

PART IV

General Learning Strategies

---

PROOF

PROOF

# 16 When Does Interleaving Practice Improve Learning?

Paulo F. Carvalho and Robert L. Goldstone

## Introduction

As you flip through the pages of this handbook you will notice that the content does not seem to be randomly organized. The content of the handbook is *sequenced* in a particular way: foundations before general strategies, background before applications, and so on. The editors envisaged a sequence of topics, the authors of each topic envisaged a sequence of information in each chapter, and so forth. We selected a particular sequence because we considered it to be effective. Deciding how to sequence information takes place all the time in educational contexts, from educators deciding how to organize their syllabus to educational technology designers deciding how to organize a piece of educational software, from handbook editors and writers deciding how to organize their materials, to students making decisions as to how to organize their study. One might imagine that as long as all students study the same materials, regardless of the sequence in which they study it, they will all learn the same information. This could not be further from the truth. In this chapter, we review evidence of how and why the sequence of study changes what is learned. In doing so, we uncover the powerful ways in which sequence can improve or deter learning.

## How Does the Sequence of Study Affect Learning and Why Does It Matter?

Cognitive psychologists have long identified the powerful effect that the sequence in which information is presented has on cognition and, specifically, on learning and memory. The main research question is, given the same amount and nature of training, how is learning affected by changing the sequence in which the information is presented? This question has a long history in cognitive science because it has the potential to inform theories of learning. If learning is the process of connecting and integrating new information into existing structures, then which information is learned first should have an impact on how later information is integrated and ultimately what is learned. Thus, studying sequencing effects can help us understand the learning *process*. Of course, these findings can also directly inform practice by helping educational designers, teachers, and students know which factors to consider when deciding how to sequence learning materials.

For example, Elio and Anderson (1984) proposed that learning should start with low-variability items (e.g., items that do not differ much from one another or from the central tendency of the category) and items with greater variability should be introduced later (for similar evidence with young learners, see Sandhofer & Dumas, 2008). When learning about mammals, one should probably start by studying deer, gazelles, moose, and impalas, and only after having covered these more similar types of mammals move to bats and whales. However, not all learners benefit from this approach. Elio and Anderson (1984) also showed that if the learners' approach to the task is to consciously generate hypotheses of category membership, the pattern of results is reversed. One possibility is that how the learner approaches the task changes what type of information gets encoded and, consequently, what information is more relevant (Elio & Anderson, 1984); when learners are asked to generate explicit hypotheses, starting with low-variability items could lead them to be biased toward incorrect or partial classification rules (in the example above, that all mammals have four legs, for example). In these conditions, starting with high-variability items helps learners to not settle on incorrect hypotheses, whereas when learners are not trying to create explicit rules, the initially reduced variability leads students to abstract common properties that are likely to characterize most, if not all, items of the category (Elio & Anderson, 1984).

Consistent with these results, it has also been suggested that for optimal transfer of category learning, the study situation should emphasize items that promote a coherent generalization based on the properties that occur more frequently (Elio & Anderson, 1981). Moreover, in situations where one needs to learn several items that promote different types of generalizations, best learning is achieved by studying items close in time that promote similar generalizations (Elio & Anderson, 1981; Mathy & Feldman, 2009).

Students often struggle to decide whether to get difficult material out of the way by studying it early or studying it last. In general, research seems to indicate that learning benefits from study with examples organized in increasing order of complexity or difficulty, i.e., from the easiest and simplest to the hardest and more complex (Hull, 1920; Terrace, 1964). But the reverse pattern has also been shown. For example, Lee and colleagues (1988) showed that learners who start by studying examples other learners classified incorrectly made fewer errors during later classification tests than learners who studied the examples in the opposite order. One possibility is that starting with difficult items might be beneficial only for concepts organized around a clear rule, whereas the reverse is true for concepts requiring the integration across different dimensions (Spiering & Ashby, 2008). Why? When learners study easy items, for which identifying the classification rule is easy, closer in time, they will identify the rule faster and with ease. This is good for learning. But when there is no such rule, studying together items that are easy to categorize together might lead learners to assume a rule that is not correct or only partially correct (Spiering & Ashby, 2008).

Three important aspects are salient from the above brief review of sequencing effects in learning: (1) learning takes place by integrating previous experience with current experience and is sensitive to which sequence is used; (2) no one way of

sequencing is optimal for everyone, every time, and with every material; and (3) the optimal sequence depends on a match between the learning process and the properties of the sequence. These conclusions will set the stage for our overview of one particular sequencing effect: the learning impact of interleaved practice.

### The Case of Blocked and Interleaved Practice

School involves learning different topics across and within several disciplines. For example, a math course might involve learning about area, perimeter, volume, and so on. The above discussion implicitly envisions sequences of study materials where one topic is introduced and finished before the next one, but it does not have to be that way. When different topics are studied separately, we have *blocked practice* – a textbook might start with a section on calculating area, practice problems on area, and only then move on to perimeter, and so on. This sequence can perhaps consider difficulty and prerequisites, as discussed above, in how to decide how to order these blocks, that is, whether area or perimeter should happen first.<sup>1</sup> By most accounts, albeit a general and quantifiable metric of content similarity has not yet been devised, most learning and practice tools (classes, textbooks, tutoring systems, etc.), are organized this way (Bjork, Dunlosky, & Kornell, 2013).

However, another possibility is to alternate practice of different topics, i.e., *interleaved practice*. This could mean, for example, introducing area, perimeter and volume in the same session and practicing them in alternating or random order. Although arguably not as common in educational practice, interleaved practice has received recent support from educational researchers, suggesting that it might yield substantial benefits over the traditional blocked practice.

Our focus in the remainder of this chapter will be in understanding when and how interleaved practice can improve learning but also when it does not and when blocked practice should be used instead. We will start with a brief overview of empirical evidence in favor of interleaved practice. We will then cover evidence of negative or neutral effects of interleaved practice and conclude with a theoretical framework to help understand – and predict – when interleaved practice should be used.

### Using Interleaved Practice to Boost Learning

Interleaved practice has been demonstrated to improve learning compared with blocked practice in many different situations, from motor behavior to learning psychology concepts in the classroom. When learners interleave practice of different topics, concepts, or actions, learners continuously need to monitor and change their

<sup>1</sup> Though Lee and colleagues' (1988) proposal of starting with harder materials before easy materials can yield learning benefits, in practice, learning must be sequenced in such a way so that all prerequisites to current learning have been previously introduced and mastered. For example, although calculus is harder than addition, one must start with addition because calculus would be impossible without knowing addition.

response with different contexts, and simultaneously maintaining several responses in working memory results in better retention at delayed tests. This process creates beneficial contextual interference (Lee & Magill, 1985; Paas, 1992; Shea & Kohl, 1991). Early evidence for this proposal comes from studies in motor learning. For example, Ste-Marie and colleagues (2004) trained elementary school children's handwriting skills for three letters (*h*, *a*, and *y*) either in a blocked or random (interleaved) sequence. Children who practice the letters interleaved (randomly presented) retained more of their training in a delayed test of letter writing skills than those who practice the letters blocked. Follow-up studies showed that children who practiced the letters interleaved were also faster at writing a new word (*hay*) in delayed tests. Similar results have been demonstrated with other motor skills (Shea & Morgan, 1979).

However, the increase in contextual interference associated with interleaved practice has an important consequence: students' performance during learning is markedly worse in interleaved than in blocked practice (Shea & Morgan, 1979; Ste-Marie et al., 2004). Because of this mismatch between study and test performance, interleaved practice has been classified as a "desirable difficulty" – a study practice that, counterintuitively, reduces performance during learning but improves long-term learning (Bjork, 1994).

Rohrer and colleagues have proposed a similar framework (Rohrer, Dedrick, & Burgess, 2014) for the demonstrated benefits of interleaved practice in the domain of mathematics learning (Rohrer, 2012, 2015; Rohrer & Taylor, 2007). The authors propose that when practice of different types of mathematics problems are interleaved, students attend more to which procedure needs to be applied for each type of problem; they not only learn the procedure better but also when to apply it (see also Li, Cohen, & Koedinger, 2013). Consistent with this proposal, research has demonstrated that students not only learn better (as measured in delayed test) following interleaved practice but also make fewer confusion errors, i.e., they are less likely to apply the wrong procedure to a problem (Rohrer & Taylor, 2007). In addition, Wahlheim, Dunlosky, and Jacoby (2011) showed that blocked study results in decreased encoding of immediate repetitions of the same category. The authors analyzed memory performance at test for items that had been studied as a function of which position in the study sequence the item had been studied (i.e., was it the first, second, third, and so on to be studied?). The results showed a decreasing function for blocked practice, with learners more accurately classifying earlier items into the correct category than later ones. For interleaved practice, however, there was no difference in categorization performance across study positions. One interpretation of the results is that, in later presentations in the blocked sequence, learners did not attend to the items to the same extent as the earlier ones and therefore do not recall them as well.

Interleaved practice has also been shown to benefit concept learning and generalization. Kornell and colleagues (Kornell & Bjork, 2008; Kornell et al., 2010) have demonstrated in laboratory studies that learners remember the style of new artists better following interleaved practice (for examples of the artists, see Figure 16.1).

## Pessani Paintings



## Stratulat Paintings



## Braque Paintings



## Cross Paintings



**Figure 16.1** Examples of paintings and artists used by Kornell and colleagues (Kornell & Bjork, 2008; Kornell et al., 2010)

More than that, they are better at identifying new paintings from those artists and telling apart studied from not studied artists.

Expanding on this work, Kang and Pashler (2012) replicated Kornell and colleagues' results and directly tested the possibility that the benefits of interleaved practice might, at least in part, be due to the increased spacing between repetitions of the same category. The authors contrasted learners' test performance following a spaced practice sequence (in which repetitions of each category were spaced in time but not interleaved – like blocked study with added temporal spacing between repetitions) with a blocked practice condition and an interleaved condition. Learners' classification of new paintings at test was best following interleaved practice. Moreover, blocked and spaced practice resulted in equivalent test performances. These results suggest that the benefits of interleaved practice are tied not to the increased temporal lag between practices of the same type of material or response



but to the temporal contiguity among items of different concepts or that require different responses.

Birnbaum and colleagues (2013) used pictures of butterflies from different species to test people's acquisition of natural animal categories (species) following interleaved and blocked practice. They found that interleaved practice results in best learning of butterfly species and improved generalization to new exemplars of each species. This study also replicated Kang and Pashler's (2012) finding that temporal spacing between presentation of different concepts during interleaved practice reduces its benefits. Interestingly, the opposite was seen for blocked practice: increasing the temporal spacing between repetitions of the same category improved learning (albeit not above the levels seen following interleaved practice; for similar results using a discrimination task with artificial stimuli, see Mitchell, Nash, & Hall, 2008).

Interleaved practice has also been shown to improve learning of verbal concepts, specifically in educational contexts. For example, Introductory Psychology students were better at identifying new situations exemplifying psychology concepts such as "foot-in-the-door technique" and "hindsight bias" when they had practiced by studying examples of these concepts interleaved (Rawson, Thomas, & Jacoby, 2015). Similarly, college students were better at identifying psychological disorders when they practiced each disorder by studying representative cases interleaved, regardless of whether these were presented visually or aurally (Zulkipli et al., 2012), and medical students were better at interpreting exams following interleaved practice with x-ray images (Rozenstein et al., 2016) and EEG results (Hatala, Brooks, & Norman, 2003). Studies have also found benefits of interleaved practice for neuroanatomy students learning to classify new structures (Pani, Chariker, & Naaz, 2012) and college students learning to classify mathematical functions (McDaniel, Fadler, & Pashler, 2013), to mention a few.

Generally, there is an assumption that interleaved practice will result in greater longer-term benefits than what would be possible with blocked practice. For instance, Rohrer, Dedrick, and Stershic (2015) created an intervention for middle school math class that included blocked or interleaved study of four types of problems followed by a review session of all the problems studied. Either one or thirty days after the end of the review session students completed a test. The results showed an overall benefit for interleaved study. Interestingly, the numerical benefit of interleaved practice over blocked practice increased with increasing retention intervals. Similar results have been found using an online tutor system to deliver practice math activities (Ostrow et al., 2015).

It is possible, however, that the benefit of retention interval seen here is orthogonal to the relative benefits of different sequences. It is possible, for example, that greater delays between the end of study and test promote better performance for all sequences of study, the difference being that one characteristic of interleaved study is that it includes a more even distribution of the problems across the entire learning sequence. Rohrer and colleagues (2014) report partial evidence for this possibility in an experiment comparing interleaved and blocked practice of mathematical problems in a naturalistic setting that did not include a review session before test.



The results show that the benefit of interleaved over blocked practice is smaller for materials studied in earlier blocks, that is, for materials for which the period between last study and test was the longest for blocked practice and increases monotonically with decreasing retention intervals between the end of blocked practice and test. These results indicate that although retention interval may play a role, it is not limited to interleaved practice.

Overall, the evidence of the benefits of interleaved practice is encouraging and has led to the development of several intervention studies and learning systems that incorporate interleaved practice as one of their key aspects. These interventions have shown improved learning and instructional outcomes in domains as diverse as geology (Andrews & Frey, 2015), biochemistry (Horn & Hernick, 2015), physiology (Linderholm, Dobson, & Yarbrough, 2016), neuroanatomy (Chariker, Naaz, & Pani, 2011; Pani et al., 2012), and plant biology (Kirchoff et al., 2014). However, we should note, there has, to date, not been a large-scale classroom randomized controlled trial of the benefits of interleaved practice as a learning and teaching tool, and the existence of null results in applied studies (e.g., Dobson, 2011) still warrants some caution.

To summarize this section, interleaved practice shows great promise as an easy-to-implement way to improve learning. It has been shown to positively change learning in a wide range of domains and populations, including in classroom studies. However, it is important to keep in mind that, as we noted before, at least in what concerns sequencing effects in learning, no single sequence seems to benefit all types of situations. Moreover, students seem to prefer blocked study when given the chance (Carvalho et al., 2016; Kornell & Bjork, 2008; Tauber et al., 2013; Yan et al., 2017) and this preference seems particularly resistant to change (Yan, Bjork, & Bjork, 2016). Thus, it is important to look not only at the uses of interleaved practice but also at the positive uses of blocked practice.

### **Using Blocked Practice to Boost Learning**

Despite what the previous section might make one think, blocked practice is not always a poor choice. There is considerable evidence that blocked practice can result in improved or equally effective learning as interleaved practice, and also across a varied set of domains, from abstract stimuli and motor behavior to psychology concepts. For example, Kurtz and Hovland (1956) demonstrated, using abstract stimuli, that people remembered the properties of studied items better and were better at describing the main characteristics of each of four categories, when study had been blocked by category. Similarly, Whitman and Garner (1963) had adult participants learn two categories organized by the relational structure of geometrical objects in a figure. The results showed that participants achieved criterion more quickly when stimuli from the same category (i.e., that shared the same relational structure) were studied blocked. More recently, benefits of blocked practice have been demonstrated with several different types of artificial stimuli (Carvalho & Goldstone, 2014b, 2017b; Zulkiply & Burt, 2013). One important idea put forward by this research is that concept learning does not have to always take place by

contrasting a to-be-learned concept with other concepts. It is possible – perhaps sometimes preferable – to learn concepts in isolation as with blocked practice. Learning the concepts in isolation allows one to focus on their positive characterizations and not on how positive examples contrast with negative examples (Goldstone, 1996).

Evidence of positive effects of blocked practice also seems to question the proposal that contextual interference benefits learning, as discussed in the previous section. In fact, research in motor behavior has also shown that blocked practice can sometimes improve learning compared with interleaved practice. Hebert, Landin, and Solmon (1996) taught undergraduate students enrolled in tennis classes different ground strokes, either blocked or interleaved. They found that low-skilled students (as measured before starting training) learned the ground strokes better with blocked practice and performed better in a delayed posttest compared with interleaved practice. No significant difference between the two practice schedules was found for high-skilled students at posttest. Similar results have been found when comparing young and older learners' acquisition of motor skills in blocked or interleaved practice (Al-Mustafa, 1989; Farrow & Maschette, 1997; Pigott & Shapiro, 1984; Pinto-Zipp & Gentile, 1995). This pattern of results, in addition to other evidence of lack of difference between the two sequences of study (see Wulf & Shea, 2002), seems to indicate that perhaps the benefits of interleaved practice are connected exactly with situations that the contextual interference theory proposes that it should benefit the least from: those in which memory and attentional demands are relatively low, such as with experts, adults, or longer study times. When attentional and memory demands are increased, as might be expected for novices, younger learners, and shorter study times, the pattern is reversed and blocked practice is frequently more beneficial (Wulf & Shea, 2002). However, Sana, Yan, and Kim (2016) recently demonstrated that when learning three different statistics concepts (Kruskal–Wallis, Wilcoxon signed-rank, and Chi-squared tests), learners with high working memory capacity benefited from blocked practice to the same extent as interleaved practice, whereas learners with lower working memory capacity benefited more from interleaved practice.

Still in the math domain, Rau, Alevén, and Rummel (2013) conducted a classroom study with 5th and 6th grade math students learning fractions using different representations (e.g., pie charts, number lines, and sets) with a math tutor. Their results demonstrate a benefit in immediate and delayed tests of blocked practice by representation compared with interleaving the representations. Additionally, these benefits were particularly pronounced for low-skill students. Thus, blocked practice may not only be more beneficial overall but also especially help those students who are more likely to struggle with the materials being presented. Interleaved study might be too challenging for low-skill students.

Although one could argue that the benefits of blocked practice are tied to specific material or individual differences, that does not seem to be the case. For example, Kost, Carvalho, and Goldstone (2015) used Kornell and Bjork's (2008) artists' styles to show that blocked practice can also improve learning of natural categories that have been shown before to benefit from interleaved practice. Kost and colleagues

(2015) demonstrated that, in a condition where learners had the opportunity to study each painting more than once and were asked to “guess” the category assignment and provided with feedback, learners were better at classifying old and new paintings from artists studied using blocked practice. When study did not include repetition or active guessing from the learners, interleaved practice was again more beneficial, as previous research had indicated.

Blocked practice can also improve verbal learning. For example, Carpenter and Mueller (2013) showed that non-French speakers learned orthographic-to-phonological mappings in French (i.e., “-eau” and the corresponding sound /o/ in the words “bateau,” “carreau,” and “corbeau,” and “-er” and the sound /e/ in the words “adosser,” “attraper,” and “baver”) better when they studied different words with the same mapping blocked (bateau, carreau, corbeau, adosser, attraper, baver, etc.) rather than when words with different mappings were studied interleaved (bateau, adosser, carreau, attraper, corbeau, baver, etc.). Similarly, Sorensen and Woltz (2016) conducted a study investigating the effect of blocked and interleaved practice in learning the association between novel words and groups of English words. For example, the new association between “things found underground” (such as cave, roots, potato, and tunnel) and the novel word “brask.” They found that blocked practice resulted in better performance in implicit (categorization of novel words, e.g., worm, gopher, well, and aquifer, into one of the new verbal categories) and explicit tests (definition of the grouping rule for the new categories). Moreover, Rawson and colleagues (2015) demonstrated that when undergraduate Introductory Psychology students were presented a definition before studying several examples of different psychological concepts such as “foot-in-the-door,” blocked practice resulted in better subsequent classification of novel examples compared with interleaved practice.

It is important to note that the benefits of blocked practice are not limited to immediate testing as one might initially predict based on results from the contextual interference literature and temporal spacing theories. For example, Carvalho and Goldstone showed benefits of blocked practice at different intervals, including from 3 minutes to up to three days, for both memory (Carvalho & Goldstone, 2017b) and categorization (Carvalho & Goldstone, 2014a).

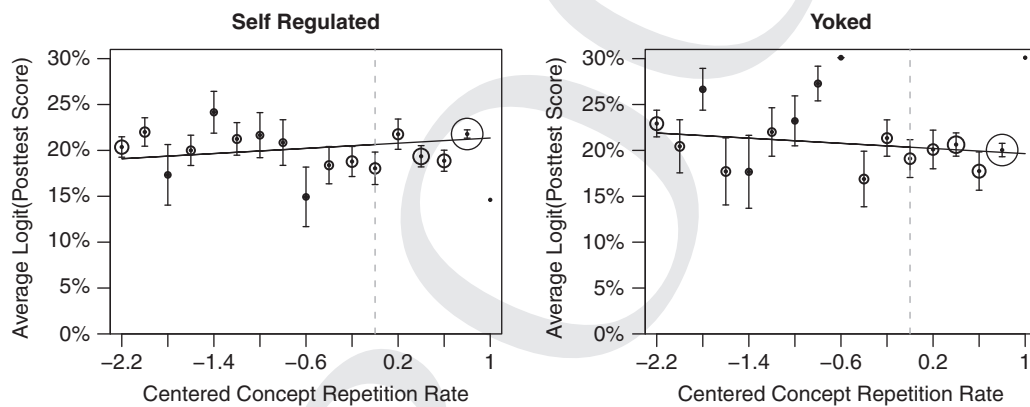
Finally, despite the widely held belief that students might not study using the best sequence, benefits of blocked study have been found in self-regulated learning. If left to organize their study, students prefer to block practice by studying items from the same category close in time (Tauber et al., 2013) and, because blocked practice is often thought to result in worse learning (Bjork et al., 2013), one might assume that students’ choices are not optimal and result in worse learning. Carvalho and colleagues (2016) tested this possibility in a classroom study with Introductory Psychology students learning the concepts of mean, median, and mode – three measures of central tendency in statistics. After completing a short pretest, students were provided with practice materials of these three concepts in an online platform. Students were free to study the materials in any order they wished. The posttest was a set of questions regarding these concepts introduced in the students’ midterm exam. As predicted, students showed a high preference to block their study. More importantly, higher rates of blocked practice were associated with higher posttest scores,

controlling for pretest differences (see Figure 16.2). This relationship was not present in a control group of students who did not choose how to organize their study but were instead yoked to a sequence chosen by another student.

To summarize this section, blocked practice shows great promise as a simple way to improve learning. It has been shown to positively affect learning in a wide range of domains and populations, including in classroom studies (Monteiro et al., 2017). However, as reviewed in the previous section, interleaved practice also has demonstrable benefits. Thus, whether interleaved or blocked practice is more beneficial cannot be answered in terms of main effects but rather in how sequencing interacts with other factors. In the next section, we discuss a theoretical framework that can help make this decision.

### How to Decide Whether to Use Interleaved or Blocked Practice?

Two important conclusions can be taken from the previous discussion. On the one hand, there might not be a single sequence that improves all learning, all the time, for



**Figure 16.2** Average posttest score by concept repetition rate for the (a) self-regulated and (b) yoked groups in Carvalho and colleagues (2016; reprinted with permission)

Students were divided into bins by their adjusted rate of repetition and average posttest scores (Logit transformed) within each bin were plotted. Concept repetition rates (Logit transformed) within each bin were plotted. Concept repetition rate was adjusted by subtracting the average rate of repetition for the entire group from the rate of repetition for the bin – a value of 0 in the x-axis indicates mean concept repetition rate (represented by the vertical dashed line) and increasing values indicate increasing difference from average. The values in the y-axis represent Logit transformed posttest scores. Each point in the graph lies at the center of a 20%-wide interval of concept repetition rates, and represents the average posttest score among students whose concept repetition rates fell in that interval. The number of students in each bin is represented by the area of the circles surrounding the data points. The regression lines represent best-fitting lines of the regression analyses assuming average values for all predictors other than concept repetition rate.

all the students. On the other hand, sequence does have the potential to improve learning and it is a relatively easy-to-implement intervention in most educational settings and beyond. Where does this leave us? We believe that the development of a theory that can account for the pattern of results observed, and predict when and why interleaved or blocked practice will be beneficial, will allow for the comprehensive use of sequence of study to make learning better.

We (Carvalho & Goldstone, 2015b, 2017b) have proposed that the effect of different sequences on learning efficacy might not be due to the sequence per se but the effect that sequencing has on a general-purpose learning process. Our account of why the sequence of study changes learning can predict most of the results described above and help guide educators, students, and learning scientists on deciding which sequence to use.

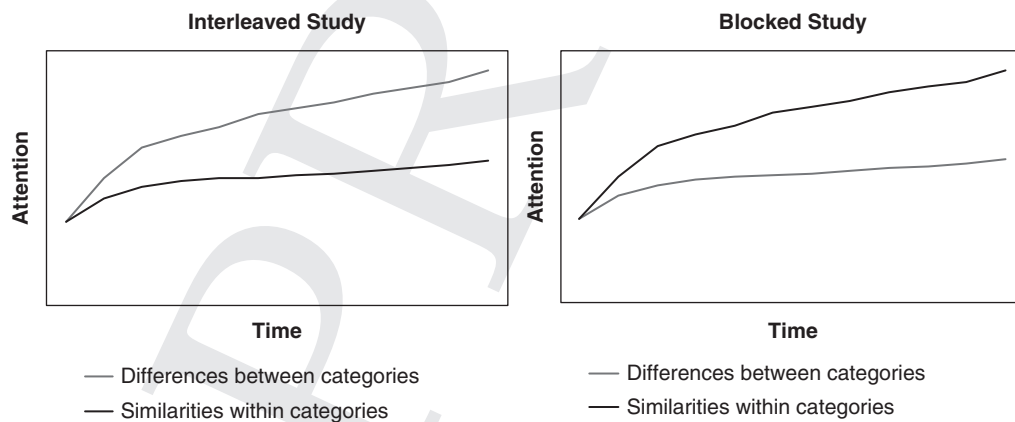
The Sequential Attention Theory (SAT; Carvalho & Goldstone, 2015b, 2017b) posits that interleaved and blocked practice emphasize different aspects of the learning materials or situation and, therefore, the best sequence of study is the one that emphasizes the most challenging aspect(s) of the learning situation. As an example, if learning is hard because it is difficult to distinguish two concepts, then one would want to use a sequence that emphasizes exactly that: interleaved study. Conversely, if learning is hard because it is difficult to identify common aspects among examples of the same concept, then one would want to use a sequence that emphasizes exactly that: blocked study. Overall, this description is consistent with previous proposals (Birnbaum et al., 2013; Carpenter & Mueller, 2013; Carvalho & Goldstone, 2014b, 2015a; Goldstone, 1996; Higgins & Ross, 2011; Higgins, 2017; Mathy et al., 2013; Noh et al., 2016; Rawson et al., 2015; Sandhofer & Dumas, 2008; Zulkiply & Burt, 2013) and seems to match the main findings in the literature we described thus far (see the sections that follow for details). However, this description by itself *does not explain why this is the case*. SAT goes one step further by describing the learning mechanism which gives rise to these differences – a mechanism that can be tested and implemented in educational tools.

SAT (Carvalho & Goldstone, 2015b, 2017b) proposes that learning is a sequential process of item-by-item comparisons to emphasize important stimulus characteristics. During learning, learners compare the current item with the previously studied one and, depending on the assignment of the previous and current items, attend to similarities or differences between the two items. This assumption is consistent with behavioral and modeling results in the category learning literature, showing that recent categorization events play a stronger role in a novel categorization decision than do older events and that categorization decisions are not based on a veridical analysis of the distribution of exemplars across time (e.g., more recent examples are more emphasized; Jones & Sieck, 2003; Stewart & Brown, 2004; Stewart, Brown, & Chater, 2002) but are context- and task-specific (Mack & Palmeri, 2015; Markman & Ross, 2003; Palmeri & Mack, 2015; Ross, 2000). Moreover, this proposal is also congruent with recent neurophysiological evidence suggesting the important role of pattern completion for learning and the role of the hippocampus in providing details not only about past events but also about the relationship between events to create learning (Mack & Preston, 2016; Schlichting & Preston, 2015; Zeithamova, Schlichting, & Preston, 2012).

During each learning moment (e.g., a trial in a laboratory task), the learner evaluates similarities and differences between the current stimulus and the recollection they have of the previous item(s), as well as the category assignment of the previous exemplar and the current one. If the previous and current items belong to the same category, attention will be directed toward their similarities. However, if they belong to different categories, attention will be directed toward their differences. Across time, attention will be increasingly shifted toward relevant within-category similarities and between-category differences. This will, in turn, affect category representations, which will affect categorization decisions and recollection. With each new learning moment, the relevant properties will be progressively better encoded whereas irrelevant ones will be poorly encoded or not encoded at all.

When categories are studied interleaved, the number of transitions between objects of different categories is frequent, which will result in attending to differences between categories on most trials by the process just described. In the same way, when categories are studied blocked, the likelihood of a within-category transition is high, which will increase attention toward within-category similarities by the same process. Furthermore, this process can also lead to encoding information that might not be central for learning the categories. For example, blocked practice would lead learners to encode similarities among items of the same category that will end up also being present in the other category and, therefore, cannot discriminate between the two categories (see Figure 16.3 for a schematic representation of this proposal).

The process of sequential shifts of attention toward similarities or differences among studied items will give rise to overall different attention patterns and this, according to SAT, is the reason why interleaved study leads to better encoding of differences between categories and blocked practice leads to better encoding of similarities among items of the same category. Carvalho and Goldstone (2017b) showed



**Figure 16.3** Schematic representation of the mechanism proposed in Sequential Attention Theory (SAT) for how each sequence leads to attending to different properties of the studied materials (from Carvalho & Goldstone, 2017b. Reprinted with permission from the American Psychological Association)



evidence for this mechanism in a series of laboratory studies. The researchers probed what learners attend to, encoded, and remembered from materials studied interleaved or blocked. If the differences seen between the two practice sequences are due to different attentional patterns during learning, using an eye tracker, we should see that learners look at different aspects of the study materials depending on the sequence of study. That is exactly what Carvalho and Goldstone found. Whereas during interleaved practice learners attended to sequential differences between study items, during blocked practice learners showed no such bias. Moreover, Carvalho and Goldstone probed learners' memory and categorization of new items and found that, following blocked practice, learners were more sensitive to changes in characteristic properties of the concepts that did not discriminate between the categories but were frequent in each category. Conversely, following interleaved practice, learners ignored those properties and were more sensitive to infrequent but discriminative properties (those that helped tell the categories apart).

Finally, SAT makes a series of predictions about factors that should modulate the benefit of interleaved and blocked practice. Next, we will briefly review some of these factors. In the final section, we will propose future directions and other modulating factors that have not yet been tested in the literature.

### **Different Types of Concepts**

The type of concepts being studied will change which aspects are harder to learn and should be attended to. Telling alligators and crocodiles apart requires fine discrimination among similar concepts whereas identifying whether a physics problem requires classical or quantum theoretical constructs requires the identification of common characteristics across a highly varied set of classical and quantum physics problems. Situations such as the former would benefit from interleaved practice, whereas the latter would benefit from blocked practice. Consistent with this proposal, interleaved practice improves learning, for example, highly similar abstract categories (Carvalho & Goldstone, 2014b, 2014a; Zulkipli & Burt, 2013), rule-based categories (Noh et al., 2016), the style of similar artists (Kang & Pashler, 2012; Kornell & Bjork, 2008; Kornell et al., 2010), confusable natural bird categories (Wahlheim et al. 2011), related mathematical and verbal concepts (Rawson et al., 2015; Rohrer et al., 2015; Rohrer & Taylor, 2007; Taylor & Rohrer, 2010), and motor behaviors (Tse & Altarriba, 2010). Conversely, blocked practice improves learning, for example, highly dissimilar abstract categories (Carvalho & Goldstone, 2014b, 2014a; Kurtz & Hovland, 1956; Zulkipli & Burt, 2013), information integration categories that require noticing several similarities among items of the same category (Noh et al., 2016), and identifying a phoneme as present in orthographically diverse words (Carpenter & Mueller, 2013).

### **Different Types of Study Activities**

Different study tasks might lead students to focus on different aspects of the same concept (Markman & Ross, 2003; Yamauchi & Markman, 2000). For example, teachers often use inference activities that present a series of different



scenarios or sentences with a common underlying principle to stimulate students to make inferences during reading. Inference-making is an important skill for reading proficiency (McNamara & Kintsch, 1996) that is part of the common core standards (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010). This activity will tend to benefit from blocked study, as learners will more easily be able to identify and infer the common aspects across multiple scenarios if these are presented close in time, instead of interleaved with other scenarios with different inferences. Conversely, when asked to complete classification activities such as sorting different animals into classes – an important skill in biology education (National Research Council, 2012) – students focus on the properties that distinguish the animals and that can help in dividing into different classes. This type of learning activity will benefit from interleaved practice (e.g., practice classifying animals of different classes together, instead of only one class at a time), as students will more easily be able to identify the differences between the classes than if they study only one item at a time.

Consistent with this prediction, Carvalho and Goldstone (2015a) demonstrated that, when using the same learning materials, interleaved practice promoted learning in a classification task whereas blocked practice promoted learning in an inference task. Similar results with different materials were found by Kost and colleagues (2015). Rawson and colleagues (2015) found a similar interaction by manipulating whether the concept definition was presented along with examples (which arguably makes it more of an inference task) or not (which arguably makes it more of a classification task).

### **Different Types of Testing Activities**

Different testing activities also recruit different knowledge. If the test activity requires knowing the discriminating properties among different options – such as the commonly used multiple-choice test – not having encoded these properties would result in low performance. Conversely, if the test task requires the student to recruit the details about how different examples fit together (i.e., the characteristics properties of a concept, as would arguably be the case with providing definitions for studied concepts), perhaps knowing only the discriminating properties would impair performance. Consistent with this proposal, Carvalho and Goldstone (2017a) demonstrated that learners provided more correct definitions for psychology concepts studied blocked than interleaved, but there were no differences between the two practice sequences for other types of tests (multiple-choice, writing new examples, matching definitions with the correct concept). Similar results were found by Carvalho and Albuquerque (2012) using a memory test. In that study, learners learned to discriminate two images along one dimension using either interleaved or blocked practice. When during a test learners had to tell apart studied items from novel items that differed along the discriminating dimension, interleaved practice improved learning. Conversely, if the memory test included distractors that varied from studied items

along a different dimension, there was no difference between the two practice sequences (see also de Zilva, Mitchell, & Newell, 2013).

### **Different Types of Learners**

SAT describes the attentional process for an average adult. However, different developmental stages, levels of expertise, and working memory capacity change how the attentional process operates, with consequences for which practice sequence is better.

There are documented developmental differences in how children and adults deploy attention. For example, younger children are more likely to distribute attention inconsistently among dimensions whereas older children are able to consistently focus on a single dimension during learning (Cook & Odom, 1992; Smith & Kemler, 1978; Strutt, Anderson, & Well, 1975; Thompson & Markson, 1998). Similarly, in categorization tasks, it has been shown that infants and children do not optimize attention toward task-relevant properties of the materials being studied whereas adults show such optimization (Best, Robinson, & Sloutsky, 2011; Best, Yim, & Sloutsky, 2013), with consequences for what is remembered from learning (Deng & Sloutsky, 2016). Thus, for the young attentional system, interleaved practice might pose special challenges because it is particularly hard to focus on specific dimensions of change among many, and attend only to that. This means that even though interleaved practice directs attention toward relevant differences among successive items, if the attentional system cannot focus on a reduced number of dimensions and is instead attracted by overall variation, then the high amount of variation during interleaved practice (varying stimuli and varying responses) can delay learning. During blocked practice, on the other hand, learners do not need to focus their attention on particular differences but rather attend to all properties of the material being studied. Successive similarities among items of the same category will be more frequent and potentially better encoded, resulting in successful learning.

This interpretation is consistent with the results already discussed, showing that young children benefit more from blocked practice whereas adults benefit more from interleaved practice for the same types of materials (Al-Mustafa, 1989; Farrow & Maschette, 1997; Pigott & Shapiro, 1984; Pinto-Zipp & Gentile, 1995). Moreover, it is consistent with results showing that, for young children, the benefits of interleaved practice increase with decreased perceptual complexity of the materials studied – less complex materials will have fewer features competing for attention, which will help children focus on the relevant feature. That is, when fewer dimensions change from one study moment to the next, children can benefit from interleaved practice, but there is no difference between the two sequences for more perceptually complex materials (Albaret & Thon, 1998).

Similarly, experts are able to focus their attention on specific components of the task and ignore others whereas novices tend to show less organized attention

(Beilock et al., 2002; Kioumourtzoglou et al., 1998; Werner & Thies, 2000). Experts are also able to select and manage which information is relevant from an array whereas novices tend to attend to all information presented (Carter et al., 1988). Thus, novices, those who do not have relevant prior knowledge that might help identify which dimensions might be relevant for learning, might not be able to benefit from interleaved practice because they are not able to organize their attention to focus on only the relevant changes from one learning moment to the next. Conversely, experts, who more easily organize their attention based on previous knowledge, can successfully benefit from interleaved practice and identify relevant differences between successive items (Hebert et al., 1996). Consistent with this proposal, Shea, Kohl, and Indermill (1990) demonstrated that increasing amounts of practice with a new complex task (from 50 to 400 practice opportunities) increased the benefits of interleaved practice. At low levels of practice (50 practice opportunities), blocked practice resulted in better learning.

Finally, for simplicity, in our description of the sequential attention process during learning – which we argued gives rise to differences between interleaved and blocked practice – we assumed that learners consider only the previous item to establish sequential comparisons and decide what to attend to in each learning moment. What would happen if learners could maintain more information in memory? Individuals with higher working memory capacity are able to maintain in memory and simultaneously operate with a large set of information (Engle & Kane, 2003; Kane et al., 2004; Unsworth & Engle, 2007). Moreover, these individuals also are better at controlled search from long-term memory and focused attention (Unsworth & Engle, 2007; Unsworth & Spillers, 2010). One prediction, then, would be that higher working memory capacity would be related with reduced differences between interleaving and blocked practice because learners can use more information in the sequential attentional learning process. The results from the work by Sana and colleagues (2016) as described are consistent with this proposal; for individuals with higher working memory spans (a measure of working memory capacity), the benefits of interleaved practice were less pronounced than for individuals with lower working memory spans (for whom, interleaved practice was more beneficial). The less information one can maintain and operate in working memory simultaneously, the more beneficial temporal contiguity of contrasting cases is, because it will result in heightened attention toward the relevant properties which would otherwise not be possible.

### **Conclusion: What Now?**

We started this chapter by arguing that sequencing decisions happen all the time in educational settings and might seem inconsequential. The evidence presented here paints a picture of an important influence of how study practice is organized, be it interleaved, blocked, hard-to-easy, easy-to-hard, and so on. It is also important to emphasize that it is quite apparent from the available evidence that there is no one-

size-fits-all solution in terms of best sequence of study. Thus, future research should focus on understanding the underlying learning process as an online process that takes place over time. As such, learning is influenced by the sequence of study in, hopefully, predictable ways. We believe real progress in this area will come from mechanistic theories and models grounded on solid laboratory and *in vivo* research. It will most likely not be enough to demonstrate that interleaved practice works or not – we should strive to explain in detail the learning process that is changed by interleaved practice. Only the existence of strong theories and models will allow successful and systematic use of different sequences of practice as a tool to improve learning. We believe SAT is a good initial model to understanding the learning process and how it is changed by sequence, but it can certainly be improved, expanded, or challenged.

Additionally, there are several areas where research is still lacking. For example, most research either does not include or ignores the fact that students often study the same materials several times, perhaps always in the same sequence. What consequences does repeating the same sequence of items have? Initial evidence from Kost and colleagues (2015) shows that repetition of the sequence of learning improves learning during blocked practice. One possibility is that the repetition allows learners to contrast the category similarities acquired during the first pass-through with other categories during subsequent repetitions. However, these more complex learning sequences are still very much an open question, even though they are common in students' practice.

Moreover, in our effort to find the best sequence between interleaved and blocked practice, we might also have missed the opportunity to find the best sequence that includes both interleaved and blocked practice. If one sequence directs attention toward similarities among successive items and the other differences between successive items, a carefully crafted version mixing the two might be the best of both worlds.

Similarly, not every blocked sequence is equal. For example, it is possible that the way blocked practice is implemented will have an impact in its effectiveness. One way blocked practice is often implemented is by studying all materials of one concept before starting the next (e.g., Kornell & Bjork, 2008; Rawson et al., 2015; Ste-Marie et al., 2004; Zulkiply & Burt, 2013). This implementation might, in effect, reduce the potential benefits of blocked practice by reducing students' opportunity to test hypotheses during study and reducing attention to the task – both factors that have been shown to matter for improved learning (Abel & Roediger, 2017; Karpicke & Blunt, 2011; Wahlheim et al., 2011). However, it is possible to equate the informativity, attention, and the ability to generate hypotheses between interleaved and blocked practice. For example, Carvalho and Goldstone (2014a, 2014b, 2015a) have implemented blocked practice by including reduced alternation between categories (25 percent of transitions are changes between concepts). Conversely, interleaved practice includes a high level of alternation between categories (75 percent transitions are changes between concepts). This approach allows students to generate hypotheses and test them in both sequences, and makes both sequences unpredictable and therefore equates for attentional demands. This approach also matches how

students often decide to block their study (Carvalho et al., 2016; Tauber et al., 2013). It is currently unclear whether, when the two sequences are implemented in such a way, it is possible to increase the range of situations that benefit from blocked practice.

There is also a lack of research that investigates differences in learning between concepts studied in different positions during blocked practice. As we mentioned earlier in this chapter, early research on sequential effects on learning focused on finding which concepts should be studied first and which should be studied afterward. However, most research on interleaved versus blocked practice compares interleaved practice with a randomized version of blocked practice, where this factor is “controlled for” by careful experimental practice. Perhaps a carefully crafted blocked practice that considers difficulty level to determine the order of the concepts is a more appropriate sequence to compare with interleaved practice. Consistent with this prediction, Patel, Liu, and Koedinger (2015) recently demonstrated that middle school students practicing fraction addition and fraction multiplication showed equivalent learning with interleaved and blocked practice if blocked practice started with fraction multiplication followed by fraction addition. When blocked practice started with fraction addition, however, students performed worse than in interleaved practice. This research is initial evidence that bringing to bear early research on concept sequencing to the study of interleaved and blocked practice can prove fruitful.

Learning is also complex in terms of multidimensionality of the learning materials, a factor often ignored in current research. For example, when sequencing practice of different concepts (addition, subtraction, multiplication) with different types of problems, should one interleave the concepts and block the types of problems, interleave both, or any other combination? To date, only one study by Rau and colleagues (2013) probed this question. The authors found that in the context of fraction learning with different types of problems and representations, best learning was achieved by interleaving type of problem and blocking type of representation. Further study is needed to identify common properties across multidimensional materials, where one dimension benefits from interleaved practice and the other from blocked practice, to extrapolate general principles to support implementation and theoretical understanding.

Another important area is investigating to what extent students’ decisions influence and shape the effect of different sequences. Although it is possible that blocked practice might not be optimal for certain situations, when students decide to block their study (and they often do), it might result in best learning. This is because the attentional and memory demands of self-regulated tasks differ from tasks where information is not controlled by the learner (Gureckis & Markant, 2012; Markant & Gureckis, 2014; Markant, Settles, & Gureckis, 2016). Most current studies have not compared students’ performance between self-regulated and yoked study to identify the impact of self-regulation on the effectiveness of different sequences of study to promote learning (however, see Carvalho et al., 2016). We see the study of sequencing in self-regulated practice

as a major next step in this line of research – not only because it can provide novel insights not currently available but also as an entry way into understanding how to change and improve metacognitive awareness to the benefits of choosing the right sequence of study for the right situation.

## References

- Abel, M. & Roediger, H. L. (2017). Comparing the testing effect under blocked and mixed practice: The mnemonic benefits of retrieval practice are not affected by practice format. *Memory and Cognition*, *45*(1), 81–92. <https://doi.org/10.3758/s13421-016-0641-8>
- Albaret, J. M. & Thon, B. (1998). Differential effects of task complexity on contextual interference in a drawing task. *Acta Psychologica*, *100*(1–2), 9–24.
- Al-Mustafa, A. A. (1989). Contextual interference: Laboratory artifact or sport skill learning related? (Doctoral dissertation).
- Andrews, S. E. & Frey, S. D. (2015). Studio structure improves student performance in an undergraduate introductory soil science course. *Natural Sciences Education*, *44*(1), 60. <https://doi.org/10.4195/nse2014.12.0026>
- Beilock, S. L., Carr, T. H., MacMahon, C., & Starkes, J. L. (2002). When paying attention becomes counterproductive: Impact of divided versus skill-focused attention on novice and experienced performance of sensorimotor skills. *Journal of Experimental Psychology: Applied*, *8*(1), 6–16. <https://doi.org/10.1037//1076-898X.8.1.6>
- Best, C. A., Robinson, C. W., & Sloutsky, V. M. (2011). The effect of labels on children's category learning. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 3332–3336). Austin, TX.
- Best, C. A., Yim, H., & Sloutsky, V. M. (2013). The cost of selective attention in category learning: Developmental differences between adults and infants. *Journal of Experimental Child Psychology*, *116*(2), 105–119. <https://doi.org/10.1016/j.jecp.2013.05.002>
- Birnbaum, M. S., Kornell, N., Bjork, E. L., & Bjork, R. A. (2013). Why interleaving enhances inductive learning: The roles of discrimination and retrieval. *Memory and Cognition*, *41*(3), 392–402. <https://doi.org/10.3758/s13421-012-0272-7>
- Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. P. Shimamura (eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.
- Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*, 417–444. <https://doi.org/10.1146/annurev-psych-113011-143823>
- Carpenter, S. K. & Mueller, F. E. (2013). The effects of interleaving versus blocking on foreign language pronunciation learning. *Memory and Cognition*, *41*(5), 671–682. <https://doi.org/10.3758/s13421-012-0291-4>
- Carter, K., Cushing, K., Sabers, D., Stein, P., & Berliner, D. (1988). Expert-novice differences in perceiving and processing visual classroom information. *Journal of Teacher Education*, *39*(3), 25–31. <https://doi.org/10.1177/002248718803900306>
- Carvalho, P. F. & Albuquerque, P. B. (2012). Memory encoding of stimulus features in human perceptual learning. *Journal of Cognitive Psychology*, *24*(6), 654–664. <https://doi.org/10.1080/20445911.2012.675322>



- Carvalho, P. F., Braithwaite, D. W., de Leeuw, J. R., Motz, B. A., & Goldstone, R. L. (2016). An in vivo study of self-regulated study sequencing in introductory psychology courses. *Plos One*, *11*(3), e0152115. <https://doi.org/10.1371/journal.pone.0152115>
- Carvalho, P. F. & Goldstone, R. L. (2014a). Effects of interleaved and blocked study on delayed test of category learning generalization. *Frontiers in Psychology*, *5*, 936. <https://doi.org/10.3389/fpsyg.2014.00936>
- (2014b). Putting category learning in order: Category structure and temporal arrangement affect the benefit of interleaved over blocked study. *Memory and Cognition*, *42*(3), 481–495. <https://doi.org/10.3758/s13421-013-0371-0>
- (2015a). The benefits of interleaved and blocked study: Different tasks benefit from different schedules of study. *Psychonomic Bulletin and Review*, *22*(1), 281–288. <https://doi.org/10.3758/s13423-014-0676-4>
- (2015b). What you learn is more than what you see: What can sequencing effects tell us about inductive category learning? *Frontiers in Psychology*, *6*, 505. <https://doi.org/10.3389/fpsyg.2015.00505>
- (2017a). The most efficient sequence of study depends on the type of test. Manuscript under review.
- (2017b). The sequence of study changes what information is attended to, encoded, and remembered during category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Chariker, J. H., Naaz, F., & Pani, J. R. (2011). Computer-based learning of neuroanatomy: A longitudinal study of learning, transfer, and retention. *Journal of Educational Psychology*, *103*(1), 19–31. <https://doi.org/10.1037/a0021680>
- Cook, G. L. & Odom, R. D. (1992). Perception of multidimensional stimuli: A differential-sensitivity account of cognitive processing and development. *Journal of Experimental Child Psychology*, *54*(2), 213–249.
- Deng, W. S. & Sloutsky, V. M. (2016). Selective attention, diffused attention, and the development of categorization. *Cognitive Psychology*, *91*, 24–62. <https://doi.org/10.1016/j.cogpsych.2016.09.002>
- de Zilva, D., Mitchell, C. J., & Newell, B. R. (2013). Eliminating the mere exposure effect through changes in context between exposure and test. *Cognition and Emotion*, *27*(8), 1345–1358. <https://doi.org/10.1080/02699931.2013.775110>
- Dobson, J. L. (2011). Effect of selected “desirable difficulty” learning strategies on the retention of physiology information. *Advances in Physiology Education*, *35*(4), 378–383.
- Elio, R. & Anderson, J. R. (1981). The effects of category generalizations and instance similarity on schema abstraction. *Journal of Experimental Psychology: Human Learning and Memory*, *7*(6), 397–417.
- (1984). The effects of information order and learning mode on schema abstraction. *Memory and Cognition*, *12*(1), 20–30.
- Engle, R. W. & Kane, M. J. (2003). *Executive attention, working memory capacity, and a two-factor theory of cognitive control*, *44*, 145–199. [https://doi.org/10.1016/S0079-7421\(03\)44005-X](https://doi.org/10.1016/S0079-7421(03)44005-X)
- Farrow, D. & Maschette, W. (1997). The effects of contextual interference on children learning forehand tennis groundstrokes. *Journal of Human Movement Studies*, *33*, 47–67.
- Goldstone, R. L. (1996). Isolated and interrelated concepts. *Memory and Cognition*, *24*(5), 608–628.



- Gureckis, T. M. & Markant, D. B. (2012). Self-directed learning: A cognitive and computational perspective. *Perspectives on Psychological Science : A Journal of the Association for Psychological Science*, 7(5), 464–481. <https://doi.org/10.1177/1745691612454304>
- Hatala, R. M., Brooks, L. R., & Norman, G. R. (2003). Practice makes perfect: The critical role of mixed practice in the acquisition of ECG interpretation skills. *Advances in Health Sciences Education: Theory and Practice*, 8(1), 17–26.
- Hebert, E. P., Landin, D., & Solmon, M. A. (1996). Practice schedule effects on the performance and learning of low- and high-skilled students: An applied study. *Research Quarterly for Exercise and Sport*, 67(1), 52–58. <https://doi.org/10.1080/02701367.1996.10607925>
- Higgins, E. J. (2017). The complexities of learning categories through comparisons. *Psychology of Learning and Motivation*, 66, 43–77. <https://doi.org/10.1016/bs.plm.2016.11.002>
- Higgins, E. J. & Ross, B. H. (2011). Comparisons in category learning: How best to compare for what? In L. Carlson, C. Holscher, & T. Shipley (eds.), *Proceedings of the 33rd annual conference of the Cognitive Science Society* (pp. 1388–1393). Austin, TX: Cognitive Science Society.
- Horn, S. & Hernick, M. (2015). Improving student understanding of lipids concepts in a biochemistry course using test-enhanced learning. *Chemical Education Research and Practice*, 16(4), 918–928. <https://doi.org/10.1039/C5RP00133A>
- Hull, L. (1920). Quantitative aspects of evolution and concepts: An experimental study. *The Psychological Monographs*, 28(1), 1–86.
- Jones, M. & Sieck, W. R. (2003). Learning myopia: An adaptive recency effect in category learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 29(4), 626–640.
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology. General*, 133(2), 189–217. <https://doi.org/10.1037/0096-3445.133.2.189>
- Kang, S. H. K. & Pashler, H. (2012). Learning painting styles: Spacing is advantageous when it promotes discriminative contrast. *Applied Cognitive Psychology*, 26(1), 97–103. <https://doi.org/10.1002/acp.1801>
- Karpicke, J. D. & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, 331(6018), 772–775. <https://doi.org/10.1126/science.1199327>
- Kioumourtzoglou, E., Kourtessis, T., Michalopoulou, M., & Derri, V. (1998). Differences in several perceptual abilities between experts and novices in basketball, volleyball and water-polo. *Perceptual and Motor Skills*, 86(3), 899–912. <https://doi.org/10.2466/pms.1998.86.3.899>
- Kirchoff, B. K., Delaney, P. F., Horton, M., & Dellinger-Johnston, R. (2014). Optimizing learning of scientific category knowledge in the classroom: The case of plant identification. *CBE Life Sciences Education*, 13(3), 425–436. <https://doi.org/10.1187/cbe.13-11-0224>
- Kornell, N. & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the “enemy of induction”? *Psychological Science*, 19(6), 585–592. <https://doi.org/10.1111/j.1467-9280.2008.02127.x>

- 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*(2), 498–503.
- Kost, A. S., Carvalho, P. F., & Goldstone, R. L. (2015). Can you repeat that? The effect of item repetition on interleaved and blocked study. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (eds.), *Proceedings of the 37th annual meeting of the Cognitive Science Society* (pp. 1189–1194). Austin, TX: Cognitive Science Society.
- Kurtz, K. H. & Hovland, C. I. (1956). Concept learning with differing sequences of instances. *Journal of Experimental Psychology, 51*(4), 239–243. <https://doi.org/10.1037/h0040295>
- Lee, E. S., MacGregor, J. N., Bavelas, A., Mirlin, L., & et al. (1988). The effects of error transformations on classification performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(1), 66–74. <https://doi.org/10.1037/0278-7393.14.1.66>
- Lee, T. D. & Magill, R. A. (1985). Can forgetting facilitate skill acquisition? In D. Goodman, R. B. Wilberg, & I. M. Franks (eds.), *Differing perspectives in motor learning, memory, and control* (pp. 3–22). Amsterdam: North-Holland.
- Li, N., Cohen, W. W., & Koedinger, K. R. (2013). Problem order implications for learning. *International Journal of Artificial Intelligence in Education, 23*(1–4), 71–93. <https://doi.org/10.1007/s40593-013-0005-5>
- Linderholm, T., Dobson, J., & Yarbrough, M. B. (2016). The benefit of self-testing and interleaving for synthesizing concepts across multiple physiology texts. *Advances in Physiology Education, 40*(3), 329–334. <https://doi.org/10.1152/advan.00157.2015>
- Mack, M. L. & Palmeri, T. J. (2015). The dynamics of categorization: Unraveling rapid categorization. *Journal of Experimental Psychology: General, 144*(3), 551–569. <https://doi.org/10.1037/a0039184>
- Mack, M. L. & Preston, A. R. (2016). Decisions about the past are guided by reinstatement of specific memories in the hippocampus and perirhinal cortex. *Neuroimage, 127*, 144–157. <https://doi.org/10.1016/j.neuroimage.2015.12.015>
- Markant, D. B. & Gureckis, T. M. (2014). Is it better to select or to receive? Learning via active and passive hypothesis testing. *Journal of Experimental Psychology: General, 143*(1), 94–122. <https://doi.org/10.1037/a0032108>
- Markant, D. B., Settles, B., & Gureckis, T. M. (2016). Self-Directed learning favors local, Rather than global, Uncertainty. *Cognitive Science, 40*(1), 100–120. <https://doi.org/10.1111/cogs.12220>
- Markman, A. B. & Ross, B. H. (2003). Category use and category learning. *Psychological Bulletin, 129*(4), 592–613. <https://doi.org/10.1037/0033-2909.129.4.592>
- Mathy, F. & Feldman, J. (2009). A rule-based presentation order facilitates category learning. *Psychonomic Bulletin and Review, 16*(6), 1050–1057. <https://doi.org/10.3758/PBR.16.6.1050>
- Mathy, F., Haladjian, H. H., Laurent, E., & Goldstone, R. L. (2013). Similarity-dissimilarity competition in disjunctive classification tasks. *Frontiers in Psychology, 4*, 26. <https://doi.org/10.3389/fpsyg.2013.00026>
- McDaniel, M. A., Fadler, C. L., & Pashler, H. (2013). Effects of spaced versus massed training in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(5), 1417–1432. <https://doi.org/10.1037/a0032184>

- McNamara, D. S. & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes*, 22(3), 247–288. <https://doi.org/10.1080/01638539609544975>
- Mitchell, C. J., Nash, S., & Hall, G. (2008). The intermixed-blocked effect in human perceptual learning is not the consequence of trial spacing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 237–242.
- Monteiro, S., Melvin, L., Manolakos, J., Patel, A., & Norman, G. (2017). Evaluating the effect of instruction and practice schedule on the acquisition of ECG interpretation skills. *Perspectives on Medical Education*, 6(4), 237–245.
- National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common Core State Standards: English Language, Arts & Literacy in History/Social Studies, Science, and Technical Subject*. [www.corestandards.org/wp-content/uploads/ELA\\_Standards1.pdf](http://www.corestandards.org/wp-content/uploads/ELA_Standards1.pdf)
- National Research Council. (2012). *A framework for K-12 science education: Practices, Crosscutting concepts, and Core ideas*. Washington, DC: National Academies Press. <https://doi.org/10.17226/13165>
- Noh, S. M., Yan, V. X., Bjork, R. A., & Maddox, W. T. (2016). Optimal sequencing during category learning: Testing a dual-learning systems perspective. *Cognition*, 155, 23–29. <https://doi.org/10.1016/j.cognition.2016.06.007>
- Ostrow, K., Heffernan, N., Heffernan, C., & Peterson, Z. (2015). Blocking vs. interleaving: Examining single-session effects within middle school math homework. In C. Conati, N. Heffernan, A. Mitrovic, & M. F. Verdejo (eds.), *Artificial intelligence in education*, Vol. 9112 (pp. 338–347). Cham: Springer. [https://doi.org/10.1007/978-3-319-19773-9\\_34](https://doi.org/10.1007/978-3-319-19773-9_34)
- Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84(4), 429–434. <https://doi.org/10.1037/0022-0663.84.4.429>
- Palmeri, T. J. & Mack, M. L. (2015). How experimental trial context affects perceptual categorization. *Frontiers in Psychology*, 6, 180. <https://doi.org/10.3389/fpsyg.2015.00180>
- Pani, J. R., Chariker, J. H., & Naaz, F. (2012). Computer-based learning: Interleaving whole and sectional representation of neuroanatomy. *Anatomical Sciences Education*, 6(1), 11–18.
- Patel, R., Liu, R., & Koedinger, K. (2015). When to block versus interleave practice? Evidence against teaching fraction addition before fraction multiplication. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (eds.), *Proceedings of the 38th annual conference of the Cognitive Science Society* (pp. 2069–2074). Austin, TX: Cognitive Science Society.
- Pigott, R. E. & Shapiro, D. C. (1984). Motor schema: The structure of the variability session. *Research Quarterly for Exercise and Sport*, 55, 41–45.
- Pinto-Zipp, G. & Gentile, A. M. (1995). Practice schedules in motor learning: Children vs. adults. *Society of Neuroscience: Abstracts*, 21, 1620.
- Rau, M. A., Alevin, V., & Rummel, N. (2013). Interleaved practice in multi-dimensional learning tasks: Which dimension should we interleave? *Learning and Instruction*, 23, 98–114. <https://doi.org/10.1016/j.learninstruc.2012.07.003>
- Rawson, K. A., Thomas, R. C., & Jacoby, L. L. (2015). The power of examples: Illustrative examples enhance conceptual learning of declarative concepts. *Educational Psychology Review*, 27(3), 483–504. <https://doi.org/10.1007/s10648-014-9273-3>

- Rohrer, D. (2012). Interleaving helps students distinguish among similar concepts. *Educational Psychology Review*, 24(3), 355–367.
- (2015). Student instruction should be distributed over long time periods. *Educational Psychology Review*, 27(4), 635–643. <https://doi.org/10.1007/s10648-015-9332-4>
- 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 and Review*, 21(5), 1323–1330. <https://doi.org/10.3758/s13423-014-0588-3>
- Rohrer, D., Dedrick, R. F., & Stershic, S. (2015). Interleaved practice improves mathematics learning. *107(3)*, 900–908.
- Rohrer, D. & Taylor, K. (2007). The shuffling of mathematics problems improves learning. *Instructional Science*, 35(6), 481–498.
- Ross, B. H. (2000). The effects of category use on learned categories. *Memory and Cognition*, 28(1), 51–63.
- Rozenstein, A., Pearson, G. D. N., 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(8), 979–984. <https://doi.org/10.1016/j.jacr.2016.03.031>
- Sana, F., Yan, V. X., & Kim, J. A. (2016). Study sequence matters for the inductive learning of cognitive concepts. *Journal of Educational Psychology*.
- Sandhofer, C. M. & Dumas, L. A. A. (2008). Order of presentation effects in learning color categories. *Journal of Cognition and Development*, 9(2), 194–221.
- Schlichting, M. L. & Preston, A. R. (2015). Memory integration: Neural mechanisms and implications for behavior. *Current Opinion in Behavioral Sciences*, 1, 1–8. <https://doi.org/10.1016/j.cobeha.2014.07.005>
- Shea, C H & Kohl, R. M. (1991). Composition of practice: Influence on the retention of motor skills. *Research Quarterly for Exercise and Sport*, 62(2), 187–195. <https://doi.org/10.1080/02701367.1991.10608709>
- Shea, C. H., Kohl, R., & Indermill, C. (1990). Contextual interference: Contributions of practice. *Acta Psychologica*, 73(2), 145–157. [https://doi.org/10.1016/0001-6918\(90\)90076-R](https://doi.org/10.1016/0001-6918(90)90076-R)
- 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(2), 179–187.
- Smith, L. B. & Kemler, D. G. (1978). Levels of experienced dimensionality in children and adults. *Cognitive Psychology*, 10, 502–532.
- Sorensen, L. J. & Woltz, D. J. (2016). Blocking as a friend of induction in verbal category learning. *Memory and Cognition*, 44(7), 1000–1013. <https://doi.org/10.3758/s13421-016-0615-x>
- Spiering, B. J. & Ashby, F. G. (2008). Initial training with difficult items facilitates information integration, but not rule-based category learning. *Psychological Science*, 19(11), 1169–1177. <https://doi.org/10.1111/j.1467-9280.2008.02219.x>
- Ste-Marie, D. M., Clark, S. E., Findlay, L. C., & Latimer, A. E. (2004). High levels of contextual interference enhance handwriting skill acquisition. *Journal of Motor Behavior*, 36(1), 115–126.
- Stewart, N. & Brown, G. D. A. (2004). Sequence effects in the categorization of tones varying in frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(2), 416–430.

- Stewart, N., Brown, G. D. A., & Chater, N. (2002). Sequence effects in categorization of simple perceptual stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(1), 3–11.
- Strutt, G. F., Anderson, D. R., & Well, A. D. (1975). A developmental study of the effects of irrelevant information on speeded classification. *Journal of Experimental Child Psychology*, 20(1), 127–135. [https://doi.org/10.1016/0022-0965\(75\)90032-6](https://doi.org/10.1016/0022-0965(75)90032-6)
- Tauber, S. K., Dunlosky, J., Rawson, K. A., Wahlheim, C. N., & Jacoby, L. L. (2013). Self-regulated learning of a natural category: Do people interleave or block exemplars during study? *Psychonomic Bulletin and Review*, 20(2), 3560363.
- Taylor, K. & Rohrer, D. (2010). The effects of interleaved practice. *Applied Cognitive Psychology*, 24(6), 837–848.
- Terrace, H. S. (1964). Wavelength generalization after discrimination learning with and without errors. *Science*, 144(361), 78–80.
- Thompson, L. A. & Markson, L. (1998). Developmental changes in the effect of dimensional salience on the discriminability of object relations. *Journal of Experimental Child Psychology*, 70(1), 1–25. <https://doi.org/10.1006/jecp.1998.2445>
- Tse, C.-S. & Altarriba, J. (2010). Does survival processing enhance implicit memory? *Memory and Cognition*, 38(8), 1110–1121.
- Unsworth, N. & Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, 114(1), 104–132. <https://doi.org/10.1037/0033-295X.114.1.104>
- Unsworth, N. & Spillers, G. J. (2010). Variation in working memory capacity and episodic recall: the contributions of strategic encoding and contextual retrieval. *Psychonomic Bulletin and Review*, 17(2), 200–205. <https://doi.org/10.3758/PBR.17.2.200>
- Wahlheim, C. N., Dunlosky, J., & Jacoby, L. L. (2011). Spacing enhances the learning of natural concepts: an investigation of mechanisms, metacognition, and aging. *Memory and Cognition*, 39(5), 750–763. <https://doi.org/10.3758/s13421-010-0063-y>
- Werner, S. & Thies, B. (2000). Is “change blindness” attenuated by domain-specific expertise? An expert-novices comparison of change detection in football images. *Visual Cognition*, 7(1–3), 163–173. <https://doi.org/10.1080/135062800394748>
- Whitman, J. R. & Garner, W. R. (1963). Concept learning as a function of form of internal structure. *Journal of Verbal Learning and Verbal Behavior*, 2(2), 195–202.
- Wulf, G. & Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic Bulletin and Review*, 9(2), 185–211.
- Yamauchi, T. & Markman, A. B. (2000). Inference using categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(3), 776–795.
- 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(7), 918–933.
- Yan, V. X., Soderstrom, N. C., Seneviratna, G. S., Bjork, E. L., & Bjork, R. A. (2017). How should exemplars be sequenced in inductive learning? Empirical evidence versus learners’ opinions. *Journal of Experimental Psychology: Applied*
- Zeithamova, D., Schlichting, M. L., & Preston, A. R. (2012). The hippocampus and inferential reasoning: Building memories to navigate future decisions. *Frontiers in Human Neuroscience*, 6, 70. <https://doi.org/10.3389/fnhum.2012.00070>

- Zulkiply, N. & Burt, J. S. (2013). The exemplar interleaving effect in inductive learning: moderation by the difficulty of category discriminations. *Memory and Cognition*, 41(1), 16–27. <https://doi.org/10.3758/s13421-012-0238-9>
- Zulkiply, 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(3), 215–221. <https://doi.org/10.1016/j.learninstruc.2011.11.002>

PROOF