

Commentary on “Beyond Time-on-Task: The Relationship between Spaced Study and Certification in MOOCs”

Zachary A. Pardos

University of California, Berkeley, USA

pardos@berkeley.edu

ABSTRACT: In Miyamoto et al. (2015, this issue) the authors looked to substantiate the presence of the spacing effect, referenced from the psychology literature, in several MOOCs. Their secondary analyses constituted a robust, empirical finding on the correspondence between session distribution and certification but with only a coarse, analogous relationship to the theory of distributed practice. This article underscores the difficulty of validating theory and inferring causation from data produced in the wild but also the resultant rigour necessitated by this difficulty.

Keywords: MOOC, spacing effect, distributed practice, sessions, time on-site, theory, commentary, online education

1 SUMMARY OF MIYAMOTO ET AL.

The authors (Miyamoto et al., 2015) take inspiration from the spacing effect results observed in the psychology literature in guiding their MOOC data explorations. In that literature, numerous experiments found that, holding time on task equal, participants who were given repeated spaced exposure to particular memorization tasks were twice as likely to demonstrate recall than those given repeated blocked (or massed) exposure. The authors investigate to what degree they could detect this effect in MOOCs with a historical data analysis approach as opposed to the traditional randomized control trial approach employed by the scholars their work was inspired by. The premise of the approach was to look, post-hoc, into student performance data from 20 online course offerings at HarvardX, and measure if participants who spent approximately the same amount of time in the course had differential certification rates depending on the number of sessions they logged. The authors assert, with some reservations, that high certification in these courses is analogous to the outcome of recall in the cited studies, and that spaced online sessions with the course is analogous to spaced repetition of memorization tasks. Given the authors’ approach, they would ideally like to have observed high variability in the number of sessions given the same time on course (or on task) so as to then look for differential certification rates with respect to a metric of spacing, such as sessions per unit time. They find, to the contrary, high correlation between session count and time on course and low variability in the number of sessions given a particular relative time on course. Nevertheless, by utilizing the size of the dataset, the authors find a way around this issue by separating participants into bins (“deciles”) corresponding to time on course and look for differences in certification given high, medium, and low terciles of session counts. They find that overall, high session counts corresponded to a 40% certification rate compared to 28% among low session count students in the same total time decile. The authors

(2015). Commentary on "Beyond Time-on-Task: The Relationship between Spaced Study and Certification in MOOCs. *Journal of Learning Analytics*, 2(2), 70–74. <http://dx.doi.org/10.18608/jla.2015.22.6>

point out that a potential confound could be that generally high performing students space their sessions out more, and that the session spacing itself is not the causal factor. To try to tease this apart, the authors look at within-student spacing effects. They find that the relative number of sessions between two courses for a given student can vary but that the relationship between spacing and certification holds, at least ruling out session count tercile membership as a course invariant attribute of the student.

2 COMMENTARY

The aspiration to connect to a seminal result in the psychology literature is a laudable goal, as is the determination to investigate the phenomenon using a data mining approach given the wealth of data at their disposal. In addition to the robust correlation between session spacing and certification, the techniques employed of controlling for identified confounds, post-hoc, by stratifying the confounding variable (total time) and the use of deciles in order to normalize measures between courses of differing lengths were valuable methodological contributions.

Given the topic of theory of this special issue, it is particularly important to focus on the substantiation between the theory of distributed practice and the authors' findings. Of primary concern is the incongruence in the material presented to subjects in the spaced repetition studies from psychology and the material in sessions in the online courses. In the psychology studies, where it was pointed out that a 2x increase in recall was found, the task being repeated was rote memorization where the repeated stimulus was a particular word or word-number pair. This type of memorization task and its respective stimuli is very far afield from the learning of general college course material and begs the question of what the repeated stimuli is across sessions within a course. The authors acknowledged this incongruence but did not go as far as arguing for why theirs remained an appropriate context to study this theory.

An additional threat to the appropriateness of the context is the lack of necessity to memorize in a MOOC. The majority of assessments, aside from rare proctored tests, are open-book and open-Internet. If information was not as well recalled with a massed study approach, students could easily compensate by referencing the forgotten materials as needed, and thus suppressing any difference between the recall (or certification outcome) of spaced versus massed practice.

It is not difficult to imagine that spaced study can better foster deeper conceptualization of material, even if a direct connection to a concrete theory is not yet substantiated. Furthermore, it is not without precedence that a finding from secondary analyses precedes theoretical explanation or experimental validation. In the Educational Data Mining field (Koedinger, D'Mello, McLaughlin, Pardos, & Rosé, 2015), after all, exploratory findings often prime the design of follow-up experimentation and further data modelling. It must be asked then that even if the spacing of sessions causes improved certification in MOOCs, then, in this unmanipulated environment, what causes students' spacing behaviour?

(2015). Commentary on “Beyond Time-on-Task: The Relationship between Spaced Study and Certification in MOOCs. *Journal of Learning Analytics*, 2(2), 70–74. <http://dx.doi.org/10.18608/jla.2015.22.6>

Interaction with students’ prior knowledge of the course material? Discipline? The authors’ results certainly create an appetite for a randomized controlled trial to manipulate spacing in a MOOC in order to develop causal claims with respect to certification.

3 RELATED WORK AND TEMPORAL DISTRIBUTION vs. SPACING

Missing from this work is an explanation for why distributed practice has been shown to be effective in particular circumstances in the literature and the parameters that experimenters have found to mediate its effect, namely; duration of each exposure (time before the next event), frequency of repetition, and number of intervening events in between exposures.

Melton (1970), cited by the authors, reflects on how his contemporaries, who have attempted to study distributed versus massed practice, moved on from “skill-learning” tasks, with crudely defined practice trials and accomplishment measures, to rote verbal-learning tasks. These rote verbal-learning tasks served as the basis for the majority of works cited in this article to substantiate the existence of a distributed practice (DP) effect (also referred to as spaced practice in the article) over massed practice (MP). The application of this theory to MOOCs is a return to the skill-learning contexts abandoned by his contemporaries and which Melton posits are inappropriate for these theories of practice.

There exists a tension, unaddressed by the authors, between temporal distribution of practice and space between practices. In the continuous paired–associate learning experiments by Peterson, Wampler, Kirkpatrick, & Saltzman (1963) and Young (1966), subjects were exposed to various word-integer pairs and then later asked to recall the number associated with the word. The number of intervening pairs between repetitions that maximized recall was seven to eight, after which the effect abated. A similar result was demonstrated in free-recall experiments by Melton & Shulman (1967), where subjects were verbally presented with a list of nouns and then later asked to recall as many as they could. In this case there was a sharp increase in percentage of recall with two to eight intervening items between repetitions with an asymptote of recall around twenty to forty. These studies point to greater space between repetitions as facilitating the effect. This may be at odds with the authors’ result of increasing outcomes correlating with the number of sessions. The more sessions (assuming uniform temporal distribution), the less space between practices. The literature suggests less space between practice will lead to less recall, while Miyamoto et al.’s results appear to be suggesting the opposite effect. It should also be pointed out that “space,” as defined in the literature, refers to other stimuli (words, word-integer pairs). The definition of space in this article is left undefined by Miyamoto et al. but certainly could mean interactions with another course or any activity that does not include interacting with the course being analyzed.

The authors drew on the mediating effects of the total time law on recall, citing Underwood (1970); however, in these cited studies, frequency of repetition, and thus total time, is varied, as opposed to the authors’ data manipulation of fixing total time and varying frequency. Is greater frequency (with lower

(2015). Commentary on “Beyond Time-on-Task: The Relationship between Spaced Study and Certification in MOOCs. *Journal of Learning Analytics*, 2(2), 70–74. <http://dx.doi.org/10.18608/jla.2015.22.6>

exposure duration) or less frequency (with higher exposure duration) better for recall? The authors do not make explicit which of these finer grained mediating factors are attributable to the observed differences in certification.

4 CONCLUSIONS

Without espousing as firm a connection to the spacing effect from the psychology literature, the empirical finding of the robust relationship between sessions per total time and certification would stand on its own as a strong data mining result. In order to elevate this relationship from discriminating between outcomes towards explaining them, authors from learning analytics must engage deeper with the relevant theories to articulate the alignment (or misalignment) of assumptions from the theoretical frameworks to the nascent educational context being studied. This articulation would then serve as a contribution back to theory. Interestingly, it appears undeniable that exploring their data under the pretense of connecting to a particularly theory drove the authors to conduct deeper, more rigorous and interesting analyses than would otherwise have been conducted by the collaboration. In this way, critically projecting their findings through the prism of theory has provided a clearer, more informative view of student behaviour than previous behavioural analyses with MOOCs have achieved. This article is evidence that the learning analytics field is well served with an eye towards theory in its practices and that growth in its impact and rigour is likely to be attained with contribution, in equal parts, from theory, methodology, experimentation, and practice.

REFERENCES

- Koedinger, K. R., D’Mello, S., McLaughlin, E. A., Pardos, Z. A., Rosé, C. P. (2015). Data mining and education. *WIREs Cognitive Science*, 6(4), 333–353. <http://dx.doi.org/10.1002/wcs.1350>
- Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior*, 9(5), 596–606. [http://dx.doi.org/10.1016/S0022-5371\(70\)80107-4](http://dx.doi.org/10.1016/S0022-5371(70)80107-4)
- Melton, A. W., & Shulman, H. G. (1967, October). *Further studies of a distributed practice effect on probability of recall in free recall*. Paper presented at Psychonomic Society, Chicago, IL.
- Miyamoto, Y. R., Coleman, C. A., Williams, J. J., Whitehill, J., Nesterko, S., Reich, J. (2015, this issue). Beyond time-on-task: The relationship between spaced study and certification in MOOCs. *Journal of Learning Analytics*, 2(2), 47–69. <http://dx.doi.org/10.18608/jla.2015.22.5>
- 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. <http://dx.doi.org/10.1037/h0046694>
- Underwood, B. J. (1970). A breakdown of the Total-Time Law in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, 9, 573–580. [http://dx.doi.org/10.1016/S0022-5371\(70\)80104-9](http://dx.doi.org/10.1016/S0022-5371(70)80104-9)

(2015). Commentary on “Beyond Time-on-Task: The Relationship between Spaced Study and Certification in MOOCs. *Journal of Learning Analytics*, 2(2), 70–74. <http://dx.doi.org/10.18608/jla.2015.22.6>

Young, J. L. (1966). *Effects of intervals between reinforcements and test trials in paired-associate learning* (Technical Report No. 101, Psychology Series). Stanford, CA: Institute of Mathematical Studies in the Social Sciences.