Title Page w/ ALL Author Contact Info.

On rest-from-deliberate practice as a mechanism for the spacing effect: Commentary on Chen et al. (2021)

Faria Sana*[1], Veronica X. Yan[2], Paulo F. Carvalho[3]

[1] Athabasca University

[2] The University of Texas at Austin

[3] Carnegie Mellon University

Author Note

Faria Sana: https://orcid.org/0000-0002-2202-7592

Veronica X. Yan: https://orcid.org/0000-0002-3988-3184

Paulo F. Carvalho: https://orcid.org/0000-0002-0449-3733

All authors contributed equally to this work and share co-first authorship. Preparation of this manuscript was supported by National Science Foundation grant # BCS-1824257 to PFC and Social Science and Humanities Research Council Insight grant #435-2021-0426 to FS and VXY. The authors declare no conflict of interest.

* Correspondence concerning this article should be addressed to Faria Sana (fsana@athabascau.ca), Center for Social Sciences, Athabasca University, 10011 109 Street Northwest, #1200, Edmonton, Alberta T5J 3S8, Canada.

Abstract

In their recent paper, Chen et al. (2021) propose that rest periods between deliberate learning characterize the spacing effect and the alternation between skills without rest characterizes the interleaving effect. In this commentary, we show that this theory is inadequate in two aspects. First, the operationalization of their constructs are problematic—their mechanism of rest-from-deliberate-learning mismatches their operationalization (e.g., they code deliberate learning activities that should not allow for working memory recovery as rest-from-deliberate-learning), and their definition of whether stimuli require discriminative contrast appears to depend on the study outcome. Second, their systematic review neglects a large body of literature that is incompatible with their theory. For example, they neglect classic spacing studies on vocabulary learning, and their theory of spacing effects as being a result of working memory recovery cannot account for lag effects or interactions found in the literature. We conclude that there are almost certainly mechanistic differences between spacing and interleaving effects, but the mechanisms are likely not mutually exclusive, as defined by Chen and colleagues.

169 words

Keywords: spaced practice; interleaved practice; discriminative contrast hypothesis; rest-from-deliberate-learning

On rest-from-deliberate-learning as a mechanism for the spacing effect: Commentary on Chen et al. (2021)

In a recent *Educational Psychology Review* article, Chen and colleagues (2021; henceforth “CPS”) argued that spacing and interleaving are undergirded by two distinct mechanisms. They define spaced practice as “periods of mental activity [that] are alternated with periods of mental rest-from-deliberate-learning” (hypothesis 1; p. 2). CPS propose that learning depletes working memory resources and rest-from-deliberate-learning allows these mental resources to recover, yielding the spacing effect. In contrast, they define interleaved practice as “periods of practicing one skill are alternated with periods of practicing a different skill” (hypothesis 2; p. 2). CPS propose that the benefits of interleaving are the result of contrastive processing between the different skills. The authors argue that support for their hypotheses comes from studies involving rest-from-deliberate-learning and studies involving the need to discriminate in the absence of rest-from-deliberate-learning. CPS claim that there is no benefit of spacing or interleaving if there is neither rest-from-deliberate-learning nor the need to discriminate.

CPS seem to recognize the different theories on the spacing and interleaving effects, assert that they differ on various dimensions, and seek to distinguish them through their proposed theory and evidence based on study classifications. Critically, by separating the two phenomena, CPS set up a situation where one needs to consider the evidence for spacing and its mechanisms, and the evidence for interleaving and its mechanisms. In our view, the rest-from-deliberate-learning theory as stated cannot effectively discriminate between the two phenomena and because of that cannot account for all of the evidence presented.

We will focus on two critical aspects of this issue: (1) problematic operationalization: major constructs such as rest and concept relatedness are not well defined, no explicit constraints of these constructs are stated, and consequently the theory does not generate testable predictions; and (2) selective use of evidence: the evidence in support of their theory consists of selectively describing results that are compatible with their theory, neglecting a large body of literature that is not compatible with their theory.

Problematic Operationalization of Key Constructs

At the core of CPS's proposal is the argument that the spacing effect is a cognitive load effect: that massed practice is load-intensive, and that spacing (rest-from-deliberate-learning) allows for the recovery of working memory resources. Conversely, the interleaving effect is the result of alternating practice of related concepts that require discrimination *in the absence of rest-from-deliberate-learning*. Critical to these hypotheses is how one defines rest-from-deliberate-learning and what concepts are considered to be related or not. CPS propose that spacing should only be considered to include rest-from-deliberate-learning if the study included rests for sleeping, play, or incidental learning activities (p. 4). However, CPS's operationalization of what constitutes rest-from-deliberate-learning is unclear and inconsistent. Similarly, what makes concepts similar or dissimilar (and thus requiring contrast or not) is also unclear and inconsistent.

First, the operationalization of rest-from-deliberate-learning does not seem to match the alleged mechanism. If the mechanism is working memory recovery, then the rest should allow working memory to recover rather than occupying learners with other tasks that use up working memory resources. Moreover, in systematic review, they coded activities as rest-from-deliberate-learning even when participants were still focused on learning content (i.e. intentional, not incidental learning), as long as that content was not related to the target content. For example, CPS include the studies from Young and colleagues (2019) as evidence in favor of their spacing-as-deliberate-learning hypothesis. The spacing intervals in these studies were filled with irrelevant text passages that participants were *instructed to encode for later retrieval*. CPS argue that these text passages were irrelevant to the target content of learning (paired associates) and therefore serve as rest-from-deliberate-learning, but it remains to be explained how the intentional study of text passages would not constitute deliberate learning, deplete working memory resources, and as such be considered *rest*.

Second, CPS suggest that words from the same language are inherently similar and necessitate discriminative contrast. As such, CPS discount the verbal learning studies that show benefits of spacing, such as the majority of the 317 experiments included in the meta-analysis by Cepeda et al. (2006) and Ebbinghaus' (1885/1964) seminal work. Yet, when describing studies that fail to find either a benefit of spacing or interleaving (Carpenter & Mueller, 2013; Ostrow & Heffernan, 2015), CPS simultaneously argue that these are cases where items from the same domain do not need discriminative contrast, but also do not constitute rest-from-deliberate-learning. Carpenter and Mueller (2013) found that interleaving French words with different endings (e.g. —eau, —ou, —is) does not benefit learning of pronunciation rules. CPS argued that this finding did not test for discrimination because the words "*were easily distinguishable by the use of different rules associating the word pairs or by their appearance*." (p.16). In our view, it seems inconsistent to argue that learning pairs in one's native language (e.g., "apple-candy", "table-chair") requires discriminative contrast while learning how to pronounce foreign words (e.g., "bateau", "genou") does not. Simultaneously, CPS argue that although learning *language* requires discriminative contrast, learning different types of *math* problems does not. CPS point to Ostrow and Heffernan (2015) as evidence: in their study, participants who practiced angles, surface area, and probability problems in an interleaved manner did not perform better on a final test than did those who practiced the problems in a blocked manner, Hedge's g = 0.22. What CPS omit, however, is that there was in fact a large interleaving benefit for low-skilled students, Hedge's g = 0.60.

Selective Use of Evidence

CPS's theory is rooted in theoretical interpretation of a systematic review of the literature. CPS argue that a majority of the evidence is consistent with their proposal and that the literature includes only a few negative examples that add support to their theory. They overlook evidence that is inconsistent with their theories or cannot be accounted for by their proposed mechanisms, from choosing to focus on only the parts of published studies that fit their theory to ignoring an entire body of literature on spacing effects in verbal learning. First, CPS overlook evidence that is inconsistent with their proposal of interleaving only being helpful if discrimination is required. For example, only half of the data from Foster et al. (2019) is reported in CPS's analyses. Foster et al. asked participants to practice four different types of mathematics problems—these problems were practiced with no rest in between each one, and the order of the problem types was either blocked or interleaved. Critically, they manipulated whether the four types of problems were similar (e.g., volumes of different geometric shapes) or dissimilar (e.g., wedge volume, exponent division, fraction addition, permutations). When participants studied four similar problem types, they found a large interleaving benefit, d = .62. CPS report this result as evidence of the discrimination mechanism (see Table 2 in Chen et al., 2021). What CPS omit, however, is that the interleaving benefit was larger with the dissimilar set of mathematics problems (d = 1.00). These dissimilar math problems were deliberately learned without rest, thus the results cannot be accounted for by a rest-from-deliberate-learning mechanism. At the same time, the concepts were dissimilar to other problems learned in the same session, and thus such an effect cannot be explained by the discriminative contract mechanism. Thus, the mechanism that led to this greater benefit is therefore a puzzle for CPS's theory.

Second, CPS's review ignores most of the evidentiary basis for the spacing and interleaving effects. Large swaths of spacing effect research have been conducted using verbal learning paradigms, in which the intervals between repetitions of a given item are filled with presentations of the other to-be-learned items. That is, the participants typically experience no rest from deliberate learning. However, they also do not fit the criteria for interleaving, as there is no reason to think that the learning of unrelated items would be benefited by discriminative contrast. Importantly, robust spacing effects are found even though there is neither rest-from-deliberate learning nor need for discrimination. For example, Schwartz (1975) compared retention of bigrams (e.g., AR-LE) following massed and spaced practice. In both conditions participants were shown two presentations of each bigram either consecutively (massed) or spaced by presentation of other bigrams (spaced). Schwartz (1975; and many others following a similar paradigm) shows a benefit for spaced practice. But because there is no rest-from-deliberate-learning or need to discriminate the bigrams, it is unclear how the hypotheses proposed by CPS would account for these results: are these not spacing effects? There is similar evidence with other materials (e.g., nonsense syllables, words, paired associates, pictures) that are studied twice at varying intervals (Cepeda et al., 2006). Given CPS's inclusion criteria, none of these studies is part of their analyses. In fact, CPS reported only 48 studies that could be classified as spacing and 67 studies that could be classified as interleaving. These counts are in sharp contrast to 317 experiments located in 184 articles reported in a now 15 year old meta-analyses of the spacing effect (Cepeda et al., 2006), and the 59 studies reported in a recent meta-analysis of the interleaving effect (Brunmair & Richter, 2019). In our view, a theory of spacing that cannot account for the large majority of the evidence in the literature falls short.

Third, although CPS's theory of spacing can account for effects that are found when spacing is compared to massing (no-spacing), it has trouble accounting for spacing effects that are found when shorter and longer intervals between repetitions are compared (i.e., lag effects). Given how quickly working memory can recover (only 10 to 15 seconds, unless the information is actively being attended to or rehearsed; Peterson & Peterson, 1959), taking a 10 min rest in between repetitions should be no different than a one day rest, which in turn should be no different than a one week rest. In fact, CPS highlight that Vlach et al. (2008) found a very large spacing effect with a rest of only 30 to 50 sec. Hence, CPS's theory should not predict a difference between spacing intervals of one day versus multiple days, weeks, or months. And yet, spacing research has also found larger benefits of spacing when those repetitions are one week apart compared to one day apart (e.g., Kapler et al., 2015), and when those repetitions are months apart compared to weeks apart (e.g., Carpenter et al., 2009). In fact, Cepeda et al (2008) found that although a spacing effect was evident at multiple lag durations (the duration between repetitions), it followed an inverted U shaped curve, suggesting that more spacing is not always best for learning; the optimal lag duration depended on the duration of the retention interval (i.e., the time between last study and test moments). A rest-from-deliberate-learning account of the spacing effect cannot account for either lag effects or non-monotonic effects.

Fourth, it is unclear how CPS's theory would account for well-documented interactions and moderators of the spacing effect. For example, Bui and colleagues (2013) found that whereas participants with lower working memory capacity benefited more from having easier intervening tasks in between repetitions of items, participants with higher working memory capacity benefited from having more difficult intervening tasks. Similarly, CPS's argument cannot account for non-monotonic lag effects—findings that benefits can be larger after shorter intervals than after longer intervals. For example, Appleton-Knapp et al. (2005) found that whereas exact repetitions benefit from a longer spacing interval, varied repetitions benefit from a shorter spacing interval. Finally, CPS's argument that spacing is connected to working memory resource depletion is also contradicted by evidence that spacing effects can in fact be larger when the intervening activity is more taxing and hence there is less opportunity for working memory recovery. For example, Bjork and Allen (1970) found that spacing benefits were larger when participants were given a more difficult intervening task than an easier intervening task in between repetitions of items. Although in this study the intervening task was digit rehearsal and thus might be considered incidental learning, the fact remains that according to CPS's proposal there should be no effect of moderators such as individual differences in working memory capacity and the difficulty of the intervening task.

A Path Forward

Since Ebbinghaus (1885/1964), the finding that spaced practice yields better retention than massed practice has been the topic of intense research and theorizing. So has interleaving, a related phenomenon wherein close alternation of related topics also improves memory and transfer (e.g., Brunmair & Richter, 2019; Carvalho & Goldstone, 2019; Yan & Sana, 2021). In the extant literature, spacing has been typically defined as any interval between repeated presentation, whether this interval takes place on the order of seconds (Petersen et al., 1962, 1963), days (Cepeda et al., 2006), or even a year (Bahrick et al., 1993; Cepeda et al., 2008). Classic spacing studies have often used word list or paired associates paradigms in which a list of unrelated words or word pairs are presented (Cepeda et al., 2006), each word/pair is repeated and the final test is a memory test (e.g., repeatedly studying the word pair "apple - tower" and being later tested on retrieving the second part of the pair "apple - ?"). On the other hand, interleaving has been typically defined as alternating the study or practice of different concepts. Classic interleaving studies have often used a paradigm in which repetitions are often varied rather than exact, and in which the final test involves transfer of learning to new examples (e.g., studying multiple examples of the concept "fraction addition" and being asked to solve novel fraction addition problems at test).

There are many different, non-mutually exclusive theories that have been proposed to account for spacing effects on retention. It is beyond the scope of this commentary to provide an elaborated discussion of these theories (see e.g., Carpenter, 2017). For example, study-phase retrieval theory of spacing argues that learning is improved when retrieval of prior presentations is difficult, but successful (Appleton-Knapp et al., 2005; Bjork & Allen, 1970; Cuddy & Jacoby, 1982; Krug et al., 1990). This theory therefore predicts an optimal spacing interval—where retrieval is not trivial nor impossible, and can account for various different moderators of the spacing effect (e.g., intervening task difficulty). Importantly, interleaving inherently includes spacing: When the examples from one category are interleaved with examples from other categories, this necessarily means that the examples from a given category are spaced apart from each other. In other words, spacing and interleaving have been used to describe slightly different paradigms: Spacing usually refers to memory studies, interleaving usually refers to category learning studies. In spacing studies, it does not matter what is inserted into the intervals between repetitions; in interleaving studies, special attention is paid to the similarity of the concepts or categories being learned.

Although interleaving different concepts necessarily involves spaced repetition of different categories, its benefit for learning has also been suggested to result from comparison and contrast of the properties of examples of different concepts (see discriminative contrast theory, Kang & Pashler, 2012; sequential attention theory, Carvalho & Goldstone, 2017). A recent meta-analysis showed support for this theory, finding that interleaving similar categories yields larger interleaving benefits than does interleaving dissimilar categories (Brunmair & Richter, 2019).

Critically, benefits that result from spaced presentation of examples from the same category and benefits that arise from discriminative contrast of different categories are not mutually exclusive, as CPS argue they are. Both can occur (Birnbaum et al., 2013) and perhaps it is time to theoretically bridge spacing and interleaving theories (Yan et al., 2020). Unfortunately, CPS's two proposed mechanisms are mutually exclusive—learners can either be engaging in discriminative contrast or resting, but not both—and hence does not account for the broader evidence base. It is, of course, possible that CPS's theory can account for much of the evidence for spacing and interleaving benefits in the literature. But to do so, it would require clearly defining what constitutes concept similarity, specifying the process for how information gets stored from working memory to long-term memory, and considering established moderators of both spacing and interleaving.

References

Appleton-Knapp, S. L., Bjork, R. A., & Wickens, T. D. (2005). Examining the spacing effect in advertising: Encoding variability, retrieval processes, and their interaction. *Journal of Consumer Research, 32*(2), 266-276. https://doi.org/10.1086/432236

Bahrick, H. P., Bahrick, L. E., Bahrick, A. S., & Bahrick, P. E. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science, 4*(5), 316–321. https://doi.org/10.1111/j.1467-9280.1993.tb00571.x

Birnbaum, M. S., Kornell, N., Bjork, E. L., & Bjork, R. A. (2013). Why interleaving enhances inductive learning: The roles of discrimination and retrieval. *Memory & Cognition*, *41*(3), 392–402. https://doi.org/10.3758/s13421-012-0272-7

Bjork, R. A., & Allen, T. W. (1970). The spacing effect: Consolidation or differential encoding. *Journal of Verbal Learning & Verbal Behavior, 9*(5), 567–572. https://doi.org/10.1016/S0022-5371(70)80103-7

Brunmair, M., & Richter, T. (2019). Similarity matters: A meta-analysis of interleaved learning and its moderators.*Psychological Bulletin*, *145*(11), 1029–1052. https://doi.org/10.1037/bul0000209

Bui, D. C., Maddox, G. B., & Balota, D. A. (2013). The roles of working memory and intervening task difficulty in determining the benefits of repetition. *Psychonomic Bulletin & Review, 20*(2), 341–347. https://doi.org/10.3758/s13423-012-0352-5

Carpenter, S. K. (2017). Spacing effects in learning and memory. In J. T. Wixted, & J. H. Byrne (Eds.), *Learning and Memory: A Comprehensive Reference* (pp. 465–485). Academic Press.

Carpenter, S. K., & Mueller, F. E. (2013). The effects of interleaving versus blocking on foreign language pronunciation learning. *Memory & Cognition, 41,* 671–682. http://dx.doi.org/10.3758/s13421-012-0291-4

Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th grade students' retention of U.S. history facts. *Applied Cognitive Psychology, 23*(6), 760– 771. https://doi.org/10.1002/acp.1507

Carvalho, P. F., & Goldstone, R. L. (2017). The sequence of study changes what information is attended to, encoded, and remembered during category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(11), 1699–1719. https://doi.org/10.1037/xlm0000406

Carvalho, P. F., & Goldstone, R. L. (2019). When does interleaving practice improve learning? In J. Dunlosky & K. A. Rawson (Eds.), *The Cambridge handbook of cognition and education* (pp. 411–436). Cambridge University Press. https://doi.org/10.1017/9781108235631.017

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354–380. https://doi.org/10.1037/0033–2909.132.3.354

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science, 19*(11), 1095–1102. https://doi.org/10.1111/j.1467–9280.2008.02209.x

Chen, O., Paas, F. & Sweller, J. (2021). Spacing and interleaving effects require distinct theoretical bases: a systematic review testing the cognitive load and discriminative-contrast hypotheses. *Educational Psychology Review*. https://doi.org/10.1007/s10648-021-09613-w

Cuddy, L. J., & Jacoby, L. L. (1982). When forgetting helps memory: An analysis of repetition effects. *Journal of Verbal Learning & Verbal Behavior, 21*(4), 451–467. https://doi.org/10.1016/S0022-5371(82)90727-7

Ebbinghaus, H. (1885/1964). *Memory: A contribution to experimental psychology* (trans. H. A. Ruger, C. E. Bussenius, & E. R. Hilgar). New York: Dover Publications.

Foster, N. L., Mueller, M. L., Was, C., Rawson, K. A., & Dunlosky, J. (2019). Why does interleaving improve math learning? The contributions of discriminative contrast and distributed practice. *Memory & Cognition, 47*(6), 1088–1101. https://doi.org/10.3758/s13421-019-00918-4

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. psyh. https://doi.org/10.1002/acp.1801

Kapler, I. V., Weston, T., & Wiseheart, M. (2015). Long-term retention benefits from the spacing effect in a simulated undergraduate classroom using simple and complex curriculum material. *Learning and Instruction, 36,* 38–45. https://doi.org/10.1016/j.learninstruc.2014.11.001

Krug, D., Davis, T. B., & Glover, J. A. (1990). Massed versus distributed repeated reading: A case of forgetting helping recall? *Journal of Educational Psychology, 82*(2), 366–371. https://doi.org/10.1037/0022-0663.82.2.366

Ostrow, K., Heffernan, N., Heffernan, C., & Peterson, Z. (2015). Blocking vs. interleaving: Examining single-session effects within middle school math homework. In *Artificial Intelligence in Education* (pp. 338–347). Springer International Publishing.

Peterson, L. R., Saltzman, D., Hillner, K., & Land, V. (1962). Recency and frequency in paired-associate learning. *Journal of Experimental Psychology, 63*(4), 396–403. https://doi.org/10.1037/h0043571

Peterson, L. R., Wampler, R., Kirkpatrick, M., & Saltzman, D. (1963). Effect of spacing presentations on retention of a paired associate over short intervals. *Journal of Experimental Psychology, 66*(2), 206–209. https://doi.org/10.1037/h0046694

Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58*(3), 193–198. https://doi.org/10.1037/h0049234

Schwartz, M. (1975). The effect of constant vs. varied encoding and massed vs. distributed presentations on recall of paired associates. *Memory & Cognition*, *3*(4), 390-394. https://doi.org/10.3758/BF03212930

Vlach, H. A., Sandhofer, C. M., & Kornell, N. (2008). The spacing effect in children's memory and category induction. *Cognition, 109*(1), 163–167. https://doi.org/10.1016/j.cognition.2008.07.013

Yan, V. X., & Sana, F. (2021). Does the interleaving effect extend to unrelated concepts? Learners' beliefs versus empirical evidence. *Journal of Educational Psychology, 113*(1), 125–137. https://doi.org/10.1037/edu0000470

Yan, V., Schuetze, B., & Eglington, L. G. (2020). A Review of the Interleaving Effect: Theories and Lessons for Future Research. Preprint. https://psyarxiv.com/ur6g7/

Young, A. P., Healy, A. F., Jones, M., & Bourne, L. E. (2019). Verbal and spatial acquisition as a function of distributed practice and code-specific interference. *Memory & Cognition, 47*(4), 779-791. https://doi.org/10.3758/s13421-019-00892