovery during rest-from-deliberate-learning, while interleaved practice can be explained by the discriminative-contrast hypothesis positing that interleaving assists learners to discriminate between topic areas. A systematic review of the literature provides evidence for this suggestion.

Keywords Cognitive load theory · Working memory resources depletion · Spaced practice · Interleaved practice · Discriminative-contrast hypothesis

The spacing effect was described by Ebbinghaus (1885) in what may be the first example of an effect by the new discipline of experimental psychology. It occurs when learning episodes that are spaced with gaps between them result in superior learning compared to the same episodes presented in massed form without spacing. Interleaving is sometimes described as an example

✉ Ouhao Chen
o.chen@lboro.ac.uk

of spacing (e.g., Taylor & Rohrer, 2010), because the learning episodes relevant to one skill are spaced by interleaved learning episodes relevant to a different skill. In this paper, we argue that the spacing and interleaving effects have distinct causes, with spaced practice having been explained by a variety of constructs including the cognitive load related effect of working memory resource depletion, and interleaved practice explained by the discriminative-contrast hypothesis providing an example of a learning-to-perceptually-discriminate effect.

We suggest two different hypothesised processes underlie these two effects:

1. In spaced practice, periods of mental activity are alternated with periods of mental rest-from-deliberate-learning. Mental rest-from-deliberate-learning offers the opportunity to restore depleted working memory resources caused by mental effort during learning.
2. In interleaved practice, periods of practicing one skill are alternated with periods of practicing a different skill. Interleaving does not involve mental rest-from-deliberate-learning but rather assists learners to discriminate between similar concepts or topics of a domain.

A detailed justification of these two hypotheses is provided next.

Cognitive Load Theory and the Spacing Effect

Cognitive load theory (CLT) is an instructional theory based on human cognitive architecture and evolutionary psychology (Sweller et al., 2011; Sweller et al., 2019). It can be described by five fundamental principles (Sweller & Sweller, 2006).

We acquire information using either the *borrowing and re-organising principle* under which information is obtained from other people, frequently teachers, or if we cannot obtain information from others, we randomly generate new information and test it for effectiveness using the *randomness as genesis principle* during problem solving. That information then is processed in a limited capacity and duration working memory by the *narrow limits of change principle* before being stored for indefinite periods of time in our large, long-term memory using the *information store principle*. Based on signals from the external environment, indefinite amounts of that stored information then can be transferred back to working memory to govern appropriate action using the *environmental organizing and linking principle*. These five principles serve as the cognitive base for instructional design. They have been used to predict that when dealing with complex, high element interactivity information that imposes a heavy working memory load, a reduction in element interactivity will enhance learning (Sweller, 2010).

Working memory resource depletion occurs when a task requires a heavy cognitive effort that depletes the limited working memory resources indicated by the narrow limits of change principle, so depressing performance on an immediately subsequent task. Substantial evidence for working memory resource depletion using multiple experiments was provided by Schmeichel (2007). There also is some evidence that working memory resource depletion may be more likely when two tasks share common cognitive components (Healey et al., 2011). The depleted resources can be restored by rest-from-deliberate-learning (e.g., Tyler & Burns, 2008). CLT assumes that inserted spacing with rest-from-deliberate-learning during spaced practice allows the restoration of depleted working memory resources, thus enhancing subsequent learning.

As indicated above, cognitive load effects usually require high element interactivity information that imposes a heavy working memory load. The claim that the spacing effect is a cognitive load effect that is obtainable with simple materials that are low in element interactivity requires further discussion, as it apparently contradicts the usual cognitive load theory assumption that cognitive load effects require complex materials that are high in element interactivity. Most cognitive load effects are caused by an instructional procedure that overwhelms working memory during learning. Learning is enhanced when compared with an alternative instructional procedure that reduces working memory load. For example, presenting information in a split-attention format requires the use of more working memory resources than presenting the same material in a physically integrated format, leading to the split-attention effect. If intrinsic cognitive load is low because element interactivity is low, the effect will not be obtained because even under split-attention conditions, total cognitive load does not exceed working memory resources.

Working memory depletion may be more likely under high element interactivity conditions caused by extraneous cognitive load and so contribute to cognitive load effects, but depletion is not required as an explanation. The effects can be explained by cognitive load theory without any recourse to depletion effects.

The spacing effect is different. The effect is not obtained because working memory is overwhelmed at a given point due to high element interactivity. Rather, we have hypothesised that the effect is obtained because working memory is depleted when learning continuously without rest-from-deliberate-learning. Currently, we do not know whether depletion is greater using high than low element interactivity information. It may well be greater but given the large number of examples of the spacing effect using simple memorisation tasks that are low in element interactivity, the effect clearly does not require high element interactivity information. The fact that most cognitive load effects require high element interactivity information does not contradict the hypothesis that working memory depletion can occur using high or low element interactivity information because most cognitive load effects do not require the incorporation of depletion into their explanation.

Chen et al. (2018) designed two experiments to test the hypothesis that the spacing effect is due to working memory resource depletion. Primary school students were assigned to either massed or spaced practice when learning fraction addition. Massed practice was completed in 1 day. The same amount of practice was completed by the spaced practice group over a period of 3 days. For both groups, after all practice sessions were completed, a working memory capacity test was administered immediately prior to the content test phase. For massed practice, all learning activities, including the working memory resources test and post-content test, were conducted within a day, whereas the spaced practice group received the working memory resources test and post-content test a day after the third day of practice. Accordingly, the working memory resources test for the massed group was administered immediately after the learning phase and immediately before the content test while the same test for the spaced group was administered after a rest-from-deliberate-learning but also just before the content test. This procedure allowed differences in the working memory test results between the two groups to indicate whether any depletion due to deliberate learning was reversed after rest and whether that reversal was associated with differences on the content test. The results showed that spaced practice resulted in superior learning to massed practice with that superiority associated with increased availability of working memory resources.

Spacing Effect and Rest-from-Deliberate-Learning

A spacing gap can be either across individual items that together constitute a single concept with those items massed or spaced, or across practice sessions (Carpenter, 2014). In all studies classified in this paper as testing the effect of spacing, the space allows rest-from-deliberate-learning rather than being filled by another task that requires deliberate learning. It does not refer to rest from other cognitive activities such as for example, subtracting from 100 in steps of 7 that may impose a heavy cognitive load but with any learning being incidental rather than deliberate. None of the studies classified as spacing effect studies in this paper intentionally presented participants with explicit learning tasks although during long periods of rest-from-deliberate-learning, some of the uncontrolled activities of participants may have included intentional learning tasks. For our purposes, the only relevant criterion to be included during the rest-from-deliberate-learning phase is that learners had periods when intentional learning did not occur. Critically, for the interleaving effect, the “resting” period is filled with deliberate learning and it is that factor that distinguishes spacing from interleaving experiments. The effect of mental rest-from-deliberate-learning on cognitive function may be analogous to resting muscles after heavy physical exercise.

Explanations for the Spacing Effect

As indicated above, based on the rest-from-deliberate-learning and working memory resource depletion assumption, Chen et al. (2018) suggested that the spacing effect is due to working memory resource depletion during learning with recovery after rest-from-deliberate-learning. There are many earlier explanations. Delaney et al. (2010) favored a study–phase hypothesis according to which the gaps between learning sessions increase forgetting, so more effortful retrieval is required for the next learning session compared to massed learning. It follows from the study–phase hypothesis that an even larger spacing effect should be obtained if the rest-from-deliberate-learning period is replaced by other, mentally taxing activities that might result in even more forgetting of the target activity and so requiring even more effortful retrieval. Accordingly, the study–phase hypothesis and cognitive load theory generate opposing predictions. These predictions are incidentally tested by interleaved practice, discussed below.

There are earlier suggestions such as *rehearsal*, *consolidation*, *voluntary–attention*, *habituation*, and *encoding–variability* (Hintzman, 1974) that have been proposed as explanations of the spacing effect. The rehearsal explanation suggests that learners continue to process the learned material during spacing and so memory of the information is consolidated. During massed practice, interference from the following information during rehearsal will occur (Atkinson & Shiffrin, 1968; Delaney et al., 2012; Greeno, 1967; Rundus, 1971). The consolidation explanation is similar to the rehearsal explanation in that if no intrusion comes from the immediately following information, the learner engages in strengthening and consolidating the memory trace of the initial event, even though the initial event is absent (Delaney et al., 2012; Landauer, 1969, 1974). These explanations contradict the working memory recovery hypothesis since they assume that the information continues to be processed during “rest” periods, suggesting that real rest is not occurring and so working memory recovery also should not occur.

The encoding–variability explanation assumes that the greater the variability of presentations of the information, the more successful is encoding. Increasing the intervals between presentations increases the chance of similar information being coded differently thus

increasing the strength of memory traces (e.g., Maddox, 2016; Martin, 1972; McFarland Jr. et al., 1979). Assuming that differential coding of the same information increases the burden on working memory, this explanation also contradicts an increase in working memory resource availability when information or activity is spaced.

Both the voluntary-attention explanation (Hintzman, 1974, 1976; Koval, 2019; Underwood, 1969, 1970) and the habituation explanation (Hintzman, 1974; Koval, 2019) suggest that there is insufficient attention paid to the following information if it arrives too soon after the initial information or activity. Both explanations are in accord with the working memory recovery hypothesis since a decrease in attention or habituation due to subsequent cognitive activity is consistent with the effect of a decrease in working memory resources due to activity.

The Discriminative-Contrast Hypothesis and Interleaved Practice

Blocked practice occurs when all episodes associated with learning skill A are presented before switching to all episodes of learning skill B (e.g., AABB). Alternatively, interleaved practice occurs when each episode of learning skill A is alternated with an episode of skill B (e.g., ABAB). Generally, research has shown that interleaved practice is superior to blocked practice (e.g., Kang & Pashler, 2012; Kornell & Bjork, 2008; Taylor & Rohrer, 2010; Rohrer & Taylor, 2007; Shea & Morgan, 1979; Wahlheim et al., 2011a, b). The interleaving effect can be explained by the *discriminative-contrast hypothesis*.

Early research on interleaved practice was built on the contextual interference hypothesis (Battig, 1972) which makes predictions identical to the discriminative-contrast hypothesis. Both hypotheses assume that increased contextual interference during practice could produce more distinctive and elaborative processes that allow learners to distinguish between similar categories of information. The discriminative-contrast hypothesis was initially proposed by Kornell and Bjork (2008) to deal with category learning, such as distinguishing types of birds or types of paintings. Kang and Pashler (2012) designed a study to test the hypothesis. Different paintings from different artists were interleaved for discrimination. In order to test the influence of discriminative practice, participants in one group viewed the paintings from the same artist that were interleaved by cartoons that were very easy to discriminate from paintings. Zulkiply and Burt (2013) used a similar design to Kang and Pashler (2012). In both studies, interleaved practice was found to be more effective for materials with a low degree of discriminability, supporting the discriminative-contrast hypothesis.

Explanations for the Interleaving Effect

A common explanation for the interleaving effect is that since interleaving automatically includes spacing, interleaving is simply an example of the spacing effect and so can be explained in an identical manner. As indicated below, we doubt that that explanation is viable. The discriminative-contrast or contextual interference hypothesis provides an alternative. Rohrer et al. (2014) extended that hypothesis. Their association hypothesis, built on the discriminative-contrast hypothesis, suggests that students not only learn to discriminate among and contrast problems but also associate a specific response strategy to individual problems when those problems are presented in interleaved form. The association hypothesis may possibly explain the interleaving effect when using problems that may be easier to discriminate

between, although it is not clear why block practice would not also facilitate associations. Possibly, novice learners do need to learn to discriminate between problems from the same field (e.g., mathematics) even if for a more expert learner, they are completely different.

Detailed evidence in the literature for the mental rest-from-deliberate-learning and discriminative-contrast hypotheses

In this study, literature about interleaving and spacing effects was systematically searched and analysed to test the hypotheses that the spacing effect is due to mental rest-from-deliberate-learning while the interleaved effect can be explained by the discriminative-contrast hypothesis. An interleaved design necessarily results in spacing, increasing the difficulty of establishing a clear boundary and definition of interleaving and spacing effects. Research studies on spaced and interleaved practices were systematically searched using Scopus, Web of Science, ERIC, and PsycARTICLES as databases. By using “Spacing effect” AND “Learning” and “Interleaving effect” AND “Learning” separately as keywords, there were a total of 657 studies found and screened from the four databases. We conducted our searches from the 10th of May 2020 to the 20th of August 2020 using Scopus and Web of Science. ERIC and PsycARTICLES searches were completed by the 6th of January 2021.

Inclusion Criteria

The following inclusion criteria for the review were used: (a) the language for publication was English; (b) a quantitative measurement of performance was included; (c) across all included studies, participants were students across all stages; (d) publications were in journals, conference proceedings, or books.

Exclusion Criteria

Some conditions were used to exclude some of the searched studies: (a) the language of publication was not English; (b) the authors did not report an experimental study; (c) the authors did not measure learning but instead measured other factors such as motivation.

Article Screening

Based on inclusion and exclusion criteria, the first author screened each study by reading the abstract initially, leaving 150 studies for further screening. Thereafter, the full texts were read, focusing more on experimental materials and procedures used, resulting in the exclusion of a further 31 studies. This procedure left 119 studies to be included for analysis.

Coding of Included Studies

Two tables were created to code studies demonstrating successful spacing and interleaving effects separately. In this systematic review, rest-from-deliberate-learning is defined as activities that did not require learners to process information in working memory that needed to be transferred to long-term memory for later use. Accordingly, rest-from-deliberate-learning activity tasks were not targeted for learning and testing. For example, the play activity between two targeted items in the study of Vlach et al. (2008) was classed as a rest-from-

deliberate-learning activity. Based on this classification of rest-from-deliberate-learning, 48 studies were classed as testing for the spacing effect (Table 1) and 67 as testing for the interleaving effect (Table 2). In addition, 2 studies of the 67 classed as interleaving studies along with another 3 studies that could not be obviously classified as either spacing or interleaving studies based on rest-from-deliberate-learning found neither spacing nor interleaving effects.

In Table 1 for the spacing effect, which included rest-from-deliberate-learning for all studies, the factors coded were (a) the authors and year of publication; (b) the major rest-from-deliberate-learning activity and its difficulty; (c) the complexity of materials; (d) the length of rest-from-deliberate-learning; (e) effect sizes. In Table 2 for the interleaving effect, none of which included rest-from-deliberate-learning, the factors coded were (a) the authors and year of publication; (b) the degree of similarity of the materials; (c) effect sizes.

Analysis of the Literature Search

The Spacing Effect and Evidence Supporting Hypothesis 1 Except for Bego et al. (2017) published in conference proceedings and Ebbinghaus (1885) published in a book, the other studies were all journal articles. All of the 48 spaced practice studies (see Table 1) obtained the spacing effect and used a rest-from-deliberate-learning period either by distributing learning sessions across multiple days or months or by inserting rest-from-deliberate-learning activities between learning sessions for 30 s to hours or days. From Table 1, it can be seen that the majority of rest-from-deliberate-learning activities comprised sleeping and taking a break between two learning sessions. The type of break varied between studies, from direct breaks with no tasks to playing or reading some non-targeted passages. In all cases, the break could result in mental rest-from-deliberate-learning, such as sleeping, allowing the restoration of working memory resources that had been depleted during the learning session. Additionally, activities that were in some cases inserted during breaks but which differed from the targeted learning materials or concepts and were not relevant to learning also fell within our definition of rest-from-deliberate-learning. The length of rest-from-deliberate-learning also varied, from 30 s to a few days. Eleven of the 48 studies found a spacing effect with very short rest-from-deliberate-learning periods (30 s but less than 12 h), but the majority of studies found a spacing effect with a long-period rest-from-deliberate-learning. Thirty of the 48 studies used simple experimental materials, such as word pairs. Although the materials were themselves simple, during the experiments, participants had to complete multiple trials continuously.

Thirty-seven of the 47 studies included sleeping as rest-from-deliberate-learning. The remaining ten studies did not include sleeping but included breaks, such as playing or reading learning-irrelevant passages to rest learners from deliberate learning. The main function of breaks was to take students' attention away from learning, similar to the breaks between classes in schools. Based on reported effect sizes in Table 1, the spacing effect is robust for varied lengths of resting, from 30 s to a few days. For example, a large effect was found for resting of 30 to 50 s (Vlach et al., 2008).

The positive effects of spaced practice have been supported by the research studies of Table 1 in many topic areas from learning motor movements (Dail & Christina, 2004a, b) to learning physics (Grote, 1995). Other studies, focusing on recalling word lists, also showed superior results of spaced practice (Cain & Willey, 1939). Rea and Modigliani (1985) found spaced practice with young children when learning multiplication facts and spelling lists.

Table 1 Summary of spaced practice studies

Study	Major type of rest-from-deliberate-learning	Materials	Length of rest-from-deliberate-learning	Effect size (Cohen's <i>d</i>)
Ebbinghaus (1885)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple	20 min or 1 day	/
Cain and Willey (1939)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple	1 to few days	.99
Grote (1995)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Complex	Few days	.66
Gluckman et al. (2014)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple and complex	1 day	1.42
Gurung and Burns (2019)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Complex	1 week	1.36
Dobson et al. (2017)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple	2 or 3 days	3.40
Chen et al. (2018)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Complex (exp 1) Complex (exp 2)	1 day	.70 .52
Kapler et al. (2015)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Complex	1 or 8 days	.61
Küpper-Tetzel et al. (2014)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple	1 to few days	.57
Toppino et al. (2018)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple	Few days	.31
Barzagar Nazari and Ebersbach (2019)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Complex	1 day	.83
Bercovitz et al. (2017)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple	1 or few days	1.24
Slone and Sandhofer (2017)	<i>Breaks</i> : Play activities or read irrelevant books, no learning, rest-from-deliberate-learning used between different learning sessions, no learning-relevant activity or no engaging in learning, but relaxing	Simple	2 min	.94
Young et al. (2019)	<i>Breaks</i> : Read irrelevant passages, no learning, rest-from-deliberate-learning used between different learning sessions, no learning-relevant activity or no engaging in learning, but relaxing	Simple	10 min	1.24
Breckwoldt et al. (2016)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Complex	1 day	.56
Schutte et al. (2015)	<i>Breaks</i> : Rest-from-deliberate-learning used between different learning sessions, no learning-relevant activity or no engaging in learning, but relaxing	Complex	2 or 3.35 h	.72
Pyc et al. (2014)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple for 5 experiments	1 day	1.55
Whitten II and Bjork (1977)	<i>Breaks</i> : Shadowing tasks, rest-from-deliberate-learning used between different learning sessions, no learning-relevant activity or no engaging in learning, but relaxing	Simple	2.2 min	/

Table 1 (continued)

Study	Major type of rest-from-deliberate-learning	Materials	Length of rest-from-deliberate-learning	Effect size (Cohen's <i>d</i>)
Bego et al. (2017)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Complex	1 week	.26
Stafford and Haasnoot (2017)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Complex	12 h	1.50
Lotfolahi and Salehi (2017)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple	1 week	1.43
Lotfolahi and Salehi (2016)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple	1 week	2.00
Spruit et al. (2015)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple and complex	1 week	.63
Kwon et al. (2015)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple	12 h	.64
Miles (2014)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Complex	1 day	2.14
Samudra et al. (2019)	<i>Breaks</i> : Rest-from-deliberate-learning used between different learning sessions, no learning-relevant activity or no engaging in learning, but relaxing	Simple	1 hour	.50
Smith and Kimball (2010)	<i>Breaks</i> : Rest-from-deliberate-learning used between different learning sessions, no learning-relevant activity or no engaging in learning, but relaxing	Simple	16 – 20 minutes	.79
Kang et al. (2014)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple	Few days	.38
Vlach et al. (2014)	<i>Breaks</i> : Play activities, no learning, rest-from-deliberate-learning used between different learning sessions, no learning-relevant activity or no engaging in learning, but relaxing	Simple	30 s or 50 s	1.24
Simone et al. (2013)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple	24 h	2.29
Jackson et al. (2013)	<i>Breaks</i> : Sitting and rest-from-deliberate-learning, no tasks given, rest-from-deliberate-learning used between different learning sessions, no learning-relevant activity or no engaging in learning, but relaxing	Complex	15 min	2.41
Goossens et al. (2012)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple	1 day	.36
Glover and Corkill (1987)	<i>Breaks</i> : Sitting and rest-from-deliberate-learning, no tasks given, rest-from-deliberate-learning used between different learning sessions, no learning-relevant activity or no engaging in learning, but relaxing	Complex	30 min	/
Gerbier and Koenig (2012)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple (exp 1) Simple (exp 2)	Few days	.33
Vlach et al. (2012)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple and complex Simple	1 day 7 days	1.02 .48

Table 1 (continued)

Study	Major type of rest-from-deliberate-learning	Materials	Length of rest-from-deliberate-learning	Effect size (Cohen's <i>d</i>)
Sobel et al. (2011)	<i>Sleeping</i> : The learning sessions were distributed across multiple days			
Budé et al. (2011)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple	1 day or few days	.83
Kerfoot et al. (2010)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Complex	1 week	.75
Goverover et al. (2009)	<i>Breaks</i> : Rest-from-deliberate-learning used between different learning sessions, no learning-relevant activity or no engaging in learning, but relaxing	Simple	5 min	.26
Vlach et al. (2008)	<i>Breaks</i> : Play activities, no learning, rest-from-deliberate-learning used between different learning sessions, no learning-relevant activity or no engaging in learning, but relaxing	Simple	30 s	2.98
Kerfoot et al. (2007)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Complex	1 week	.68
Ambridge et al. (2006)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple for 2 experiments	1 day	1.20
Dail and Christina (2004a, b)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Complex	Few days	/
Bahrick et al. (1993)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple	Few days	/
Rohrer et al. (2014)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Complex	Few days	1.04
Rohrer et al. (2015)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Complex	Few days	.79
Cepeda et al. (2008)	<i>Sleeping</i> : The learning sessions were distributed across multiple days	Simple	Few days	1.3
Yan and Sana (2020)	<i>Sleeping</i> : The learning sessions were distributed across multiple days (exp 1, 2)	Simple (exp 1)	1 day	.59
		Simple (exp 3)		.44

Note: The missing effect sizes are due to either a lack of information or no response from authors concerning raw data sets for computing Cohen's *d*

Although some studies found superior results of spaced practice with very short study interventions (seconds: e.g., Whitten II & Bjork, 1977), others have obtained the effect with longer interventions (days or weeks: Dobson et al., 2017; Chen et al., 2018). The duration of an intervention seems not to be a moderator for spaced practice (Gerbier et al., 2015; Godbole et al., 2014; Kapler et al., 2015).

Two types of spaced practice that are discussed are uniform spaced practice and altering spacing intervals. Whereas uniform spaced practice consists of fixed or equal intervals between two or more learning episodes, in expanding spaced practice, intervals are gradually increased while in contracting spaced practice, intervals are gradually reduced. Küpper-Tetzel et al. (2014) investigated the effectiveness of uniform, contracting and expanding spaced

Table 2 Summary of interleaved practice studies

Study	Materials	Effect size (Cohen's <i>d</i>)	Study	Materials	Effect size (Cohen's <i>d</i>)
Kang and Pashler (2012)	<i>Similar</i> : Contrasting and discriminating paintings with similar styles and painting subject	.56	Kornell and Bjork (2008)	<i>Similar</i> : Contrasting and discriminating paintings from the same artist or other artists	1.60
Taylor and Rohrer (2010)	<i>Similar</i> : Contrasting and discriminating different types of prism problems and their formulae	1.21	Shea and Morgan (1979)	<i>Similar</i> : Contrasting and discriminating different orders of barrier to be knocked down	/
Wahlheim et al. (2011a, b)	<i>Similar</i> : Contrasting and discriminating different birds from the similar families	.97	Foster et al. (2019)	<i>Similar</i> : Contrasting and discriminating different types of volume problems and their formulae	.62
Carson and Wiegand (1979)	<i>Similar</i> : Contrasting and discriminating different bean bags with varied weights	/	Hall et al. (1994)	<i>Similar</i> : Contrasting and discriminating three types of pitches	1.00
Keller et al. (2006)	<i>Similar</i> : Contrasting and discriminating types of shooting tasks with varied distances etc.	.15	Landin et al. (1993)	<i>Similar</i> : Contrasting and discriminating types of basketball shooting tasks with varied distances etc.	1.37
Kornell et al. (2010)	<i>Similar</i> : Contrasting and discriminating paintings from the same artist or other artists	1.00	Zulkipli and Burt (2013)	<i>Similar</i> : Contrasting and discriminating paintings from the same artist or other artists	.92
Carvalho and Goldstone (2014)	<i>Similar</i> : Contrasting and discriminating blob figures (created by randomly generating curvilinear segments) sharing most features with each other	/	Wong et al. (2020)	<i>Similar</i> : Contrasting and discriminating paintings from the same artist or other artists	.34
Pan et al. (2019)	<i>Similar</i> : Contrasting and discriminating Spanish verbs with types of past tense	.53	Rozenshtein et al. (2016)	<i>Similar</i> : Contrasting and discriminating 12 common chest radiographic patterns	/
Zulkipli (2013)	<i>Similar</i> : Contrasting and discriminating types of cases of psychopathological disorders	1.03	Wang et al. (2020a)	<i>Similar</i> : Contrasting and discriminating word pairs or picture pairs	/
Guzman-Munoz (2017)	<i>Similar</i> : Contrasting and discriminating paintings from the same artist or other artists	1.54	Zulkipli (2015)	<i>Similar</i> : Contrasting and discriminating paintings from the same artist or other artists	/
Clapper (2015)	<i>Similar</i> : Contrasting and discriminating verbal characteristics	2.25	Ostrow et al. (2015)	<i>Similar</i> : Contrasting and discriminating 3	.61

Table 2 (continued)

Study	Materials	Effect size (Cohen's <i>d</i>)	Study	Materials	Effect size (Cohen's <i>d</i>)
	supposedly describing fictitious species of insects or trees (12 verbal features plus a verbal label)			similar mathematics skills (Exp 1)	
Sana et al. (2017)	<i>Similar</i> : Contrasting and discriminating 3 statistical concepts	.36	Perruchet (1989)	<i>Similar</i> : Contrasting and discriminating words	1.25
Toppino and DiGeorge (1984)	<i>Similar</i> : Contrasting and discriminating a list of pictures	/	Balota et al. (2006)	<i>Similar</i> : Contrasting and discriminating word pairs with individuals with dementia of the Alzheimer type	1.04
Godbole et al. (2014)	<i>Similar</i> : Contrasting and discriminating 60 medium-frequency nouns	.61	Wang et al. (2020b)	<i>Similar</i> : Contrasting and discriminating paintings from the same artist or other artists	.54
Kim et al. (2018)	<i>Similar</i> : Contrasting and discriminating 175 one-syllable nouns taken from the MRC Psycholinguistic Database	.58	Mettler et al. (2016)	<i>Similar</i> : Contrasting and discriminating the outlines of different African countries	.57
Maass et al. (2015)	<i>Similar</i> : Contrasting and discriminating sentences with statistical concepts	/	Metcalf and Xu (2016)	<i>Similar</i> : Contrasting and discriminating paintings from the same artist or other artists	.43
Zigterman et al. (2015)	<i>Similar</i> : Contrasting and discriminating word pairs	1.49	Kornmeier et al. (2014)	<i>Similar</i> : Contrasting and discriminating German – Japanese pairs	/
Middleton et al. (2019)	<i>Similar</i> : Contrasting and discriminating 700-item proper noun picture corpus	/	Vlach et al. (2019)	<i>Similar</i> : Contrasting and discriminating pictures of common objects	.51
Wang et al. (2017)	<i>Similar</i> : Contrasting and discriminating unfamiliar faces	/	Swehla et al. (2016)	<i>Similar</i> : Contrasting and discriminating word pairs	1.13
Vlach and Kalish (2014)	<i>Similar</i> : Contrasting and discriminating novel objects and novel linguistic labels	.91	Green et al. (2014)	<i>Similar</i> : Contrasting and discriminating a list of nouns	2.5
Verkoeijen and Bouwmeester (2014)	<i>Similar</i> : Contrasting and discriminating paintings from the same artist or other artists	.37	Rawson and Dunlosky (2013)	<i>Similar</i> : Contrasting and discriminating eight key terms and their definitions	1.56
Logan et al. (2012)	<i>Similar</i> : Contrasting and discriminating 19 medium-frequency words	.71	Zulkipli et al. (2012)	<i>Similar</i> : Contrasting and discriminating types of cases of psychopathological disorders	1.55

Table 2 (continued)

Study	Materials	Effect size (Cohen's <i>d</i>)	Study	Materials	Effect size (Cohen's <i>d</i>)
Xue et al. (2011)	<i>Similar</i> : Contrasting and discriminating unfamiliar faces	/	Toppino and Cohen (2010)	<i>Similar</i> : Contrasting and discriminating a list of nouns	/
Son (2010)	<i>Similar</i> : Contrasting and discriminating 60 synonym pairs	/	Kornell (2009)	<i>Similar</i> : Contrasting and discriminating 40 synonyms	.48
Toppino et al. (2009a, b)	<i>Similar</i> : Contrasting and discriminating word pairs	1.10	Toppino et al. (2009b)	<i>Similar</i> : Contrasting and discriminating drawings of common objects	/
Karpicke and Roediger III (2007)	<i>Similar</i> : Contrasting and discriminating word pairs	.50	Son (2004)	<i>Similar</i> : Contrasting and discriminating 60 synonym pairs	.39
Mammarella et al. (2002)	<i>Similar</i> : Contrasting and discriminating unfamiliar faces and nonwords	6.63	Toppino and Bloom (2002)	<i>Similar</i> : Contrasting and discriminating 22 high-frequency nouns and 22 high-frequency adjectives	/
Russo and Mammarella et al. (2002)	<i>Similar</i> : Contrasting and discriminating words and nonwords	.65	Braun and Rubin (1998)	<i>Similar</i> : Contrasting and discriminating words	/
Russo et al. (1998)	<i>Similar</i> : Contrasting and discriminating nonwords	1.56	Shea and Kohl (1991)	<i>Similar</i> : Contrasting and discriminating physical tasks on converting physical force into a voltage	/
Shaughnessy (1976)	<i>Similar</i> : Contrasting and discriminating words	/	Bray et al. (1976)	<i>Similar</i> : Contrasting and discriminating word pairs	/
Shaughnessy et al. (1974)	<i>Similar</i> : Contrasting and discriminating word pairs	1.62	Cermak et al. (1996)	<i>Similar</i> : Contrasting and discriminating 68 high-frequency nouns	/
Middleton et al. (2016)	<i>Similar</i> : Contrasting and discriminating picture corpus of 700 entities named with proper nouns	/	Rohrer et al. (2014)	<i>Similar</i> : Contrasting and discriminating 4 types of mathematics problems (considering learners' low level of expertise)	1.05
Rohrer et al. (2015)	<i>Similar</i> : Contrasting and discriminating types of mathematics problems	.79	Maddox and Balota (2015)	<i>Similar</i> : Contrasting and discriminating word pairs	11.51
Rohrer and Taylor (2007)	<i>Similar</i> for 2 Experiments: Contrasting and discriminating multiple permutation tasks	1.34			

Note: The missing effect sizes are due to either a lack of information or no response from authors concerning raw data sets for computing Cohen's *d*

practice. Twenty-eight word pairs were created by using 56 concrete and highly familiar nouns with no obvious semantic associations among them. For uniform spaced practice, the learning

intervals were 3 days, and for expanding spaced practice, the first learning interval was 1 day and then 5 days with contracting spaced practice moving from 5 to 1 day. Results indicated that expanding and uniform spaced practices were better for long retention intervals compared to contracting spaced practice, which was better for shorter retention intervals. A clear superiority of expanding spaced practice was found by Toppino et al. (2018) and Vlach and Sandhofer (2012). Generally speaking, expanding spaced practice is superior to uniform spaced practice (Cepeda et al., 2006) and contracting spaced practice is the least effective for long-term retention.

The results of these studies are consistent with the hypothesis that spaced practice is a cognitive load effect due to working memory resource depletion. Chen et al. (2018) provided evidence that working memory resources decreased after cognitive effort and recovered after rest-from-deliberate-learning. Based on the definition of rest-from-deliberate-learning, all included studies were classified as spaced practice without necessarily following the original definition of spacing or interleaving in the reported study.

The Interleaving Effect and Evidence Supporting Hypothesis 2 As indicated in Table 2, Maass et al. (2015) and Ostrow et al. (2015) reported an interleaving effect in conference proceedings with all other experiments reported in journal articles. All 67 studies demonstrated the interleaving effect and used very similar interleaved concepts during learning. They shuffled targeting concepts or inserted lags (i.e., intervening concepts) between two targeted concepts without *no mental* rest-from-deliberate-learning. A lag effect is categorised as a spacing effect in many studies, as different concepts or skills were spaced, but based on our definition of rest-from-deliberate-learning, introducing a lag should be classed as interleaving, as no rest-from-deliberate-learning was inserted although there was spacing. The 67 studies testing for interleaved practice (see Table 2) are in accord with the hypothesis that interleaved practice is due to learning to discriminate and contrast. The general design testing for the interleaving effect is as follows:

1. Shuffling: ABCBCACAB..... [where A, B, C are targeted skills]
2. Lagging: AaaaBbbbCccc..... [where A, B, C are targeting skills that are from a similar domain to the skills associated with a, b, c]

An interleaving effect has been found on inductive learning in which learners acquire a concept or category by observing exemplars. In Kornell and Bjork's (2008) study, a given artist's paintings were presented in blocked practice or interleaved with other artists' paintings as interleaved practice. Participants were required to discriminate between paintings by different artists and indicate which of the previously seen artists, if any, painted each of a series of new paintings. Similar results were reported by Kornell et al. (2010). Wahlheim et al. (2011a, b) presented bird species either individually or in pairs. Participants were better able to recognise different bird species in pairs. Carvalho and Goldstone (2014) designed a series of blobs, created by randomly generating curvilinear segments, to investigate interleaved practice. When designing those blobs, they varied not only similarity between—but also within—categories. The results again found that interleaved practice was more effective for high-similarity categories.

In addition to category learning, some research studies in other domains, such as Mathematics, Music, and Language, also supported the discriminative-contrast hypothesis with interleaved practice (Rohrer & Taylor, 2007; Taylor & Rohrer, 2010). Rohrer

and Taylor (2007) compared interleaved practice with blocked practice in two experiments. The results of both studies showed interleaved practice to be more effective than blocked practice. Taylor and Rohrer (2010) designed four types of mathematics problems that were either blocked or interleaved. They contrasted and controlled the spacing factor in order to test the interleaving effect only. Interleaved practice greatly enhanced learning performance by improving students' ability to associate each problem with its correct solution. Wong et al. (2020) chose music pieces composed by six musicians. Participants were either in the interleaved condition where the music pieces from different composers were alternated, whereas, in the blocked condition, participants listened to all music pieces from one composer then switched to listening to all music pieces from another composer. Interleaved practice was found to be more effective than blocked practice in distinguishing styles of different composers. A study of interleaved practice in foreign language learning was reported by Pan et al. (2019). In their experiment, interleaved practice was designed for learning Spanish grammar and found to be superior to blocked practice. Rozenshtein et al. (2016) presented six examples of 12 common chest radiographic patterns in either interleaved or blocked practice format. Interleaved practice enabled participants to better distinguish and recognise differences in patterns than blocked practice, which supported the discriminative-contrast hypothesis.

It needs to be noted that for some studies, such as Rohrer et al. (2015), it is not possible to disentangle spacing and interleaving effects. In these studies, learning sessions were distributed across days, allowing sleeping/breaks for restoring working memory resources (testing the spacing effect), and in addition, including interleaving within each assignment for contrasting and discriminating purposes (testing for the interleaving effect). It also should be noted that in the study of Rohrer et al. (2014), while dissimilar mathematics problems were used, students were novices with respect to those problems and so had to learn which solution was appropriate, an activity which required them to contrast and discriminate between the problems leading to the interleaving effect.

Null Effects of Spacing and Interleaving Tables 1 and 2 list positive examples of spacing and interleaving, respectively. They are in accord with the two hypotheses that the spacing effect requires a mental rest-from-deliberate-learning period while the interleaving effect requires two tasks that students need to learn to discriminate between. Of course, stronger evidence for the mental rest-from-deliberate-learning hypothesis would be obtained if it could be shown that spacing without mental rest-from-deliberate-learning reduced or failed to demonstrate the spacing effect while interleaving domains that were obviously different reduced or failed to demonstrate the interleaving effect.

There were five studies that (1) *all* used spacing between different concepts or skills with no rest-from-deliberate-learning, resulting in no spacing effect; (2) *all* used materials that were very dissimilar, resulting in no need to contrast and discriminate materials and so finding no interleaving effect. These null effects support both hypotheses. The relevant studies were all interleaving experiments in which the domain sets were obviously different and equally difficult. Thus, all the studies used experimental designs in which there was spacing with no rest-from-deliberate-learning and interleaving without the need to learn to discriminate between domains. With the exception of the study of Ostrow and Heffernan (2015) that was published in a conference proceeding, the other four studies were all journal articles.

Ostrow and Heffernan (2015) compared an interleaved design (A1, A2, B1, B2, C1, C2, A3, B3, C3, B4, C4, A4) with a blocked design (A1, A2, A3, A4, B1, B2, B3, B4,

C1, C2, C3, C4). They found no difference between conditions with knowledgeable learners. There were three skills taught: Complementary and Supplementary Angles, Surface Area of a Pyramid, and Compound Probability without Replacement. Although the different skills were automatically spaced in the interleaved design, no rest-from-deliberate-learning was allowed, and so the spacing effect was not generated, in accord with Hypothesis 1. In addition, as learners were knowledgeable in the domain, there was no need to contrast and discriminate similar concepts and skills; therefore, the interleaving effect was not generated, in accord with Hypothesis 2.

De Croock and Van Merriënboer (2007) investigated the effects of blocked practice and interleaved practice on troubleshooting skills in three dissimilar scenarios: One set dealt with injury cases, one with damage cases, and one with traffic cases, and found no difference between groups. A design similar to Ostrow and Heffernan (2015) was used. The failure to find the effects due to either spacing or interleaving with neither rest-from-deliberate-learning nor a need to contrast or discriminate again supported both hypotheses.

Carpenter and Mueller (2013) presented a series of French words that represented different pronunciation rules. The words were arranged either blocked by rule or interleaved by rule. The performance of blocked practice was superior to that of interleaved practice. The explanation was that there was no discrimination required when learning those words and rules. No spacing effect was found as rest-from-deliberate-learning was not allowed. Similarly, when asking students to learn word pairs, such as anatomy names and Indonesian words, in either interleaved or massed practice, a null effect was found by Hausman and Kornell (2014). They suggested that word pairs might be very easily distinguished, so the discriminative-contrast hypothesis was not tested. Since there was no rest-from-deliberate-learning period, the spacing effect was again not obtained.

Yan and Sana (2020) investigated the interleaving effect by interleaving either one level (domain or concept) or two levels (domain and concept); the concepts were chosen from two unrelated domains: physics and statistics. When the two levels were both interleaved, there was no significant difference between an interleaved design and blocked design. This result again supports our hypotheses since there was neither rest-from-deliberate-learning nor difficult to discriminate domains or concepts.

Similarity of Materials The materials used in Table 2 and the five studies indicating no effect may need further discussion. With respect to Table 2, all materials were chosen from *the same domain or area*, which were not easily distinguishable by novice learners. Often, participants had to contrast and discriminate between formulae. Paintings were another type of materials used by many of the studies in Table 2. Participants needed to contrast and discriminate the paintings by the style of each painter, a very subtle distinction. Other studies used words or word pairs from the same language as stimuli for experiments. Many studies used words or word pairs that were unfamiliar to participants, such as nonsense words, or low - or medium-frequency words. Of the five studies indicating no effect, Carpenter and Mueller (2013) and Hausman and Kornell (2014) also used word pairs, but they were easily distinguishable by the use of different rules associating the word pairs or by their appearance. In some cases, anatomy names and foreign words needed to be distinguished. In the five studies indicating no effect, the experimental materials were clearly different reducing the need for learners to learn to discriminate between them.

Limitations and Future Studies

The concept of mental rest-from-deliberate-learning requires further empirical evidence to determine relevant parameters. Some factors, such as the length of resting and the levels of cognitive demands of the tasks that are being rested from, need empirical investigation. A meta-analysis also would be useful.

For the spacing effect, Chen et al. (2018) did not control learner activities during resting. While the rest was sufficiently long to ensure some rest-from-deliberate-learning did occur, more strictly controlled experiments would be useful to test whether working memory depletion is directly linked to controlled resting.

We have suggested that the spacing effect with its rest-from-deliberate-learning procedure under spaced conditions allows depleted working memory resources to be replenished. While results supporting this suggestion were obtained by Chen et al. (2018) in two experiments, we are not aware of any other studies that assessed working memory resources. Future research should directly test the hypothesis that the spacing effect is a cognitive load effect due to working memory depletion while the interleaving effect can be explained by the discriminative-contrast hypothesis.

Contributions of the Study

This systematic review aimed to provide evidence that the spacing and interleaving effects cannot be explained by the same theoretical base. The results of this study suggest that the two effects should be seen as theoretically distinct with the spacing effect as a cognitive load effect and the interleaving effect as a perceptual effect. From a practical perspective, the study provides practitioners with guidance on when to apply spaced instruction or when to apply interleaved instruction in classrooms. The conditions appear to be very different. Spacing should be used when learners need to mentally rest from learning while interleaving should be used when learners need to discriminate between apparently similar but in fact different instructional areas.

Conclusion

The data of this review suggest that the spacing effect needs a rest-from-deliberate-learning period between learning episodes. A requirement for a rest-from-deliberate-learning period can be directly tested using the procedures of the interleaving effect. Interleaving provides spacing but without rest-from-deliberate-learning because what would otherwise be a rest-from-deliberate-learning period is filled with the interleaved activity. Accordingly, if a rest-from-deliberate-learning period is required for the spacing effect, the interleaving effect should not occur. The fact that the interleaving effect does occur leads to the next question—are the interleaving and spacing effects caused by the same cognitive processes?

We suggest the requirement for a rest-from-deliberate-learning period is demonstrated by a particular category of experiments on the interleaving effect. Differences between interleaved

and blocked practice can be more difficult to obtain depending on the discriminability of the interleaved materials. For experiments with very obviously different materials that are immediately discriminable, the interleaving effect is unlikely to be obtained. It is this class of experiments that provide evidence that rest-from-deliberate-learning between episodes, not just spacing, is required for the spacing effect. Other experiments, with materials that are more difficult to discriminate between, provide clear evidence of the interleaving effect. The effect is more likely to be obtained if interleaved information is not readily discriminable. Interleaving may assist learners to discriminate between areas that are more difficult to distinguish apart, resulting in superiority of interleaving over blocking.

Based on this analysis, we suggest that rest-from-deliberate-learning periods permit recovery of depleted working memory resources as indicated by the results of Chen et al. (2018). That mechanism is not usually available in interleaving studies because in those studies, the spacing is due to interleaving rather than rest-from-deliberate-learning and so the interleaving effect is likely to have different causal factors. The suggested mechanism that drives the results of successful interleaving studies is the discriminative-contrast hypothesis. That hypothesis is not relevant in spacing studies because multiple areas that learners need to discriminate between are not used.

By distinguishing between studies that did and did not use rest-from-deliberate-learning, on the one hand, and did and did not use multiple, difficult to discriminate learning areas, on the other hand, the two effects can be clearly separated by two different cognitive mechanisms. The spacing effect can be explained by working memory depletion after cognitive effort and recovery during rest-from-deliberate-learning. In contrast, the explanation of interleaved practice focuses on discrimination between highly similar concepts with evidence showing that interleaved practice is not moderated by varying working memory capacity (Sana et al., 2018).

One of the consequences of interleaved practice automatically involving spaced practice is that the two procedures frequently and understandably are conflated (e.g., Taylor & Rohrer, 2010). Some research studies have tried to distinguish the two practices by inserting time delays within blocked practice to match the time delays of interleaved practice (e.g., Kang & Pashler, 2012; Taylor & Rohrer, 2010). However, in most cases, interleaved practice automatically combines interleaving and spacing manipulations (Kornell & Bjork, 2008), a fact that has erroneously led to the assumption that both effects have the same cause.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

(* indicating studies included in the review)

*Ambridge, B., Theakston, AL., Lieven, EV., & Tomasello, M. (2006). The distributed learning effect for children's acquisition of an abstract grammatical construction. *Cognitive Development*, 21, 174-193, 2.

- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The Psychology of Learning and Motivation* (Vol. 2, pp. 89–05). Academic Press, 968.
- *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, 5.
- *Balota, D. A., Duchek, J. M., Sergent-Marshall, S. D., & Roediger III, H. L. (2006). Does expanded retrieval produce benefits over equal-interval spacing? Explorations of spacing effects in healthy aging and early stage Alzheimer's disease. *Psychology and Aging*, 21, 19–31, 1.
- *Barzagar Nazari, K., & Ebersbach, M. (2019). Distributing mathematical practice of third and seventh graders: Applicability of the spacing effect in the classroom. *Applied Cognitive Psychology*, 33, 288–298, 2.
- Battig, W. F. (1972). Intrasubject interference as a source of facilitation in transfer and retention. In R. F. Thompson & J. F. Voss (Eds.), *Topics in learning and performance*. Academic Press.
- *Bego, C. R., Lyle, K. B., Ralston, P. A., & Hieb, J. L. (2017). Retrieval practice and spacing in an engineering mathematics classroom: Do the effects add up? In *2017 IEEE Frontiers in Education Conference (FIE)* (pp. 1–5). IEEE.
- *Bercovitz, K. E., Bell, M. C., Simone, P. M., & Wiseheart, M. (2017). The spacing effect in older and younger adults: Does context matter? *Aging, Neuropsychology, and Cognition*, 24, 703–716, 6.
- *Braun, K., & Rubin, D. C. (1998). The spacing effect depends on an encoding deficit, retrieval, and time in working memory: Evidence from once-presented words. *Memory*, 6, 37–65, 1.
- *Bray, J. F., Robbins, D., & Witcher, W. B. (1976). Encoding variability theory and the spacing effect in associate learning. *Memory & Cognition*, 4, 548–552, 5.
- *Breckwoldt, J., Ludwig, J. R., Plener, J., Schröder, T., Gruber, H., & Peters, H. (2016). Differences in procedural knowledge after a “spaced” and a “massed” version of an intensive course in emergency medicine, investigating a very short spacing interval. *BMC Medical Education*, 16, 249–258, 1.
- *Budé, L., Imbos, T., van de Wiel, M. W., & Berger, M. P. (2011). The effect of distributed practice on students' conceptual understanding of statistics. *Higher Education*, 62, 69–79, 1.
- *Cain, L. F., & Willey, R. D. V. (1939). The effect of spaced learning on the curve of retention. *Journal of Experimental Psychology*, 25, 209–214, 2.
- Carpenter, S. K. (2014). Spacing and interleaving of study and practice. *Applying the science of learning in education: Infusing psychological science into the curriculum*, 131–141.
- *Carpenter, S. K., & Mueller, F. E. (2013). The effects of interleaving versus blocking on foreign language pronunciation learning. *Memory & Cognition*, 41, 671–682, 5.
- *Carson, L. M., & Wiegand, R. L. (1979). Motor schema formation and retention in young children: A test of Schmidt's schema theory. *Journal of Motor Behavior*, 11, 247–251, 4.
- *Carvalho, P. F., & Goldstone, R. L. (2014). Putting category learning in order: Category structure and temporal arrangement affect the benefit of interleaved over blocked study. *Memory & Cognition*, 42, 481–495, 3.
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132(3), 354–380.
- *Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridge line of optimal retention. *Psychological Science*, 19, 1095–1102, 11.
- *Cermak, L. S., Verfaellie, M., Lanzoni, S., Mather, M., & Chase, K. A. (1996). Effect of spaced repetitions on amnesia patients' recall and recognition performance. *Neuropsychology*, 10, 219–227, 2.
- *Chen, O., Castro-Alonso, J. C., Paas, F., & Sweller, J. (2018). Extending cognitive load theory to incorporate working memory resource depletion: evidence from the spacing effect. *Educational Psychology Review*, 30, 483–501, 2.
- *Clapper, J. P. (2015). The impact of training sequence and between-category similarity on unsupervised induction. *The Quarterly Journal of Experimental Psychology*, 68, 1370–1390, 7.
- *Dail, T. K., & Christina, R. W. (2004a). Distribution of practice and metacognition in learning and long-term retention of a discrete motor task. *Research Quarterly for Exercise and Sport*, 75, 148–155, 2.
- Dail, T. K., & Christina, R. W. (2004b). Distribution of practice and metacognition in learning and long-term retention of a discrete motor task. *Research Quarterly for Exercise and Sport*, 75(2), 148–155.
- *De Croock, M. B., & Van Merriënboer, J. J. (2007). Paradoxical effects of information presentation formats and contextual interference on transfer of a complex cognitive skill. *Computers in Human Behavior*, 23, 1740–1761, 4.
- Delaney, P. F., Verkoeijen, P. P., & Spigel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. In *Psychology of Learning and Motivation* (Vol. 53, pp. 63–147). Academic Press.
- Delaney, P. F., Spigel, A. S., & Toppino, T. C. (2012). A deeper analysis of the spacing effect after “deep” encoding. *Memory & Cognition*, 40(7), 1003–1015.

- *Dobson, J. L., Perez, J., & Linderholm, T. (2017). Distributed retrieval practice promotes superior recall of anatomy information. *Anatomical Sciences Education*, *10*, 339-347, 4.
- *Ebbinghaus, H. (1885). *Memory* (translated by H. A. Ruger and C. E. Bussenius). New York, Teachers College, 1913. Paperback ed., New York, Dover, 1964.
- *Foster, N. L., Mueller, M. L., Was, C., Rawson, K. A., & Dunlosky, J. (2019). Why does interleaving improve math learning? The contributions of discriminative contrast and distributed practice. *Memory & Cognition*, *47*, 1088-1101, 6.
- *Gerbier, E., & Koenig, O. (2012). Influence of multiple-day temporal distribution of repetitions on memory: A comparison of uniform, expanding, and contracting schedules. *The Quarterly Journal of Experimental Psychology*, *65*, 514-525, 3.
- Gerbier, E., Toppino, T. C., & Koenig, O. (2015). Optimising retention through multiple study opportunities over days: The benefit of an expanding schedule of repetitions. *Memory*, *23*(6), 943-954.
- *Glover, J. A., & Corkill, A. J. (1987). Influence of paraphrased repetitions on the spacing effect. *Journal of Educational Psychology*, *79*, 198-199, 2.
- *Gluckman, M., Vlach, H. A., & Sandhofer, C. M. (2014). Spacing simultaneously promotes multiple forms of learning in children's science curriculum. *Applied Cognitive Psychology*, *28*, 266-273, 2.
- *Godbole, N. R., Delaney, P. F., & Verkoijen, P. P. (2014). The spacing effect in immediate and delayed free recall. *Memory*, *22*, 462-469, 5.
- *Goossens, N. A., Camp, G., Verkoijen, P. P., Tabbers, H. K., & Zwaan, R. A. (2012). Spreading the words: A spacing effect in vocabulary learning. *Journal of Cognitive Psychology*, *24*, 965-971, 8.
- *Goverover, Y., Hillary, F. G., Chiaravalloti, N., Arango-Lasprilla, J. C., & DeLuca, J. (2009). A functional application of the spacing effect to improve learning and memory in persons with multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology*, *31*, 513-522, 5.
- *Green, J. L., Weston, T., Wiseheart, M., & Rosenbaum, R. S. (2014). Long-term spacing effect benefits in developmental amnesia: Case experiments in rehabilitation. *Neuropsychology*, *28*, 685-694, 5.
- Greeno, J. G. (1967). Paired-associate learning with short-term retention: Mathematical analysis and data regarding identification of parameters. *Journal of Mathematical Psychology*, *4*(3), 430-472.
- *Grote, M. G. (1995). The effect of massed versus spaced practice on retention and problem-solving in high school physics. *The Ohio Journal of Science*, *95*, 243-247.
- *Gurung, R. A., & Burns, K. (2019). Putting evidence-based claims to the test: A multi-site classroom study of retrieval practice and spaced practice. *Applied Cognitive Psychology*, *33*, 732-743, 5.
- *Guzman-Munoz, F. J. (2017). The advantage of mixing examples in inductive learning: a comparison of three hypotheses. *Educational Psychology*, *37*, 421-437, 4.
- *Hall, K. G., Domingues, D. A., & Cavazos, R. (1994). Contextual interference effects with skilled baseball players. *Perceptual and Motor Skills*, *78*, 835-841, 3.
- *Hausman, H., & Kornell, N. (2014). Mixing topics while studying does not enhance learning. *Journal of Applied Research in Memory and Cognition*, *3*, 153-160, 3.
- Healey, M. K., Hasher, L., & Danilova, E. (2011). The stability of working memory: Do previous tasks influence complex span? *Journal of Experimental Psychology: General*, *140*(4), 573-585.
- Hintzman, D. L. (1974). Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium*. Halsted.
- Hintzman, D. L. (1976). Repetition and memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 10). Academic Press.
- *Jackson, C. E., Maruff, P. T., & Snyder, P. J. (2013). Massed versus spaced visuospatial memory in cognitively healthy young and older adults. *Alzheimer's & Dementia*, *9*, S32-S38, 1S.
- *Kang, S. H., & Pashler, H. (2012). Learning painting styles: Spacing is advantageous when it promotes discriminative contrast. *Applied Cognitive Psychology*, *26*, 97-103, 1.
- *Kang SH, Lindsey RV, Mozer MC, Pashler H (2014). Retrieval practice over the long term: should spacing be expanding or equal-interval? *Psychonomic Bulletin Review*, *21*, 1544-1550, 6.
- *Kapler, I. V., Weston, T., & Wiseheart, M. (2015). Spacing in a simulated undergraduate classroom: Long-term benefits for factual and higher-level learning. *Learning and Instruction*, *36*, 38-45.
- *Karpicke, J. D., & Roediger III, H. L. (2007). Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 704-719.
- *Keller, G. J., Li, Y., Weiss, L. W., & Relyea, G. E. (2006). Contextual interference effect on acquisition and retention of pistol-shooting skills. *Perceptual and Motor Skills*, *103*, 241-252, 1.
- *Kerfoot, B. P., Baker, H. E., Koch, M. O., Connelly, D., Joseph, D. B., & Ritchey, M. L. (2007). Randomized, controlled trial of spaced education to urology residents in the United States and Canada. *The Journal of Urology*, *177*, 1481-1487, 4.

- *Kerfoot, B. P., Fu, Y., Baker, H., Connelly, D., Ritchey, M. L., & Genega, E. M. (2010). Online spaced education generates transfer and improves long-term retention of diagnostic skills: a randomized controlled trial. *Journal of the American College of Surgeons*, *211*, 331-337, 3.
- *Kim, A. S., Saberi, F. M., Wiseheart, M., & Rosenbaum, R. S. (2018). Ameliorating episodic memory deficits in a young adult with developmental (congenital) amnesia. *Journal of the International Neuropsychological Society*, *24*, 1003-1012, 9.
- *Kornell, N. (2009). Optimising learning using flashcards: Spacing is more effective than cramming. *Applied Cognitive Psychology*, *23*, 1297-1317, 9.
- *Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the “enemy of induction”? *Psychological Science*, *19*, 585-592, 6.
- *Kornell, N., Castel, A. D., Eich, T. S., & Bjork, R. A. (2010). Spacing as the friend of both memory and induction in young and older adults. *Psychology and Aging*, *25*, 498-503, 2.
- *Kommeier, J., Spitzer, M., & Susic-Vasic, Z. (2014). Very similar spacing-effect patterns in very different learning/practice domains. *PLoS One*, *9*, 1-11.
- Koval, N. G. (2019). Testing the deficient processing account of the spacing effect in second language vocabulary learning: Evidence from eye tracking. *Applied Psycholinguistics*, *40*(05), 1103-1139.
- *Küpper-Tetzel, C. E., Kapler, I. V., & Wiseheart, M. (2014). Contracting, equal, and expanding learning schedules: The optimal distribution of learning sessions depends on retention interval. *Memory & Cognition*, *42*, 729-741, 5.
- *Kwon, Y. H., Kwon, J. W., & Lee, M. H. (2015). Effectiveness of motor sequential learning according to practice schedules in healthy adults; distributed practice versus massed practice. *Journal of Physical Therapy Science*, *27*, 769-772, 3.
- Landauer, T. K. (1969). Reinforcement as consolidation. *Psychological Review*, *76*(1), 82-96.
- Landauer, T. K. (1974). Consolidation in human memory: Retrograde amnesia effects of confusable items in paired-associate learning. *Journal of Verbal Learning and Verbal Behavior*, *13*(1), 45-53.
- *Landin, D. K., Hebert, E. P., & Fairweather, M. (1993). The effects of variable practice on the performance of a basketball skill. *Research Quarterly for Exercise and Sport*, *64*, 232-237, 2.
- *Logan, J. M., Castel, A. D., Haber, S., & Viehman, E. J. (2012). Metacognition and the spacing effect: the role of repetition, feedback, and instruction on judgments of learning for massed and spaced rehearsal. *Metacognition and Learning*, *7*, 175-195, 3.
- *Lotfolahi, A. R., & Salehi, H. (2016). Learners' perceptions of the effectiveness of spaced learning schedule in L2 vocabulary learning. *SAGE Open*, *6*, 1-9.
- *Lotfolahi, A. R., & Salehi, H. (2017). Spacing effects in vocabulary learning: Young EFL learners in focus. *Cogent Education*, *4*, 1-10.
- *Maass, J. K., Pavlik, P. I., & Hua, H. (2015, June). How spacing and variable retrieval practice affect the learning of statistics concepts. In *International Conference on Artificial Intelligence in Education* (pp. 247-256). Springer.
- Maddox, G. B. (2016). Understanding the underlying mechanism of the spacing effect in verbal learning: A case for encoding variability and study-phase retrieval. *Journal of Cognitive Psychology*, *28*(6), 684-706.
- *Maddox, G. B., & Balota, D. A. (2015). Retrieval practice and spacing effects in young and older adults: An examination of the benefits of desirable difficulty. *Memory & Cognition*, *43*, 760-774, 5.
- *Mammarella, N., Russo, R., & Avons, S. E. (2002). Spacing effects in cued-memory tasks for unfamiliar faces and nonwords. *Memory & Cognition*, *30*, 1238-1251, 8.
- Martin, E. (1972). Stimulus encoding in learning and transfer. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Winston.
- McFarland Jr., C. E., Rhodes, D. D., & Frey, T. J. (1979). Semantic feature variability and the spacing effect. *Journal of Verbal Learning and Verbal Behavior*, *18*(2), 163-172.
- *Metcalfe, J., & Xu, J. (2016). People mind wander more during massed than spaced inductive learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 978-984.
- *Mettler, E., Massey, C. M., & Kellman, P. J. (2016). A comparison of adaptive and fixed schedules of practice. *Journal of Experimental Psychology: General*, *145*, 897-917, 7.
- *Middleton, E. L., Schwartz, M. F., Rawson, K. A., Traut, H., & Verkuilen, J. (2016). Towards a theory of learning for naming rehabilitation: Retrieval practice and spacing effects. *Journal of Speech, Language, and Hearing Research*, *59*, 1111-1122, 5.
- *Middleton, E. L., Rawson, K. A., & Verkuilen, J. (2019). Retrieval practice and spacing effects in multi-session treatment of naming impairment in aphasia. *Cortex*, *119*, 386-400.
- *Miles, S. W. (2014). Spaced vs. massed distribution instruction for L2 grammar learning. *System*, *42*, 412-428.
- *Ostrow, K., Heffernan, N., Heffernan, C., & Peterson, Z. (2015, June). Blocking vs. interleaving: Examining single-session effects within middle school math homework. In *International Conference on Artificial Intelligence in Education* (pp. 338-347). Springer.

- *Pan, S. C., Tajran, J., Lovelett, J., Osuna, J., & Rickard, T. C. (2019). Does interleaved practice enhance foreign language learning? The effects of training schedule on Spanish verb conjugation skills. *Journal of Educational Psychology, 111*, 1172–1188, 7.
- *Perruchet, P. (1989). The effect of spaced practice on explicit and implicit memory. *British Journal of Psychology, 80*, 113–130, 1.
- *Pyc, M. A., Balota, D. A., McDermott, K. B., Tully, T., & Roediger, H. L. (2014). Between-list lag effects in recall depend on retention interval. *Memory & Cognition, 42*, 965–977, 6.
- *Rawson, K. A., & Dunlosky, J. (2013). Relearning attenuates the benefits and costs of spacing. *Journal of Experimental Psychology: General, 142*, 1113–1129, 4.
- Rea, C. P., & Modigliani, V. (1985). The effect of expanded versus massed practice on the retention of multiplication facts and spelling lists. *Human Learning: Journal of Practical Research & Applications.*
- *Rohrer, D., & Taylor, K. (2007). The shuffling of mathematics problems improves learning. *Instructional Science, 35*, 481–498, 6.
- *Rohrer, D., Dedrick, R. F., & Burgess, K. (2014). The benefit of interleaved mathematics practice is not limited to superficially similar kinds of problems. *Psychonomic Bulletin & Review, 21*, 1323–1330.
- *Rohrer, D., Dedrick, R. F., & Stershic, S. (2015). Interleaved practice improves mathematics learning. *Journal of Educational Psychology, 107*, 900–908, 3.
- *Rozenshtein, A., Pearson, G. D., Yan, S. X., Liu, A. Z., & Toy, D. (2016). Effect of massed versus interleaved teaching method on performance of students in radiology. *Journal of the American College of Radiology, 13*, 979–984, 8.
- Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology, 89*(1), 63–77.
- *Russo, R., Parkin, A. J., Taylor, S. R., & Wilks, J. (1998). Revising current two-process accounts of spacing effects in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 161–172.
- *Samudra, P. G., Wong, K. M., & Neuman, S. B. (2019). Promoting low-income preschoolers' vocabulary learning from educational media: Does repetition support memory for learned word knowledge? *Journal of Cognitive Education and Psychology, 18*, 160–173, 2.
- *Sana, F., Yan, V. X., & Kim, J. A. (2017). Study sequence matters for the inductive learning of cognitive concepts. *Journal of Educational Psychology, 109*, 84–98, 1.
- Sana, F., Yan, V. X., Kim, J. A., Bjork, E. L., & Bjork, R. A. (2018). Does working memory capacity moderate the interleaving benefit? *Journal of Applied Research in Memory and Cognition, 7*(3), 361–369.
- Schmeichel, B. J. (2007). Attention control, memory updating, and emotion regulation temporarily reduce the capacity for executive control. *Journal of Experimental Psychology: General, 136*(2), 241–255.
- *Schutte, G. M., Duhon, G. J., Solomon, B. G., Poncy, B. C., Moore, K., & Story, B. (2015). A comparative analysis of massed vs. distributed practice on basic math fact fluency growth rates. *Journal of School Psychology, 53*, 149–159, 2.
- *Shaughnessy, J. J. (1976). Persistence of the spacing effect in free recall under varying incidental learning conditions. *Memory & Cognition, 4*, 369–377, 4.
- *Shaughnessy, J. J., Zimmerman, J., & Underwood, B. J. (1974). The spacing effect in the learning of word pairs. *Memory & Cognition, 2*, 742–748, 4.
- *Shea, C. H., & Kohl, R. M. (1991). Composition of practice: Influence on the retention of motor skills. *Research Quarterly for Exercise and Sport, 62*, 187–195, 2.
- *Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 179–187.
- *Simone, P. M., Bell, M. C., & Cepeda, N. J. (2013). Diminished but not forgotten: Effects of aging on magnitude of spacing effect benefits. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 68*, 674–680, 5.
- *Slone, L. K., & Sandhofer, C. M. (2017). Consider the category: The effect of spacing depends on individual learning histories. *Journal of Experimental Child Psychology, 159*, 34–49.
- *Smith, T. A., & Kimball, D. R. (2010). Learning from feedback: Spacing and the delay–retention effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 80–95.
- *Sobel, H. S., Cepeda, N. J., & Kapler, I. V. (2011). Spacing effects in real-world classroom vocabulary learning. *Applied Cognitive Psychology, 25*, 763–767, 5.
- *Son, L. K. (2004). Spacing one's study: Evidence for a metacognitive control strategy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 601–604.
- *Son, L. K. (2010). Metacognitive control and the spacing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 255–262.
- *Spruit, E. N., Band, G. P., & Hamming, J. F. (2015). Increasing efficiency of surgical training: Effects of spacing practice on skill acquisition and retention in laparoscopy training. *Surgical Endoscopy, 29*, 2235–2243, 8.

- *Stafford, T., & Haasnoot, E. (2017). Testing sleep consolidation in skill learning: A field study using an online game. *Topics in Cognitive Science*, 9, 485-496, 2.
- *Swehla, S. E., Burns, M. K., Zaslofsky, A. F., Hall, M. S., Varma, S., & Volpe, R. J. (2016). Examining the use of spacing effect to increase the efficiency of incremental rehearsal. *Psychology in the Schools*, 53, 404-415, 4.
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous and germane cognitive load. *Educational Psychology Review*, 22(2), 123-138.
- Sweller, J., & Sweller, S. (2006). Natural information processing systems. *Evolutionary Psychology*, 4, 434-458.
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.
- Sweller, J., Van Merriënboer, J., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, 31(2), 261-292.
- *Taylor, K., & Rohrer, D. (2010). The effects of interleaved practice. *Applied Cognitive Psychology*, 24, 837-848, 6.
- *Toppino, T. C., & Bloom, L. C. (2002). The spacing effect, free recall, and two-process theory: A closer look. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 437-444.
- *Toppino, T. C., & Cohen, M. S. (2010). Metacognitive control and spaced practice: Clarifying what people do and why. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 1480-1491.
- *Toppino, T. C., & DiGeorge, W. (1984). The spacing effect in free recall emerges with development. *Memory & Cognition*, 12, 118-122, 2.
- *Toppino, T. C., Cohen, M. S., Davis, M. L., & Moors, A. C. (2009a). Metacognitive control over the distribution of practice: When is spacing preferred? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 1352-1358.
- *Toppino, T. C., Fearow-Kenney, M. D., Kiepert, M. H., & Teremula, A. C. (2009b). The spacing effect in intentional and incidental free recall by children and adults: Limits on the automaticity hypothesis. *Memory & Cognition*, 37, 316-325, 3.
- *Toppino, T. C., Phelan, H. A., & Gerbier, E. (2018). Level of initial training moderates the effects of distributing practice over multiple days with expanding, contracting, and uniform schedules: Evidence for study-phase retrieval. *Memory & Cognition*, 46, 969-978, 6.
- Tyler, J. M., & Burns, K. C. (2008). After depletion: The replenishment of the self's regulatory resources. *Self and Identity*, 7(3), 305-321.
- Underwood, B. J. (1969). Some correlates of item repetitions in free recall learning. *Journal of Verbal Learning and Verbal Behavior*, 8(1), 83-94.
- Underwood, B. J. (1970). A breakdown of the total-time law in free recall learning. *Journal of Verbal Learning and Verbal Behavior*, 9(5), 573-580.
- *Verkoeijen, P., & Bouwmeester, S. (2014). Is spacing really the "friend of induction"? *Frontiers in Psychology*, 5, 259.
- Vlach, H. A., & Kalish, C. W. (2014). Temporal dynamics of categorization: forgetting as the basis of abstraction and generalization. *Frontiers in Psychology*, 5, 1021.
- *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, 4.
- *Vlach, H. A., Sandhofer, C. M., & Kornell, N. (2008). The spacing effect in children's memory and category induction. *Cognition*, 109, 163-167, 1.
- *Vlach, H. A., Ankowski, A. A., & Sandhofer, C. M. (2012). At the same time or apart in time? The role of presentation timing and retrieval dynamics in generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 246-254.
- *Vlach, H. A., Sandhofer, C. M., & Bjork, R. A. (2014). Equal spacing and expanding schedules in children's categorization and generalization. *Journal of Experimental Child Psychology*, 123, 129-137.
- *Vlach, H. A., Bredemann, C. A., & Kraft, C. (2019). To mass or space? Young children do not possess adults' incorrect biases about spaced learning. *Journal of Experimental Child Psychology*, 183, 115-133.
- Wahlheim, C. N., Dunlosky, J., & Jacoby, L. L. (2011a). Spacing enhances the learning of natural concepts: An investigation of mechanisms, metacognition, and aging. *Memory & Cognition*, 39(5), 750-763.
- *Wahlheim, C. N., Dunlosky, J., & Jacoby, L. L. (2011b). Spacing enhances the learning of natural concepts: An investigation of mechanisms, metacognition, and aging. *Memory & Cognition*, 39, 750-763, 5.
- *Wang, Y., Xu, Q., Liao, S., & Jia, D. (2017). Unfamiliar faces in recognition memory: spaced learning enhances subsequent recognition memory by reducing repetition priming. *NeuroReport*, 28, 872-878, 14.
- *Wang, J., Liu, Z., Xing, Q., & Seger, C. A. (2020a). The benefit of interleaved presentation in category learning is independent of working memory. *Memory*, online advanced publish.
- *Wang, N. Y. H., Morris, J., & Howard, D. (2020b). Associative learning in people with aphasia: Exploring spacing of practice as a potential facilitator. *Aphasiology*, 34, 557-579, 5.

- *Whitten II, W. B., & Bjork, R. A. (1977). Learning from tests: Effects of spacing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 465-478, 4.
- Wiseheart, M., D'Souza, A. A., & Chae, J. (2017). Lack of spacing effects during piano learning. *PLoS One*, *12*, 1–13.
- *Xue, G., Mei, L., Chen, C., Lu, Z. L., Poldrack, R., & Dong, Q. (2011). Spaced learning enhances subsequent recognition memory by reducing neural repetition suppression. *Journal of Cognitive Neuroscience*, *23*, 1624-1633, 7.
- *Yan, V. X., & Sana, F. (2020). Does the interleaving effect extend to unrelated concepts? Learners' beliefs versus empirical evidence. *Journal of Educational Psychology*. Online Advanced Publish
- *Young, A. P., Healy, A. F., Jones, M., & Bourne, L. E. (2019). Verbal and spatial acquisition as a function of distributed practice and code-specific interference. *Memory & cognition*, *47*, 779-791, 4.
- *Zigterman, J. R., Simone, P. M., & Bell, M. C. (2015). Within-session spacing improves delayed recall in children. *Memory*, *23*, 625-632, 4.
- *Zulkipli, N. (2013). Effect of interleaving exemplars presented as auditory text on long-term retention in inductive learning. *Procedia Social and Behavioral Sciences*, *97*, 238-245.
- *Zulkipli, N. (2015). The Role of Bottom-Up vs. Top-Down Learning on the Interleaving Effect in Category Induction. *Pertanika Journal of Social Sciences & Humanities*, *23*, 933-944.
- *Zulkipli, N., & Burt, J. S. (2013). The exemplar interleaving effect in inductive learning: Moderation by the difficulty of category discriminations. *Memory & Cognition*, *41*, 16-27, 1.
- *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, 3.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Ouhao Chen¹ · Fred Paas^{2,3} · John Sweller⁴

¹ Mathematics Education Centre, Loughborough University, Loughborough, UK

² Department of Psychology, Education and Child Studies, Erasmus University Rotterdam, Rotterdam, the Netherlands

³ School of Education/Early Start, University of Wollongong, Wollongong, Australia

⁴ School of Education, University of New South Wales, Sydney, Australia