



Reply to Sana et al.'s (2022) Commentary on Rest-from-Deliberate-Learning as a Mechanism for the Spacing Effect

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

Accepted: 6 April 2022 / Published online: 15 April 2022
© The Author(s) 2022

Abstract

Sana and colleagues (2022) have raised a number of challenges regarding the operationalisation of constructs and selection of articles to Chen et al.'s (Educational Psychology Review 33:1499–1522, 2021) suggestion that resting from cognitive activity could possibly allow for working memory recovery and so explain some of the data on the spacing effect. In our response, we indicate that the goal of our proposed framework was to try to resolve some mixed results of the spacing and interleaving effects and offer an alternative explanation for those mixed results, rather than proposing a theory of everything. We acknowledge that there are other important factors, which does not however, provide grounds for rejecting our hypothesis. Additional empirical studies are needed to determine whether rest and its effect on working memory are important when analysing the spacing effect.

Keywords Spacing effect · Interleaving effect · Working memory resource depletion · Discriminative-contrast · Rest-from-deliberate-learning

Introduction

Sana et al. (2022) have provided a useful critique of Chen et al. (2021) paper in which we suggested the possibility that resting from cognitive activity can allow working memory recovery and so provide an explanation of the spacing effect while the discriminative-contrast hypothesis can provide an explanation of the interleaving

✉ Ouhao Chen
o.chen@lboro.ac.uk

¹ Department of Mathematics Education, Loughborough University, Loughborough, UK

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

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

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

effect. Before addressing the specific criticisms of Sana et al., we would like to indicate the general goals of our 2021 paper.

The first point to note is that we did not invent the discriminative-contrast hypothesis to explain the interleaving effect. That hypothesis was proposed by Kornell and Bjork (2008) and we feel it is the best hypothesis available. We agree that there may well be other causal factors also leading to the interleaving effect but nevertheless, suggest that our review indicates that there is strong evidence that the discriminative-contrast hypothesis provides a major causal factor. We also note that the field is chaotic with respect to the relation between the spacing and interleaving effect with some researchers treating them as the same effect and others as distinct effects. The discriminative-contrast hypothesis assumes they are distinct effects because the effect is more likely to be obtained using more similar than less similar interleaved tasks (Kang & Pashler, 2012; Zulkiply & Burt, 2013). Ultimately, of course, this issue must be resolved empirically.

We have provided a new explanation of the spacing effect (Chen et al., 2018) based on the assumption that working memory resources diminish with cognitive effort and replenish with rest, an explanation that is incompatible with the interleaving effect. Furthermore, Chen et al. (2018) provided empirical support for this hypothesis, support that Sana et al. (2022) unfortunately ignored.

Can the effect of rest on working memory explain all of the enormous database associated with the spacing effect? It clearly cannot but neither can any other single explanation. Our only claim is that resting can explain at least some and possibly a relatively large part of the data on the spacing effect. In doing so, we believe we can substantially reduce the chaos associated with this effect. We will next respond to specific criticisms raised by Sana et al. (2022).

Definitions of Levels of Rest and Levels of Similarity

Sana et al. (2022) object to the way we describe levels of rest and levels of similarity as follows:

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.1502). 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. For example, CPS include the studies from Young et al. (2019) as evidence in favor of their rest-from-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. (p. 2 – p. 3)

We think the comments of Sana et al. (2022) concerning our classification of Young et al. (2019) are invalid and specifically reject their statement above “...that participants were instructed to encode for later retrieval”. Participants in the spaced group were not informed those irrelevant passages would be tested prior to the actual test. By the end of the experiment, a multiple-choice test was conducted on those passages but since participants were not previously informed of the test, it is not clear that reading those passages resulted in any intentional learning or cognitive effort beyond that required to read. Without intentional learning, there is no reason to assume a working memory load. We classified reading without intent to learn as rest.

Sana et al. (2022) go on to say:

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. 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). (p. 4)

We reject the validity of this criticism. In a verbal learning task, if intervals between repetitions of a bi-gram are filled with other bi-grams to be learned, then learners must discriminate between the bi-grams, with no self-evident rule available to indicate the category to which the different bi-grams belong. In other words, each bi-gram had to be specifically discriminated from the other bi-grams, a difficult task. That contrast provides a clear test of the discriminative-contrast hypothesis and possibly a reason why the originators put forward the hypothesis in the first place. The Schwartz (1975) study is a very good example for our proposed framework and the study actually tested an interleaving effect rather than a spacing effect, as no incidental learning was involved, and participants needed to contrast different bi-grams and associated word pairs for the final recall test.

When It Is Easy or Difficult to Discriminate Between Categories

Sana et al. (2022) seem to imply that all discrimination tasks are equivalent:

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.1514). 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. (p. 3)

The difference between learning pairs in one’s native language or learning how to pronounce words is that learning pairs is a purely cognitive activity while learning how to pronounce words is primarily a sensory-motor task. We do not know whether acquiring sensory-motor skills requires the same cognitive activities as memorizing word pairs but if they are different, as is highly plausible, it would hardly be surprising if the discrimination-contrast hypothesis applied to one but not the other. On the other hand, it would be unacceptable if the discrimination-contrast hypothesis was entirely rejected simply because it does not apply to sensory-motor tasks. We note that Sana et al. (2022) do not themselves provide an explanation for the contrasting results.

With respect to this issue, Sana et al. (2022) go on to indicate:

CPS point to Ostrow et al. (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 $g = 0.60$. (p. 3)

Since this result accords precisely with our predictions, we were remiss in not pointing it out. High-skilled students are likely to be already familiar with the different categories of problems and so do not need to learn to discriminate between them. Low-skilled students are less likely to be similarly familiar with the categories and so it is advantageous for them to be placed in a condition that facilitates learning the appropriate discriminations.

The next point made by Sana et al. (2022) may have more validity:

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, Cohen's $d = 0.62$. CPS report this result as evidence of the discrimination mechanism (See Table 2 from page 1509 to page 1511 in Chen et al., 2021). What CPS omit, however, is that the interleaving benefit was larger with the dissimilar set of mathematics problems (Cohen's $d = 1.00$). (p. 4)

This argument is certainly stronger and may be valid. It is certainly not as obviously invalid as the previous arguments that we have rejected. Nevertheless, the expertise levels of the learners are not clear, and for novices, all the categories may be indistinguishable. The authors found no significant differences on prior knowledge across groups, namely participants' expertise was equally randomised, but they did not clearly report participants' prior knowledge of the given topic.

Selected Evidence for Systematic Review

Sana et al. (2022) challenged the way we selected evidence for our proposed theoretical framework:

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-analysis 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. (p. 4)

Here are our inclusion and exclusion criteria:

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

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

Using these criteria, we thoroughly searched the literature from SCOPUS, Web of Science and other major databases. Based on our definition and proposed framework, we re-categorised some interleaving studies as spacing studies and some spacing studies as interleaving studies, which may account for our increased number of interleaving studies and partially account for the decreased number of spacing studies, such as Toppino and DiGeorge (1984) and Russo et al. (1998), which were re-categorised as interleaving studies.

Our inclusion and exclusion criteria clearly were different from Cepeda et al. (2006), especially our exclusion criteria. We had more exclusion criteria while Cepeda et al. (2006) only excluded studies using clinical participants, which inevitably led us to have fewer studies/experiments. For example, Cepeda et al. (2006)

included dissertations such as Actkinson (1977), that we excluded for our search, and research about mood and connotation which are closer to motivation than learning, such as Elmes et al. (1984), that we also excluded. In any case, Sana et al. (2022) provide no evidence that our conclusions would be invalidated had we used the Cepeda et al. criteria.

The Goal of Rest-from-Deliberate-Learning Framework

Sana et al. (2022) have doubts concerning the rest-from-deliberate-learning hypothesis but ignore the one study that specifically tested the hypothesis (Chen et al., 2018):

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. 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....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). (p. 3 – p. 4)

We agree that we cannot account for all the evidence but neither can any other theory. Our only claim is more of the evidence can be accounted for by the two hypotheses that we discussed.

We acknowledged other explanations for the spacing effect and indicated how the working memory resource depletion explanation matches or does not match those other explanations. The goal of our proposed framework was to try to resolve some mixed results of the spacing and interleaving effects and offer an alternative explanation for those mixed results, rather than proposing a theory of everything.

We agree we cannot explain lag effects but doubt anyone else can either if only because the data are inconsistent. We certainly make no claim that we know the ideal interval times. At this point, we do not know how rapidly working memory depletes or recovers under various circumstances.

Sana et al. (2022) also state:

it is unclear how CPS's theory would account for well-documented interactions and moderators of the spacing effect. For example, Bui et al. (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. (p. 5)

The issue is not whether working memory depletion and recovery explains everything but whether it can explain some patterns of results. We believe it can but are

equally certain that there are other important factors. The existence of those other factors does not provide grounds for rejecting our hypothesis.

Task Complexity

Sana et al. (2022) also use task complexity to dispute the rest-from-deliberate-learning hypothesis:

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. 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. (p. 5)

Bjork and Allen's (1970) participants have presented some trigrams followed by some digit shadowing tasks at varying levels of difficulty that did not require any learning. There is no evidence nor reason to believe that these differing shadowing tasks differentially affected memory structures. In our proposed framework, we are concerned with learning tasks rather than simple processing tasks and have no difficulty accepting that the rest-from-deliberate-learning hypothesis does not apply to processing tasks unrelated to learning.

A Path Forward

We believe our review has provided strong evidence concordant with the rest-from-deliberate-learning hypothesis, evidence that does not exclude other possible causative factors for the spacing effect. Nevertheless, whether rest and its effect on working memory is important when analysing the spacing effect must ultimately be determined empirically. The only published study that we know that has tested the hypothesis is the Chen et al. (2018) study which provided evidence for the hypothesis. If the major point made by Sana et al. (2022) is that additional work is required, we agree.

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

References

- Atkinson, T. R. (1977). The effects of structural, presentation, and response variables on storage and recall of simple prose materials (Doctoral dissertation, Texas Christian University, 1976). *Dissertation Abstracts International*, 37, 5388.
- Bjork, R. A., & Allen, T. W. (1970). The spacing effect: Consolidation or differential encoding. *Journal of Verbal Learning & Verbal Behavior*, 9, 567–572.
- Brunmair, M., & Richter, T. (2019). Similarity matters: A meta-analysis of interleaved learning and its moderators. *Psychological Bulletin*, 145, 1029–1052.
- 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, 341–347.
- Carpenter, S. K., & Mueller, F. E. (2013). The effects of interleaving versus blocking on foreign language pronunciation learning. *Memory & Cognition*, 41, 671–682.
- 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, 354–380.
- Chen, O., Castro-Alonso, J. C., Paas, F., & Sweller, J. (2018). Extending cognitive load theory to incorporate working memory resource depletion: Evidence from the spacing effect. *Educational Psychology Review*, 30, 483–501.
- 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*, 33, 1499–1522.
- Ebbinghaus, H. (1885/1964). *Memory: A contribution to experimental psychology* (trans. H. A. Ruger, C. E. Bussenius, & E. R. Hilgar). New York: Dover Publications.
- Elmes, D. G., Chapman, P. F., & Selig, C. W. (1984). Role of mood and connotation in the spacing effect. *Bulletin of the Psychonomic Society*, 22, 186–188.
- Foster, N. L., Mueller, M. L., Was, C., Rawson, K. A., & Dunlosky, J. (2019). Why does interleaving improve math learning? The contributions of discriminative contrast and distributed practice. *Memory & Cognition*, 47, 1088–1101.
- Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the “enemy of induction”? *Psychological Science*, 19, 585–592.
- Kang, S. H., & Pashler, H. (2012). Learning painting styles: Spacing is advantageous when it promotes discriminative contrast. *Applied Cognitive Psychology*, 26, 97–103.
- 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.
- Russo, R., Parkin, A. J., Taylor, S. R., & Wilks, J. (1998). Revising current two-process accounts of spacing effects in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 161–172.
- Sana, F., Yan, V. X., & Carvalho, P. F. (2022). On rest-from-deliberate-learning as a mechanism for the spacing effect: Commentary on Chen et al. (2021). *Educational Psychology Review*
- The effect of constant vs. varied encoding and massed vs. distributed presentations on recall of paired associates. *Memory & Cognition*, 3, 390–394.
- Toppino, T. C., & DiGeorge, W. (1984). The spacing effect in free recall emerges with development. *Memory & Cognition*, 12, 118–122.
- Young, A. P., Healy, A. F., Jones, M., & Bourne, L. E. (2019). Verbal and spatial acquisition as a function of distributed practice and code-specific interference. *Memory & Cognition*, 47, 779–791.
- Zulkipli, N., & Burt, J. S. (2013). The exemplar interleaving effect in inductive learning: Moderation by the difficulty of category discriminations. *Memory & Cognition*, 41, 16–27.

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