Notes for picture researcher: interleaving, learning with space delays

<CT>The application of spacing and interleaving approaches in the classroom

Jonathan Firth, teacher of psychology in Glasgow and in the Outer Hebrides, and teacher-educator at the University of Strathclyde, UK

Spacing and interleaving approaches have attracted significant interest over the past few years, with research suggesting they have the potential to support both retention and transfer of learning. The small-scale study reported here, in conjunction with a review of the literature, was undertaken to explore the effect of interleaving and spacing in authentic classroom learning of concepts.

<A>An overview of interleaving and spacing

Interleaving refers to the benefits of sequencing learning tasks so that similar items – two examples of the same concept, say – are interspersed with different types of items rather than being consecutive. This results in a more variable and challenging task but is associated with benefits in terms of memory and transfer, which apply to concept learning as well as other domains (Kang, 2016).

Interleaving tends to be contrasted with classroom tasks that are scheduled in blocks of similar items, with the latter termed a 'blocked' arrangement. For example, in Figure 1, item 'A' is interleaved with items 'B' and 'C'.

<INSERT FIGURE 1 HERE>

The spacing effect, meanwhile, refers to the benefit of incorporating time delays between learning and practice, leading to improved performance over educationally relevant time periods (Cepeda et al., 2008), compared to 'massed' items, where practice sessions occur close together. An interleaved presentation of material is inevitably spaced to some extent, given that sequencing items so that they are non-consecutive leads to a time delay between one example and the next. For example, in Figure 1, above, there are larger gaps between the instances of item A in the interleaving sequence than in the blocked sequence.

However, a key difference between the two effects is that when items are spaced, the gaps between learning and practice needn't include related material. Indeed, the benefits of interleaving seem to depend on the mixing of related items, such as examples from similar categories. This may be because such scheduling puts different items side by side, improving the perception of differences between them (Kornell and Bjork, 2008). This is known as the discriminative-contrast hypothesis, and is supported by research into interleaving of unrelated categories. One such study (Hausman and Kornell, 2014) interleaved anatomy terms with

Indonesian vocabulary and found no benefit. Interleaving of related items is effective, especially when differences between items are subtle (Carvalho and Goldstone, 2014).

<INSERT FIGURE 2 HERE>

Both interleaving and spacing are what Bjork and Bjork (2011) term ‘desirable difficulties’, i.e. strategies that make learning more difficult, but in a way that is beneficial. They are both widely recommended among those who aim to apply cognitive psychology to education, appearing among the ‘Six Strategies for Effective Learning’ in the popular Learning Scientists blog (www.learningscientists.org).

The two strategies differ in their primary benefits. The spacing effect boosts memory – practice or restudy of material is more effective if spaced out over time – while interleaving boosts inductive category learning and later transfer. Recent demonstrations of this include the categorisation of chemicals into types (Eglington and Kang, 2017) and the conceptual learning of science categories or examples (e.g. Rawson et al., 2015).

However, there are some caveats. Smith and Scarf (2017) note that for spacing learning across days to be helpful, a minimum initial level of experience is required. Davis et al. (2017) have found that frequent switching between studying and test questions can be detrimental, while Kang (2016) reasons that a hybrid approach can be beneficial, with new learning occurring via blocked practice, and interleaving used in a practice or consolidation phase.

<A>Research study

To explore the effects of spacing and interleaving in an authentic classroom context, a classroom-based research project was undertaken. An opportunity sample of 31 school pupils between 16 and 17 years of age was used. Data was gathered during an end-of-year taster session, during which pupils sampled several subjects. They were entirely new to the topic being learned.

In order to make the tasks as authentic as possible, all materials were based around a school psychology specification, featuring psychological theories of phobias. The experiment aimed to reproduce the range of activities in a typical school class, and so learners were taught both the concepts (types of phobia) and relevant research evidence.

Tasks were delivered via an online protocol. The target learning material was presented in two main phases. In the first, a research study pertaining to phobias was shown on two screens, one with a description of the study and one with evaluation points, with the latter either presented immediately (massed condition) or after the second phase (spaced condition). The precise time delay for each participant therefore depended on their reading speed during the second phase (reading 353 words on screen); pilot testing had indicated a delay of two to three minutes.

During the second phase, types of phobia (specific phobia, agoraphobia and social anxiety) were defined, with key diagnostic information given; concepts and information were either presented together (blocked) or mixed with information about different types of phobia on the same screen (interleaved). For example, in the blocked condition, a participant would view three items relating to agoraphobia, while in the interleaved condition they would view a key feature of each of the three types. The online tasks also featured a test, comprising multiple choice questions about the research studies and the categorisation of novel examples of each type of phobia.

Participants sat at individual PCs, and a teacher oversaw the session. After a general briefing, each completed an on-screen consent form, followed by viewing the material presented in an order that depended on allocation to experimental conditions, which was decided via random numbers. As soon as participants had completed the task, the software automatically initiated the test.

Ethics approval followed the school's framework; as a research-focused school, it had set up its own in-school ethics board, with an academic panel providing oversight.

<A>Findings

The mean percentages of correct answers on the end-of-task test for the interleaved and blocked conditions are shown in Figure 3. A between-subjects ANOVA was carried out. This analysis revealed a significant main effect of spacing (performance in the spaced condition being worse than the massed condition, with mean scores of 12.25 vs 9.45, $p = .002$), while interleaving did not have a significant main effect.

<INSERT FIGURE 3 HERE>

<B>Discussion

The findings demonstrated that spacing had a harmful effect on outcomes in the immediate test, while the main effect of interleaving was neutral. The results fit with the idea that these are 'desirable difficulties', with the potential to impede learning in the short term. Soderstrom and Bjork (2015) describe how such strategies often lead to performance being slower and more error-prone, but improve learning over longer intervals.

Nevertheless, increased errors within a short learning session could suggest inefficiency in the learning process, and raise questions about the use of spacing, in particular with new concept learning. In a related finding, Donovan and Radosevich (1999) found that spacing was not beneficial for complex tasks; complexity interacts with learner experience and, when learning a new concept, complexity for a learner can be high. Such an explanation suggests that desirable difficulties interact with learner skill, as proposed by McDaniel and Butler (2011).

One way to get around short-term difficulties is to utilise a hybrid schedule, with interleaved or spaced practice being utilised subsequent to an initial learning phase. This fits with the recommendation of Rawson and Dunlosky (2011) that learners should first automise recall of concepts, and later retrieve items three times at widely spaced intervals. Interestingly, Yan et al. (2017) found that a blocked-to-interleaved schedule was not superior to pure interleaving, which fits the current findings that interleaving alone did not cause short-term difficulties.

An interesting finding which is worthy of further investigation, and is discussed in more depth in the online version of this article, was the significant ($p = .009$) interaction between the two variables (spacing vs interleaving), indicating that interleaving had a mediating or protective effect against the difficulties caused by spacing. On the face of it, this result does not fit well with Birnbaum et al.'s (2013) finding of a negative interaction between spacing and interleaving; however, it is important to note that unlike in that experiment, the two interventions here were used in different task phases, and spacing therefore didn't prevent discriminative contrast.

<B>Limitations and areas for further research

The present study was limited in terms of its scope and sample. A small number of pupils at the same stage of education were tested; any discussion of the effects of spacing and interleaving must take account of individual learner differences, and future studies should increase sample size and diversity.

As the learners in this study were new to the material being studied, a control group was not deemed necessary, given that prior concept knowledge was assumed to be absent. Future studies could confirm this assumption by presenting the test phase alone to a comparable group of pupils.

The present study focused on a relatively brief set of tasks within a single lesson, and future work could follow learners over a longer period to see how desirable difficulties play out across the learning of a topic. It is also essential to establish the extent to which the negative short-term classroom effects of interleaving and spacing would be counteracted by improved long-term ability to remember and transfer learning.

<A>References

Birnbaum MS, Kornell N, Bjork E et al. (2013) Why interleaving enhances inductive learning: The roles of discrimination and retrieval. *Memory & Cognition* 41(3): 392–402.

Bjork EL and Bjork RA (2011) Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In: Gernsbacher MA, Pew RW, Hough LM et al. (eds) *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society*. New York: Worth, pp. 56–64.

Carvalho PF and Goldstone RL (2014) Putting category learning in order: Category structure and temporal arrangement affect the benefit of interleaved over blocked study. *Memory & Cognition* 42(3): 481–495.

Cepeda NJ, Vul E, Rohrer D et al. (2008) Spacing effects in learning a temporal ridgeline of optimal retention. *Psychological Science* 19(11): 1095–1102.

Davis SD, Chan JCK and Wilford MM (2017) The dark side of interpolated testing: Frequent switching between retrieval and encoding impairs new learning. *Journal of Applied Research in Memory and Cognition.* Epub ahead of print.

Donovan JJ and Radosevich DJ (1999) A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology* 84: 795–805.

Eglington LG and Kang SH (2017) Interleaved presentation benefits science category learning. *Journal of Applied Research in Memory and Cognition* 6(4): 475–485.

Hausman H and Kornell N (2014) Mixing topics while studying does not enhance learning. *Journal of Applied Research in Memory and Cognition* 3(3): 153–160.

Kang SH (2016) The benefits of interleaved practice for learning. In: Horvath JC, Lodge JM and Hattie J (eds) *From the Laboratory to the Classroom: Translating Science of Learning for Teachers.* London: Routledge, pp. 79–93.

Kornell N and Bjork RA (2008) Learning concepts and categories: Is spacing the 'enemy of induction'? *Psychological Science* 19: 585–592.

McDaniel MA and Butler AC (2011) A contextual framework for understanding when difficulties are desirable. In: Benjamin AS (ed) *Successful Remembering and Successful Forgetting: A Festschrift in Honor of Robert A. Bjork*. New York: Psychology Press, pp. 175–198.

Rawson KA and Dunlosky J (2011) Optimizing schedules of retrieval practice for durable and efficient learning: How much is enough? *Journal of Experimental Psychology: General* 140(3): 283–302.

Rawson KA, Thomas RC and Jacoby LL (2015) The power of examples: Illustrative examples enhance conceptual learning of declarative concepts. *Educational Psychology Review* 27: 483–504.

Smith CD and Scarf D (2017) Spacing repetitions over long timescales: A review and a reconsolidation explanation. *Frontiers in Psychology* 8: 962.

Soderstrom NC and Bjork RA (2015) Learning versus performance: An integrative review. *Perspectives on Psychological Science* 10(2): 176–199.

Yan VX, Soderstrom NC, Seneviratna GS et al. (2017) How should exemplars be sequenced in inductive learning? Empirical evidence versus learners' opinions. *Journal of Experimental Psychology: Applied*. Epub ahead of print.