Boosting learning by changing the order and timing of classroom tasks: Implications for professional practice.

Jonathan Firth

School of Education, University of Strathclyde, Glasgow, United Kingdom

Correspondence address: Jonathan Firth, School of Education, The University of Strathclyde, 141 St James Road, Glasgow, Scotland, G4 0LT. Tel: +44 (0)141 444 8069. Email: jonathan.firth@strath.ac.uk.

Boosting learning by changing the order and timing of classroom tasks: Implications for professional practice.

The timing of what occurs in the classroom can affect how successfully new concepts are learned. This paper analyses two promising ways of modifying the schedule of tasks or examples – the spacing effect and interleaving. The spacing effect refers to improvements in long-term retention if practice sessions are separated by delays. Interleaving refers to the benefits of mixing different types of examples or problems within a series, such that contrasting items are side by side. These techniques are supported by a body of evidence which suggests that they can boost the durability and transferability of learning, and both are straightforward to implement given that they only require manipulating the schedule of learning experiences. However, research into metacognition suggests that classroom use of these techniques is likely to be undermined by misconceptions about how learning and memory work, and that teachers may not become aware of such misconceptions through experience alone. This leads to a potential disparity between research evidence and popular assumptions about learning and memory. While this can in part be addressed by teachers engaging more with evidence-based practice, use of techniques such as spacing and interleaving is highly context-specific. As discussed in the paper, this issue implies a need for focused and contextualised professional learning, which could be enhanced via the application of the same techniques to the professional learning activities carried out by new and in-service teachers.

Keywords: spacing effect; interleaving; instructional practices, teacher beliefs; practice-based teacher education.

1. Introduction

1.1 Learning and the role of memory

Learning involves a process of taking in new facts, skills, information and concepts in a way that is durable and transferrable (Bransford, Brown & Cocking, 2000). However, the learning process itself is not always intuitive. There is a tendency to mistake performance for learning (Soderstrom & Bjork, 2015); a teacher may think that learning has been accomplished at the end of a class or lecture, and fail to take account of forgetting. As such, they may overlook certain useful classroom techniques, or even actively avoid them.

This is due in part to the difficulty that teachers and learners alike have in understanding the changes that occur over the course of new learning; most lack an accurate mental model of memory processes (Pan & Bjork, 2020). While learning should not be reduced to memorisation, any new learning experiences do need to be retained in some way if they are later to be used (Kirschner, Sweller & Clark, 2006), and the relevance of those prior experiences also needs to be recognised and the details retrieved in order to facilitate the application of knowledge (Kaminske, Kuepper-Tetzel, Nebel, Sumeracki & Ryan, 2020). This retention and retrieval is based on human memory.

While the concept of memory draws heavily on positivist-influenced work in experimental psychology, it is also strongly influenced by constructivism in that the knowledge in a person's long-term memory is typically seen as being composed of interconnected 'schemas', each representing a discrete concept (Bransford et al., 2000). The long-term memory store is also widely seen as a dynamic system, with existing memories prone to being modified and reinterpreted through the lens of new experiences (Loftus, 2019). As psychologist Robert Bjork once put it, "*an item can seldom, if ever, be retrieved from memory without modifying the representation of that item in memory in significant ways*." (Bjork, 1975, p.123).

As such, the role of memory in learning is not as a passive store but rather a system which is subtle and fluid in its functioning. As well as existing memories being affected by new experiences, the reverse is true in as far as a learner creatively utilises relevant knowledge during their learning – a fundamental tenet of constructivism (Perkins, 2006). The context and timing of classroom experiences can therefore play an important role in how successfully these experiences bring about new conceptual understanding.

This paper focuses on two classroom techniques which can help in promoting successful learning via their effects on memory – spacing and interleaving. It also addresses the need for professional learning of these techniques to overcome popular misconceptions, and the potential of the techniques to be applied to practice-based teacher education itself, helping what is learned by teachers to likewise be durable and transferrable.

1.2 Desirable difficulties

Among memory-related approaches to teaching and learning, one group stands out as especially subject to misconceptions – the so-called 'desirable difficulties', manipulations which boost learning over the long term while impeding short-term performance (Bjork and Bjork, 2011). These techniques result in slower and/or more error-ridden performance during practice and to a greater subjective sense of challenge (and are therefore difficulties), but they are beneficial for memory durability and transfer (and as such, desirable, when learning is the priority).

Examples of techniques which bring about or entail desirable difficulties include self-testing (Roediger & Karpicke, 2006), active generation of material (Slamecka & Graf, 1978), or variations in the learning context/surroundings (Smith, Glenberg & Bjork, 1978). They

also include the spacing effect and interleaving, which are discussed in detail below (Section 2).

Bjork and Bjork (1992) explain the distinction between performance and learning through their 'new theory of disuse'. According to this theory, the *retrieval strength* of a memory (its accessibility at a given time) can be dissociated from its *storage strength* (how well it is stored in memory). Something can be well stored but hard to retrieve – for example, a childhood memory. Likewise, an item can be easily retrieved but not well stored – for example, something novel which was practised just a few minutes ago.

From this perspective, the teaching of students of any age (here I will use the term 'students' to refer to those in formal education, and 'learners' when referring to learning more broadly) should aim to boost storage strength of their memories. However, many classroom tasks feature massed repetition over the short term, impacting on *performance* but not *learning*. This raises a counterintuitive but important implication of the new theory of disuse: a period of forgetting can make later learning more effective, as retrieving and revising less accessible items boosts their storage strength, while more immediate repetition would only boost retrieval strength (Yan, Clark & Bjork, 2016). As a result, practice or re-study would best be delayed until information is on the point of being forgotten (Bjork, 2011).

Another implication of this theory is that a new trace in long-term memory is not fixed. Most obviously, any initial conceptual learning experience in the classroom will be subject to rapid and large scale forgetting over subsequent weeks, and cannot in itself be considered the end of the learning process, even if a formative assessment indicates that a student 'knows' the material. The classroom dynamics of learning and forgetting are better explained by this theory than by the better-known behaviourist and information-processing frameworks which emphasise short-term repetition.

It follows that simply changing the timing of when teaching or practice take place can affect how likely material is to be later successfully recalled and transferred by students, without requiring major modifications to the information itself or the task used, or any additional learning time (and indeed, less study time may be required overall if the timing is optimised). Accordingly, two techniques which focus on the schedule of learning experiences – spacing and interleaving – are among a group of educational interventions which Roediger and Pyc (2012, p. 242) describe as "low hanging fruit" – easily applied techniques with little outlay for the education system and a high chance of success, in comparison to more uncertain and costly alternatives. These techniques will be discussed next.

2. Techniques which modify item order and timing

2.1 Spacing

As highlighted above, teachers' efforts to promote effective learning may be undermined by the rapid and large-scale forgetting that follows. The *spacing effect* is a reliable and well-evidenced phenomenon which has been demonstrated in a wide range of contexts, including the learning of physical skills, facts, and conceptual knowledge (Cepeda, et al. 2006), although it should be noted that the bulk of the research consists of laboratory studies of adults. The essential finding is that information is more effectively encoded to long-term memory if it is practiced in a spaced (i.e. distributed) fashion than when practiced close together in time (usually referred to as a 'massed' presentation).

It appears to be important that items are well learned in the first study session. Rawson and Dunlosky (2011) found that during initial learning, students benefit from at least three successful attempts to retrieve concepts, after which re-study should occur at widely spaced

intervals. It is likely that if students only briefly look at a set of information and do not fully absorb or understand it, a long delay before a review will be unhelpful.

In a typical experiment into the effect, there is a gap between two study sessions which is either shorter or longer across the different experimental conditions. There is also a controlled retention interval between the end of the second study session in all conditions and a final test, meaning that the period of time during which forgetting could occur is kept constant. This is illustrated in Figure 1.

Condition 1, massed: Session 1 - Session 2 - - - - - - - - - - Test

Condition 2, spaced: Session 1 - - - - - - - - Session 2 - - - - - - - - - - Test

Figure 1**.** Typical design of an experiment into spacing. Each dash represents a time period which could be brief (e.g. 1 minute) or longer (e.g. 1 week).

Many of the early laboratory studies into spacing (e.g. Landauer & Bjork, 1978) involved short retention intervals, but there have since been field studies using delays relevant to the timescale of classroom learning. For example, in a field study of Canadian undergraduates, Kapler, Weston and Wiseheart (2015) followed a science lecture with a review delayed by either one day (condition 1) or eight days (condition 2), and a final test five weeks after the review; this study found superior performance in the eight-day delay condition.

It is difficult to put an exact time on the optimal duration for a delay, but in a study of trivia facts, Cepeda, Vul, Rohrer, Wixted and Pashler (2008) found that the benefits of spacing could extend to up to a year, and concluded that most educators are likely to be spacing too little, as the benefits of a larger delay outweighs the risk of increased forgetting. Importantly, they also concluded that the optimal delay depends on the retention interval; the longer a learner needs to remember the information, the more widely the initial study sessions

should be spaced. Therefore, assuming that it is an educational aim for students to retain information for a lengthy period after initial learning (ideally lifelong), then the delays between practice sessions should be as long as can be practically accommodated.

Another factor that may affect optimal spacing is the level of understanding that a student has gained. Items which are more deeply understood – i.e. those with better-developed links to existing schemas – are better remembered at a later point (Brainerd et al., 1990). This implies that forgetting has been slowed; from the point of view of the new theory of disuse, meaningful items remain more accessible for longer. As optimal spacing depends on the speed of forgetting, it is therefore meaning-dependent, too. Less inherently meaningful information (such as arbitrary vocabulary and terminology) may require more rapid follow-up. However, given that meaningful understanding is an interaction between the students and the material (as the extent to which something is meaningful to a student depends in part on their existing concept knowledge), optimal timing is a highly context-specific judgement for teachers.

2.2 Interleaving

Interleaving refers to the mixing of types of learning item – i.e. problems, images or examples – such that each item is preceded and followed by an item of a different type. It can be contrasted with a 'blocked' sequence, where several items of the same type are presented together. For example, in arithmetic, a set of four tasks each involving addition would be considered a blocked presentation, while four similar tasks each involving different skills (such as one addition, one subtraction, one multiplication and one division) would be considered to be interleaved. If an individual was learning about different species of birds,

interleaving would involve alternating or mixing examples of different species, while blocking would involve viewing several examples of the same species consecutively.

Figure 2 shows the typical way that interleaving and blocking are contrasted in an experimental design, keeping the overall items presented and the study time constant.

Condition A, blocked sequence: xxxxyyyyzzzz

Condition B, interleaved sequence: xyzxyzxyzxyz

Figure 2. The order of items in typical interleaved and blocked sequences.

The benefit of interleaving (sometimes termed the 'interleaving effect') has been found to apply to the practice of mathematics skills (e.g. Rohrer, Dedrick & Stershic, 2015), and to the learning of social science concepts (Rawson, Thomas & Jacoby, 2015), modern artists' styles (Kornell & Bjork, 2008), and animal species (Birnbaum, Kornell, Bjork & Bjork, 2013), to name but a few. As with other desirable difficulties, interleaving appears to make learning more difficult at first but more durable over the long term (Yan, Clark & Bjork, 2016).

There is a connection between interleaving and the spacing effect, because interleaving inevitably affects the delay between two items of the same type (in Figure 2, for example, instances of each item type are more widely spaced in Condition B than in Condition A). Indeed, when Kornell and Bjork (2008) rekindled interest in the concept with their laboratory study of interleaved learning of artists' styles, they initially attributed their findings to spacing. Since then, however, several studies have kept overall spacing of new examples constant while still comparing blocking versus interleaving – usually by inserting delays or filler items – and found that the benefit of interleaving persists (e.g. Kang & Pashler, 2012).

A key factor which results in this benefit is the increased opportunity for a learner to compare and contrast different types of item when they are presented consecutively.

Accordingly, delays in between interleaved items can reduce the benefit (Birnbaum et al., 2013), interleaving unrelated items is not beneficial (Hausman & Kornell, 2014), and the greatest benefit comes when the difference between items are subtle and thus likely to be missed if the items are presented in separate study sessions (Carvalho & Goldstone, 2014). This subtlety, however, will depend on a learner's state of knowledge; something which is obvious to an expert may be subtle to a novice. As such, this matter implies a need for judgement by the teacher in terms of where a particular class of students may benefit from items being contrasted.

Interleaving typically involves the inductive learning of a category. For example, Birnbaum et al. (2013) presented example images of bird and butterfly species followed by a test with novel images, while Eglington and Kang (2017) used a similar methodology when teaching categories based on types of hydrocarbon to American undergraduates. In such cases, it is through exposure to multiple examples that category learning occurs – and this happens best when the examples shown include items which both do and do not exemplify the category to be learned. Overall then, interleaving appears to be a valuable technique for developing transferrable concept knowledge through mixing of examples and problems.

2.3 Practical considerations

A major practical difference between spacing and interleaving is that the former tends to involve re-studying a given set of material, while interleaving tends to involve presenting several different examples of a category during an inductive learning phase and then presenting novel items during a test phase (instead of or as well as testing memory for the original items). This has educational implications; spacing is likely to be particularly beneficial when items studied are exactly the same, such as for the practice of foreign

language vocabulary or times tables. Interleaving, in contrast, stands to benefit not only memory but also transfer. It is likely to be useful in situations where students have to identify novel instances of previously studied concepts, such as identifying signs of glaciation in a previously unseen landscape, or recognising social psychology phenomena during an everyday encounter.

However, there is at least some evidence that spacing can also impact on transfer, perhaps because it supports the factual knowledge required for transfer to be successful. For example, in a study by Foot-Seymour, Foot and Wiseheart (2019), 9 to 12-year-old Canadian students were taught to judge the credibility of online sources either in three consecutive days (massed) or over three weeks (spaced). The spaced group later showed better factual retention but also better critical thinking than did the massed group. This study also highlights that although the evidence-base tends to draw on studies of university students (Cepeda et al., 2006), both spacing and interleaving have been demonstrated with younger learners (see Slone & Sandhofer, 2017; Vlach, Sandhofer & Bjork, 2014 for further examples).

Another practical consideration is that school systems in many jurisdictions tend to judge teacher performance via the performance of classes on test scores. Given that the techniques described benefit long-term learning rather than performance, there is a danger that spacing may be avoided because its benefits are slower to be realised; interleaving, in contrast, has been demonstrated over relatively short timescales. While shifting to longer-term measures of judging teacher effectiveness would be desirable, a more immediate fix would be to compress the spaced practice schedule used. The optimal delay between practice sessions depends on the timescale that information needs to be retained (Cepeda et al., 2008); teachers and classes who are working towards assessments over a timescale of days and weeks rather than years can still space their practice, but the optimal delays will be shorter.

As noted, the literature on both spacing and interleaving is dominated by laboratory studies of undergraduates, but many use highly authentic materials such as language vocabulary or science examples. Field studies have begun to show that the techniques have potential to apply to more complex, higher-order learning as well as to basic facts, and can generalise to school-aged children (Kapler et al., 2015; Foot-Seymour et al., 2019), though lower levels of background knowledge among younger students could impact on the optimal timing of spacing and on the types of contrast that would most benefit from interleaving.

Clearly, then, both techniques have potential to be applied to a range of learning experiences, particularly where there is institutional support in the form of systems that recognise that learning is best defined in terms of flexible and long-term competencies rather than performance on short-term tests. However, their use faces a further challenge – higher levels of subjective difficulty tends to prompt students to eschew such techniques, while teachers, too, may seek to avoid short-term difficulties. The next sections focus on the underpinnings of such misconceptions, and on ways to address them through professional learning.

3. Barriers due to learner and teacher metacognition

3.1 Student metacognition

When studying material with a view to learning or developing new concepts, students think not only about the material but also about their own thinking. They attempt to judge what they have learned, and make decisions about what to study next (Kornell & Bjork, 2007). The

processes involved draw on metacognition, a concept which can usefully be divided into two main areas when it comes to education (Nelson & Narens, 1994):

- Metacognitive knowledge: learned information about strategies that can be brought to bear on future learning.
- Metacognitive monitoring: focus on a particular learning task as it proceeds.

These two key strands of metacognition clarify why students and their teachers may avoid spacing and interleaving: due to mistakenly believing that they are less effective, and because of failing to recognise successful learning.

Metacognitive knowledge includes beliefs about learning and about one's own attainment level, and these are characterised by inaccuracies. The workings of human memory are not easy to figure through practice alone (Bjork, 2011), and do not fit with 'common sense' assumptions (Simons & Chabris, 2011), or with the way memory is typically portrayed in the media (Karpicke, 2016).

Metacognitive monitoring can also be inaccurate because introspection is a limited way of theorising about thinking. Indeed, students are generally poor at judging what they do and do not understand; tasks requiring 'judgement of learning' – a person's subjective anticipation of being able to recall the information at a later point in time – tend to be biased by the ease of retrieval, without making sufficient allowances for the context of this retrieval (Benjamin & Bjork, 1996).

Multiple laboratory studies of spacing (e.g. Zechmeister & Shaughnessy, 1980) and interleaving (e.g. Kornell & Bjork, 2008) have confirmed that many people mistakenly think that these strategies are less effective than the alternative (i.e. massing or blocking the same practice). Students also appear to favour sub-optimal learning strategies outside of the laboratory. For example, Hartwig and Dunlosky (2012) found that over 60% of their student sample favoured ineffective revision strategies such as re-reading, cramming and

underlining/highlighting, and that the choice of such strategies was statistically related to a poorer grade point average.

Accurate metacognition undoubtedly helps with evaluating one's own work and learning from mistakes in the right context, and as such it is seen as a potentially powerful education intervention (EEF, 2018). In the case of metacognitive knowledge, this could include teachers directly discussing metacognition with their classes, using accurate terminology and guiding them towards effective learning strategies (Pintrich, 2002). Of course, to do so, they would have to possess this knowledge themselves.

Timing appears to play a role in the accuracy of student metacognitive monitoring. A study by Nelson and Dunlosky (1991) compared judgements of learning for word pairs either immediately or after ten minutes. The ten-minute delay greatly improved the accuracy of learner judgements, which was little better than chance level when done immediately. This suggests that as well as its role in learning, spacing may improve metacognition. Indeed, Bahrick and Hall (2005) have suggested that this is one of the primary causes of the spacing effect.

Overall, then, there is evidence that students will tend to lack insight into the techniques under discussion, and there is potentially a very important role for teachers to guide them towards more effective learning behaviours. An important question, therefore, is whether teachers have sufficient expertise in these processes. The next section explores this issue.

3.2 Teacher metacognition

As argued at the start of this paper, effective learning depends upon the schedule of learning experiences; it therefore also depends on teachers making good choices about what material or tasks to present and when to present them. It may appear to stand to reason that most

teachers will have a thorough and accurate conception of how memory works to guide these choices. However, the points made above about flawed metacognition need not apply only to students. Teachers, too, may harbour beliefs about the learning processes which fail to match the evidence.

Misconceptions about the mind and learning are widespread among adults. Furnham (2018) questioned adult participants about popular myths from developmental and neuropsychology such as 'people suffering from amnesia typically cannot recall their own name or identity', and found that a large number were rated as definitely or partly true across all demographic groups. Simons and Chabris (2011) found large differences between answers from the general public and those of memory researchers on questions about memory; 63% of members of the public agreed or strongly agreed with the idea that memory works like a video camera, while 48% agreed that once you have experienced an event and formed a memory of it, that memory does not change. None of the researchers surveyed endorsed either belief.

Clearly, professionals differ from members of the public in important ways, and have been given specific training to carry out their role. However, there is considerable variability in what is taught to new teachers (Carter, 2015), meaning that there is no guarantee that what they have been told about learning will be accurate. Guides for new teachers often fail to include information on strategies such as spacing (Surma, Vanhoyweghen, Camp & Kirschner, 2018), and school textbooks tend not to feature interleaved tasks (Rohrer, Dedrick & Hartwig, 2020), meaning that such approaches are unlikely to be seen as standard practice. Meanwhile the popularity of neuromyths (Howard-Jones, 2014) – over 97% of teachers surveyed in Turkey subscribe to the idea that people learn better when taught according to their preferred 'learning style', while 91% of UK teachers agreed that differences between

students can be explained by their being ‘left brained’ or ‘right brained’ – shows that fads and misconceptions can spread easily.

Firth (2018) has found that despite having more accurate views of memory overall than those found among the public by Simon and Chabris (2011), teachers subscribed to a number of misconceptions. In particular, those studied were highly out of line with the scientific consensus on desirable difficulties such as the spacing effect, with only 12.2% agreeing that ‘as a teacher, it is wise to wait until learners have almost forgotten things before you go over them again’.

There are also reasons to doubt that experience alone will correct misconceptions about memory. In the same study by Firth (2018), the extent of misconceptions did not vary with experience; teachers within the first five years of their career showed the same level of misconceptions as those with 20 or more years of experience. Similarly, Morehead et al. (2016) found no relationship between teachers’ endorsement of myths such as learning styles and their number of years of teaching experience.

Indeed, despite experience being near synonymous with skill in some contexts, there is not a strong statistical relationship between a teacher’s years of classroom experience and student outcomes in their classes, in part because improvements tend to plateau within five years (Rivkin et al., 2005), with further experience beyond this point having a negligible impact on classroom performance. This fits with what is known from other domains; as Ericsson, Prietula and Cokely (2007) note in relation to sports expertise, practice alone can lead to the consolidation and automation of mediocre performance.

It therefore appears that simply accumulating more time in the classroom is likely to be insufficient to develop accurate metacognition regarding how memory works. This strongly suggests that teachers will not spontaneously develop metacognitive accuracy and come to

adopt evidence-based strategies such as spacing and interleaving – exactly as is the case with students (e.g. Rohrer & Pashler, 2010).

The next section of this article looks at some possible approaches to tackling this gap in professional learning.

4. Implications for professional learning

4.1 The role of evidence-based practice

The techniques discussed highlight a disparity between research evidence and popular assumptions about learning and memory that is likely to affect both students and teachers. An obvious interim conclusion is that teachers – both new trainees and in-service professionals – should engage more with research evidence, boosting their students' learning by using and recommending techniques such as spacing and interleaving.

However, as discussed earlier (see Section 2), applications of such techniques are highly context-specific. A student's existing knowledge impacts on how and when techniques should be used, and this interaction is specific to the learner and to the material. The application and use of spacing and interleaving therefore depends on judgements made in a particular classroom, even if the principles do hold more generally.

The implication of this is that teachers will need to build in delays between practice sessions (sometimes by taking existing tasks and dividing them across different sessions), but delays alone are not enough – there must be practice before and after the delay. For any practice task, students should continue until they have achieved an initial mastery of the material. Further practice at this point (i.e. overlearning) is inefficient, and should instead

occur after a delay (Rohrer & Taylor, 2006). As learning becomes more secure, the delays between separate practice sessions can increase.

When it comes to interleaving, the key factor is contrast. Again, teachers' judgement is important, as they are in the best position to know which concepts are likely to cause confusion. Such concepts should be illustrated with multiple interleaved examples, rather than being studied in separate sessions.

Teachers will also need to build metacognitive knowledge among their students, and manage any motivational impact of techniques which increase the difficulty of short-term practice. Understanding of learning and memory processes could be seen, then, as a part of the teacher's professional toolkit, rather than as something that can be imposed from above (Firth, 2017) – a fact that may assuage concerns about the potential of evidence-based practice to interfere with a teacher's professional judgement (e.g. see Biesta, 2007).

4.2 Professional learning

For in-service teachers, how exactly should professional learning about spacing and interleaving (or related desirable difficulties) take place? Unfortunately, simply providing teachers with information is likely to have only a limited impact (Wiliam, 2010). A useful direction comes from a study by Will, Masad, Vlach and Kendeou (2019) which featured common false beliefs such as 'reading in dim light causes nearsightedness' (p. 113). All participants were given information that contradicted the myths, but only one condition featured a direct refutation; participants in this condition later showed more accuracy and less circularity in their written explanations. A classroom implication would be that that simply presenting teachers with information about the evidence base may be insufficient if the

explanation does not also directly highlight areas of current practice which are suboptimal, with alternatives directly contrasted.

Just as classroom learning is more effective when it is spaced rather than massed, professional learning is more effective when it is sustained over a period of time; lessons from one-off CPD events tend to be rapidly forgotten upon returning to the classroom (Donaldson, 2011). Instead, teachers need opportunities to put their new learning to work in the classroom, perhaps gradually and with some scaffolding of the process. They may also be able to acquire naturally-occurring classroom evidence, for example class test scores, to gain a further insight in a process of ongoing reflection and enquiry (Wall & Hall, 2019). Extended professional learning also provides more opportunities to integrate new ideas into classroom practice (Darling-Hammond, Hyler, & Gardner, 2017), thus supporting the process of transfer and application of new knowledge.

One model of spaced professional learning is to engage in practitioner research projects, as such projects are naturally distributed over a period of time. A teacher research project could also engage curiosity, and be highly motivating. While this point again applies mainly to in-service teachers, a similar format can be used to ensure that learning from initial teacher education is spaced out over time, to minimise forgetting.

Interleaving, too, can be applied to professional learning. As discussed, interleaved presentations provide opportunities for students to compare and contrast concepts, thus allowing more accurate categorisation and the development of new categories, and is particularly valuable if differences are subtle and easily missed. One practical example of how this could occur is through brief presentations of classroom learning examples by multiple colleagues, for example at staff meetings. Such formats would promote teachers' ability to compare and contrast varied examples, rather than spending an entire session learning about a single technique.

There is also the potential for increased interleaving through the mixing of real observed classroom experiences. For example, viewing multiple short video clips of lessons may be preferable, in terms of concept and skill learning, to infrequent observations of entire lessons. To increase contrast, such clips should be as short as possible, and presented side-by-side without intervening delays (Birnbaum et al., 2013).

On the basis of the research evidence into interleaving (which tends to feature novel categories), there is no reason that the same principle could not be applied to initial teacher education as well. Indeed, interleaving of classroom video clips could be an excellent way of raising trainees' awareness of contrasts that are subtle and hard to notice, even before they have had the opportunity to spend time in real classrooms. It's worth noting again that the ease of perceiving a contrast will depend on the individual's prior learning; as teachers become more skilled, they will become better able to learn from subtle contrasts.

Overall, just as information does not come to us neatly categorised in the real world, so the professional learning process may benefit from increased variability and more delays, boosting teachers' ability to retain and transfer new knowledge.

5. Concluding comments

The evidence presented here explores the benefits of spacing and interleaving – both easy, quick and cheap techniques to apply. While much of the research on the two techniques is founded on theories of human long-term memory, it is important to recognise that modern cognitive psychology sees memory as a flexible and constantly-changing system which is highly sensitive to context and which provides the foundation for meaningful understanding. New experiences, in this view, should be encoded into a learner's memory in a way that best activates and links to prior knowledge. This process is sensitive both to timing and to the

opportunities available for a learner to perceive contrast between one task or example and another. Improvements to the schedule of classroom tasks can therefore boost students' ability to apply the concepts they study to new situations (Kaminske et al., 2020).

However, metacognitive research strongly suggests that we cannot rely on either students or their teachers to have accurate beliefs about the optimal timing or order of learning tasks. Those involved in supporting professional learning should be aware that many of the same myths and misconceptions present among the general population can be seen in the teaching profession, and experience alone appears insufficient to counter this problem. In addition, in-service professional learning that is interleaved and spaced out over time is more likely to be impactful.

As has been seen, much of the evidence on these techniques draws on laboratory studies of adults. Further research to guide their application to classroom practice could include a field research programme that assesses the use of spacing or interleaving across different topics, using previous years' attainment as baseline data. It would also be helpful to carry out qualitative research to explore teachers' views and judgements, for example by presenting interviewees with example tasks and asking them to talk through their views on the best schedules to use.

Finally, if preparation of new teachers were to include an overview of relevant areas of memory and misconceptions such as the ones identified above, this would increase awareness of pitfalls such as judging learning on the basis of short-term performance and students' tendency to metacognitive inaccuracy.

References

Bahrick, Harry P., and Lynda K. Hall. 2005. “The Importance of Retrieval Failures to Long-Term Retention: A Metacognitive Explanation of the Spacing Effect.” *Journal of Memory and Language* 52 (4): 566–577. doi: 10.1016/j.jml.2005.01.012

Benjamin, Aaron S., and Robert A. Bjork. 1996. “Retrieval Fluency as a Metacognitive Index.” In *Implicit Memory and Metacognition*, edited by Lynn M. Reder, 309–338. Mahwah, NJ: Erlbaum.

Biesta, Gert. 2007. “Why “What Works” Won’t Work: Evidence-based Practice and the Democratic Deficit in Educational Research.” *Educational Theory* 57 (1): 1–22. doi: 10.1111/j.1741-5446.2006.00241.x

Birnbaum, Monica S., Nate Kornell, Elizabeth Ligon Bjork, and Robert A. Bjork. 2013. “Why Interleaving Enhances Inductive Learning: The Roles of Discrimination and Retrieval. *Memory & Cognition* 41 (3): 392–402. doi: 10.3758/s13421-012-0272-7

Bjork, Elizabeth Ligon, and Robert A. Bjork. 2011. “Making Things Hard on Yourself, but in a Good Way: Creating Desirable Difficulties to Enhance Learning.” In *Psychology and The Real World: Essays Illustrating Fundamental Contributions to Society*, edited by Morton A. Gernsbacher, Richard W. Pew, Leatta M. Hough, and James R. Pomeranz, 56–64. New York, NY: Worth.

Bjork, Robert A. 1975. “Retrieval as a Memory Modifier.” In *Information Processing and Cognition: The Loyola Symposium*, edited by Robert L. Solso, 123–144. Hillside, N.J.: Lawrence Erlbaum.

Bjork, Robert A. 2011. “On the Symbiosis of Remembering, Forgetting, and Learning.” In *Successful Remembering and Successful Forgetting: A Festschrift in Honor of Robert A. Bjork*, edited by Aaron S. Benjamin, 1–22. New York, NY: Psychology Press.

Bjork, Robert A., and Elizabeth Ligon Bjork. 1992. “A New Theory of Disuse and an Old Theory of Stimulus Fluctuation.” In *From Learning Theory to Connectionist Theory: Essays in Honor of William K. Estes*, edited by Alice F. Healy, Stephen M. Kosslyn and Richard M. Shiffrin, 35–67. Hillsdale, NJ: Lawrence Erlbaum.

Brainerd, C. J., V. F. Reyna, M. L. Howe, J. Kingma, and Robert F. Guttentag. 1990. “The Development of Forgetting and Reminiscence.” *Monographs of the Society for Research in Child Development* 55 (3/4): i+iii+v-vi+1–109.

Bransford, John D., Anne L. Brown, and Rodney R. Cocking, eds. 2000. *How People Learn: Brain, Mind, Experience and School*. Washington, DC: National Academy Press.

Carter, David. 2015. *Carter Review of Initial Teacher Training* (ITT). London, UK: Department for Education.

Carvalho, Paulo F., and Robert L. Goldstone. 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. doi: 10.3758/s13421-013-0371-0

Cepeda, Nicholas J., Harold Pashler, Edward Vul, John T. Wixted, and Doug Rohrer. 2006. “Distributed Practice in Verbal Recall Tasks: A Review and Quantitative Synthesis.” *Psychological Bulletin* 132 (3): 354–380. doi: 10.1037/0033-2909.132.3.354

Cepeda, Nicholas J., Edward Vul, Doug Rohrer, John T. Wixted, and Harold Pashler. 2008. “Spacing Effects in Learning: A Temporal Ridgeline of Optimal Retention.” *Psychological Science* 19 (11): 1095–1102. doi: 10.1111/j.1467-9280.2008.02209.x

Darling-Hammond, Linda, Maria E. Hyler, and Madelyn Gardner. 2017. “*Effective Teacher Professional Development*.” Learning Policy Institute. https://learningpolicyinstitute.org/sites/default/files/product-files/Effective_Teacher_Professional_Development_REPORT.pdf

Donaldson, G. 2011. *Teaching Scotland's Future: Report of a Review of Teacher Education in Scotland*. Edinburgh, UK: Scottish Government.

EEF [Education Endowment Foundation]. 2018. "*Metacognition and Self-regulated Learning: Guidance Report.*" EEF Website. https://educationendowmentfoundation.org.uk/public/files/Support/Publications/Metacognition/EEF_Metacognition_and_self-regulated_learning.pdf

Eglington, Luke G., and Sean H. K. Kang. 2017. "Interleaved Presentation Benefits Science Category Learning." *Journal of Applied Research in Memory and Cognition* 6 (4): 475–485. doi: 10.1016/j.jarmac.2017.07.005

Ericsson, K. Anders, Michael J. Prietula, and Edward T. Cokely. 2007. "The Making of an Expert." *Harvard Business Review* 85, (7/8): 114–21.

Firth, Jonathan. 2017. "Experts in Learning." In *Flip the System UK: A Teachers' Manifesto,* edited by Lucy Rycroft-Smith and J. L. Dutaut, 20–28. Abingdon, Oxon: Routledge.

Firth, Jonathan. 2018. "Teachers' Beliefs about Memory: What are the Implications for In-Service Teacher Education?". *Psychology of Education Review* 42 (2), 15–22.

Foot-Seymour, Vanessa, June Foot, and Melody Wiseheart. 2019. "Judging Credibility: Can Spaced Lessons Help Students Think More Critically Online?." *Applied Cognitive Psychology* 33 (6): 1032–1043. doi: 10.1002/acp.3539

Furnham, Adrian. 2018. "Myths and Misconceptions in Developmental and Neuro-Psychology." *Psychology* 9 (2): 249–259. doi: 10.4236/psych.2018.92016

Hartwig, Marissa K., and John Dunlosky. 2012. "Study Strategies of College Students: Are Self-testing and Scheduling Related to Achievement?." *Psychonomic Bulletin & Review* 19 (1): 126–134. doi: 10.3758/s13423-011-0181-y

Hausman, Hannah, and Nate Kornell. 2014. “Mixing Topics While Studying Does Not Enhance Learning.” *Journal of Applied Research in Memory and Cognition* 3 (3): 153–160. doi: 10.1016/j.jarmac.2014.03.003

Howard-Jones, Paul A. 2014. “Neuroscience and Education: Myths and Messages.” *Nature Reviews Neuroscience* 15 (12): 817–824. doi: 10.1038/nrn3817

Kaminske, Althea N., Carolina Kuepper-Tetzel, Cindy Nebel, Megan Sumeracki, and Sean Ryan. 2020. “Transfer: A Review for Biology and the Life Sciences.” *CBE—Life Sciences Education* 19 (3): es9. doi: 10.1187/cbe.19-11-0227

Kang, Sean H. K., and Harold Pashler. 2012. “Learning Painting Styles: Spacing is Advantageous when it Promotes Discriminative Contrast.” *Applied Cognitive Psychology* 26 (1): 97–103. doi: 10.1002/acp.1801

Kapler, Irina V., Tina Weston, and Melody Wiseheart. 2015. “Spacing in a Simulated Undergraduate Classroom: Long-term Benefits for Factual and Higher-level Learning." *Learning and Instruction* 36: 38–45. doi: 10.1016/j.learninstruc.2014.11.001

Karpicke, Jeffrey D. 2016. “A Powerful Way to Improve Learning and Memory: Practicing Retrieval Enhances Long-term, Meaningful Learning.” *Psychological Science Agenda.* http://www.apa.org/science/about/psa/2016/06/learning-memory.aspx

Kirschner, Paul A., John Sweller, & Richard E. Clark. 2006. “Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching.” *Educational Psychologist* 41 (2): 75–86. doi: 10.1207/s15326985ep4102_1

Kornell, Nate, and Robert A. Bjork. 2007. “The Promise and Perils of Self-regulated Study.” *Psychonomic Bulletin & Review* 14 (2): 219–224. doi: 10.3758/BF03194055

Kornell, Nate, and Robert A. Bjork. 2008. “Learning Concepts and Categories: Is Spacing the ‘Enemy of Induction’?.” *Psychological Science* 19 (6): 585–592. doi: 10.1111/j.1467-9280.2008.02127.x

Landauer, Thomas K., and Robert A. Bjork. 1978. “Optimum Rehearsal Patterns and Name Learning. In *Practical Aspects of Memory,* edited by Michael M. Gruneberg, Peter E. Morris and Robert N. Sykes, 625–632. London: Academic Press.

Loftus, Elizabeth F. 2019. “Eyewitness Testimony.” *Applied Cognitive Psychology* 33 (4): 498–503. doi: 10.1002/acp.3542

Morehead, Kayla, Matthew G. Rhodes, and Sarah DeLozier. 2016. “Instructor and Student Knowledge of Study Strategies.” *Memory* 24 (2): 257–271. doi: 10.1080/09658211.2014.1001992

Nelson, Thomas O., and John Dunlosky. 1991. “When People's Judgments of Learning (JOLs) are Extremely Accurate at Predicting Subsequent Recall: The “Delayed-JOL Effect.” *Psychological Science* 2 (4): 267–271. doi: 10.1111/j.1467-9280.1991.tb00147.x

Nelson, Thomas O., and Louis Narens. 1994. “Why investigate metacognition?” In *Metacognition: Knowing about Knowing*, edited by Janet Metcalfe and Arthur P. Shimamura, 1–25. Cambridge, MA: MIT Press.

Pan, Steven C., and Robert A. Bjork. 2020. “Acquiring an Accurate Mental Model of Learning: Towards an Owner’s Manual.” In *Oxford Handbook of Learning & Memory: Foundations and Applications*, edited by Anthony Wagner and Michael Kahana. Oxford: Oxford University Press.

Pintrich, Paul R. 2002. “The Role of Metacognitive Knowledge in Learning, Teaching, and Assessing.” *Theory into Practice* 41 (4): 219–225. doi: 10.1207/s15430421tip4104_3

Rawson, Katherine A., and John Dunlosky. 2011. "Optimizing Schedules of Retrieval Practice for Durable and Efficient Learning: How Much is Enough?." *Journal of Experimental Psychology: General* 140 (3): 283–302. doi: 10.1037/a0023956

Rawson, Katherine A., Ruthann C. Thomas, and Larry L. Jacoby. 2015. "The Power of Examples: Illustrative Examples Enhance Conceptual Learning of Declarative Concepts." *Educational Psychology Review* 27: 483–504. doi: 10.1007/s10648-014-9273-3

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica* 73 (2): 417–458. doi: doi.org/10.1111/j.1468-0262.2005.00584.x

Roediger, Henry L., and Jeffrey D. Karpicke. 2006. "Test-enhanced Learning: Taking Memory Tests Improves Long-term Retention. *Psychological Science* 17 (3): 249–255. doi: 10.1111/j.1467-9280.2006.01693.x

Roediger, Henry L., and Mary Pyc. 2012. "Inexpensive Techniques to Improve Education: Applying Cognitive Psychology to Enhance Educational Practice." *Journal of Applied Research in Memory and Cognition* 1 (4): 242–248. doi: 10.1016/j.jarmac.2012.09.002

Rohrer, Doug, Robert F. Dedrick, and Marissa K. Hartwig. 2020. "The Scarcity of Interleaved Practice in Mathematics Textbooks." *Educational Psychology Review*. doi: 10.1007/s10648-020-09516-2

Rohrer, Doug, Robert F. Dedrick, and Sandra Stershic. 2015. "Interleaved Practice Improves Mathematics Learning." *Journal of Educational Psychology* 107 (3): 900–908. doi: 10.1037/edu0000001

Rohrer, Doug, and Harold Pashler. 2010. "Recent Research on Human Learning Challenges Conventional Instructional Strategies. *Educational Researcher* 39 (5): 406–412. doi: 10.3102/0013189X10374770

Rohrer, Doug, and Kelli Taylor. 2006. “The Effects of Overlearning and Distributed Practice on the Retention of Mathematics Knowledge.” *Applied Cognitive Psychology* 20: 1209–1224. doi: 10.1002/acp.1266

Simons, Daniel J., and Christopher F. Chabris. 2011. “What People Believe about how Memory Works: A Representative Survey of the US Population.” *PloS one* 6 (8): e22757. doi: 10.1371/journal.pone.0022757

Slamecka, Norman J., and Peter Graf. 1978. “The Generation Effect: Delineation of a Phenomenon.” *Journal of Experimental Psychology: Human Learning and Memory* 4 (6): 592–604. doi: 10.1037/0278-7393.4.6.592

Slone, Lauren K., and Catherine M. Sandhofer. 2017. “Consider the Category: The Effect of Spacing Depends on Individual Learning Histories." *Journal of Experimental Child Psychology* 159: 34–49. doi: 10.1016/j.jecp.2017.01.010

Smith, Steven M., Arthur Glenberg, and Robert A. Bjork. 1978. “Environmental Context and Human Memory.” *Memory & Cognition* 6 (4): 342–353. doi: 10.3758/BF03197465

Soderstrom, Nicholas C., and Robert A. Bjork. 2015. “Learning Versus Performance: An Integrative Review.” *Perspectives on Psychological Science* 10 (2): 176–199. doi: 10.1177/1745691615569000

Surma, Tim, Kristel Vanhoyweghen, Gino Camp, and Paul A. Kirschner. 2018. “The Coverage of Distributed Practice and Retrieval Practice in Flemish and Dutch Teacher Education Textbooks. *Teaching and Teacher Education* 74: 229–237. doi: 10.1016/j.tate.2018.05.007

Vlach, Haley A., Catherine M. Sandhofer, and Robert A. Bjork. 2014. “Equal Spacing and Expanding Schedules in Children’s Categorization and Generalization.” *Journal of Experimental Child Psychology* 123 (2): 129–137. doi: 10.1016/j.jecp.2014.01.004

Wall, Kate, and Elaine Hall. 2019. *Research Methods for Understanding Professional Learning*. London: Bloomsbury.

Wiliam, D. 2010. "*Teacher Quality: Why It Matters, and How to Get More of It."* The Spectator, March. http://www.dylanwiliam.org/Dylan_Wiliams_website/Papers_files/Spectator%20talk.doc

Will, Kelsey K., Andalusia Masad, Haley A. Vlach, and Panayiota Kendeou. 2019. "The Effects of Refutation Texts on Generating Explanations." *Learning and Individual Differences* 69: 108–115. doi: 10.1016/j.lindif.2018.12.002

Yan, Veronica X., Courtney M. Clark, and Robert A. Bjork. 2016. "Memory and Metamemory Considerations in the Instruction of Human Beings Revisited: Implications for Optimizing Online Learning." In *From the Laboratory to the Classroom,* edited by Jared Cooney Horvath, Jason M. Lodge and John Hattie, 73–90. Abingdon: Routledge.

Zechmeister, Eugene B., and John J. Shaughnessy. 1980. "When You Know That You Know and When You Think That You Know but You Don't." *Bulletin of the Psychonomic Society* 15 (1): 41–44. doi: 10.3758/BF03329756