INVESTIGATING REPEATABLE NUDGE EFFECTS WITH SPACED REPETITION IN USING MCQs

Yew Fei Tang, Wilson Huan

School of Engineering, Nanyang Polytechnic, Singapore

ABSTRACT

The contemporary learner's learning habit is very different now due to their exposure to new technology at a young age. They have a lot on their proverbial plates; by having too many things on one's mind made them easily distracted and unable to retain the information. This led to "binge and purge" learning, a common practice whereby learners cram for an assessment by consuming subject matter in a large lump (binge) and then spitting it back in the assessment (purge). Due to their proverbial plates being full, this led to learners favoring "binge and purge" learning to achieve academic results, but not the knowledge and skills needed for the future. A constant nudging to learners to revise and test their knowledge throughout lessons could be more appropriate in evaluating these contemporary learner's academic achievement. In this paper, an initial study was carried out with weekly e-quizzes to "nudge" learners to study, and to find the optimal number of questions to be repeated from their previous lesson. The intention of these spaced repetition learning was to enhance learners' knowledge retention of key concepts. The optimal number of questions, as defined in our study, was the number of cumulative repeated questions that would help learners to retain key concepts but would not overwhelm learners with too many questions. A subsequent study was then conducted to validate the findings and its scalability with different groups of learners learning different aspects of networking concepts. The paper ends with a discussion on the effect of spaced repetition on learners' motivation and performance in learning data communication and networking by using the data gathered from the MUSIC Model of Motivation survey and learners' average scores in e-quizzes respectively.

KEYWORDS

Nudge Theory, Pedagogy for Modern Learners, Spaced Repetition, Sustainable Development, E-learning, Assessment of Learning, Standards 8

INTRODUCTION

It had been observed that more and more contemporary learners have cultivated "binge and purge" learning. These learners usually use their short-term memory to consume subject matter in a large lump (binge) and reproduced them in the assessment (purge). This memory work would not last, and the assessment credentials would be at stake. The ability of a learner to remember key concepts is vital for Engineering learners; learners are required to remember key concepts for them to innovate and apply to real-world situations. Remembering is different

from memorizing. Memorizing tends to be based on short term memory, whereas remembering is not just the process of committing information to memory but also the process of understanding, retaining, and recalling the information. The process of remembering requires the learner to understand a concept, retain it over a time period, and then recall the concept when it is needed. The weekly E-Quizzes would nudge learners to revise what they were taught for the week. The incorporation of spaced repetition learning with spaced presentation should enhance the process for learners in remembering the key concepts for all the taught topics. The "nudging effect" is expected to change learners' behavior and motivate learners to learn; and, in the long run, cultivate their interest in engineering when they see an improvement and ability to perform.

Literature Review

Nudge theory is about positive reinforcement and indirect suggestions as ways to influence the behavior and decision making of a human being. Nudging could be vital in influencing learners' behavioral and psychological factors when they made education decisions (Jabbar, 2011 and Koch et al., 2015). During learning, there were evidence that interventions with nudging as a factor could address a set of specific learning behavioral challenges (Lavecchia et al., 2016). Empirical evidence simplifying the transition to higher education is present. (French & Oreopoulos, 2017). These interventions were using new education technology and could influence one's behavior (Escueta et al., 2017). Recent empirical work from a social welfare perspective suggested too much nudging could sometimes backfire (Carroll et al., 2009; Handel, 2013; Damgaard & Gravert, 2018). Thus, it would be useful to quantify the result for the circumstances under which nudging may or may not be successful. Empirical studies had also revealed that nudges might have very heterogeneous effects (Allcott, 2011), and as a result, it may be desirable to use targeted nudges rather than universal nudges. Furthermore, behavioral interventions may be particularly relevant and effective when individuals face economic or social scarcity because it occupied the attention and potentially impedes good decision making (Mullainathan & Shafir, 2013). A reasonable degree of "nudging" depending on individual learning ability and circumstance is deemed appropriate to each individual and would influence one's behavior, thus motivating learners to learn.

Spaced repetition was well researched and had shown promising results. Spaced repetition was first introduced in Iowa (Spitzer, 1939), and research had shown manipulation of repetition space could improve recall (Melton, 1970 and Landauer & Bjork, 1978). Recent research had also shown positive results with spaced repetition (Kang, 2016 and Kelley & Whatson, 2013). According to the principle of "spaced repetition," instead of massed learning, remembering and the practice of skills were more efficient if each item's practices were spread out over time (Bloom & Shuell, 1981). Concepts that are difficult should appear more often and materials that are easy, less often, with difficulty defined according to ease with which the user could remember the material. Incorporation of interleaving between different topics was ideal as it has the potential to stretch learners beyond information retrieval to making sense of newly taught concepts (Brown et al., 2014). The difference between consistent and expanding duration between spacing had found to produce insignificant improvement in retrieval, while more repetitions were found to be more important in producing improvement in retention (Thalheimer, 2006). Spaced learning also had the potential to impact policy and curriculum planning since it could produce improved learning outcomes and higher learning per hour compared to conventional teaching methods, backed by evidence from neuroscience on rapid memory processes in humans (Kelley & Whatson, 2013). The basis that the ability of the brain to retain memory decreases overtime is based on the forgetting curve of Ebbinghaus, as shown in Figure 1. It is a theory that humans start losing the memory of learned knowledge

over time, in a matter of days or weeks, unless the learned knowledge is consciously reviewed at constant time intervals. After each revision, repetition space could be spaced apart long before the retention rate drops to 80%. If knowledge is consciously reviewed at constant time intervals, the knowledge will eventually be remembered.

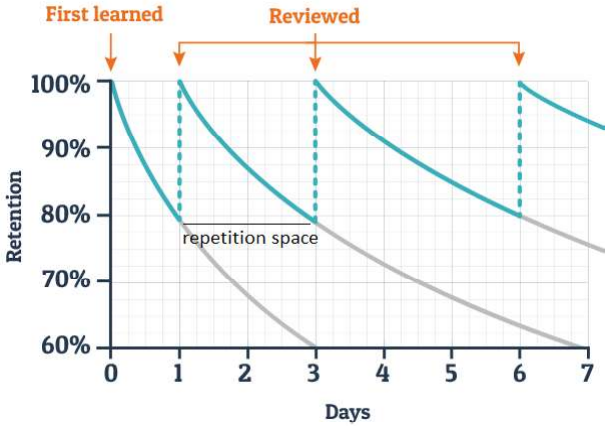

Figure 1. Typical Forgetting Curve for Newly Learned Information (Schneider, 2014)

The MUSIC model of academic motivation (Jones, 2009) is a framework used to measure and engage learners in learning. It consists of five categories of teaching strategies derived from research and theory that were critical to learners' engagement in academic settings: eMpowerment, Usefulness, Success, Interest, and Caring (MUSIC is an acronym).

Usefulness and Success relate to how learners believe the topics are relevant and important in helping them succeed, respectively. Interest indicates the level where learners like and are curious about the topic.

Evidence suggests that lecturers who address the MUSIC components were more likely to be successful at motivating their learners to engage in learning (Jones, 2013, 2015). We adopted the three categories of the MUSIC model (Usefulness, Success, and Interest) to evaluate learners' academic motivation in our investigation. We aim to gain insights into the impact of spaced repetition on learners' motivation.

METHODOLOGY

By deploying spaced repetition in the Data Communication and Networking module, we seek to investigate whether such teaching methodology help learners to better remember the technical terms and concepts to improve their academic performances in the module. Our study focus is to determine the "nudging effect" and to find the optimal number of questions to repeat in networking topics and the effect of the number of allowed attempts. The methodology is shown in Table 1.

"Nudging" in our context is to conduct weekly E-Quizzes on the topics they have learned for the week. The idea is to keep the revision materials on a weekly basis, bit-size, and manageable to revise. In having weekly E-Quizzes, the learners would be tested on their knowledge on the current material before new material is taught in the coming week's lesson. Several past E-Quizzes questions would be added in a cumulative manner to the current

material, thereby nudging them to revise on a constant basis. The assessment will contribute 20% to their total grading. The effect of the "nudging effect" would be measured in terms of their motivation.

Cohen's d is an effect size used to indicate the standardized difference between two means. It is widely used in meta-analysis. Cohen's d is an appropriate effect size for the comparison between two means.

In our case, we used Cohen's d to determine the effect size for the comparison between the mean of learners' performance in the E-Quizzes of the Experimental Control Group (EG) in the initial study and each of the experimental group's means. Cohen's d suggested that d=0.2 be considered a 'small' effect size, 0.5 represents a 'medium' effect size and 0.8 a 'large' effect size (Cohen, 1988, 1992). The rational in designing the experiment was to have the Experimental Control Group (EG) to serve as a baseline, Experimental Group 1 (E1) was to investigate the effect of having two repeat questions, Experimental Group 2 (E2) was to investigate the effect of having three repeat questions and Experimental Group 3 (E3) was to investigate the effect of the number of tries.

Table 1. Methodology for initial study

G	SS	Methodology
G	39	A weekly E-Quizzes based on the lesson conducted for the week. The E-Quiz consists of 10 multiple-choice questions (MCQs) with unlimited tries.
E1	58	Same assessment method, as stated in the control group but two questions from the previous weeks added to subsequent weeks in a cumulative manner.
E2	106	Same assessment method, as stated in the control group but three questions from the previous weeks added to subsequent weeks in a cumulative manner.
E3	98	Same assessment method, as stated in experimental group 2, but the number of tries is restricted to 3.

Legends: Group (G), Sample Size (SS), Experimental Control Group (EG), Experimental Group 1 (E1), Experimental Group 2 (E2), Experimental Group 3 (E3). Remarks: This experiment was conducted with first-year learners doing Data Communication and Network module in a networking diploma

MUSIC survey was conducted at the end of the last E-Quiz for all the experimental groups. The MUSIC survey questions were shown in Table 2.

Table 2. MUSIC Survey Questions

Q	C	Question
1	U	The repeated questions in the after-lab e-quiz was beneficial to me as I could better recall the concepts taught in class.
2	U	In general, the after-lab e-quiz was useful to me as it summarized important concepts taught in class.
3	U	I found the after-lab e-quiz to be useful for other modules in DMIT.
4	U	I will be able to use the knowledge I gained in this module.
5	U	The knowledge I gained in this module is important for my future.
6	S	I was confident that I could succeed in the after-lab e-quiz.
7	S	Throughout the module, I felt that I could be successful on the after-lab e-quiz.
8	S	I felt I could be successful in meeting the academic challenges in this module.
9	S	I am capable of getting a high grade in this module.
10	I	The after-lab quiz held my attention.
11	I	The after-lab e-quiz was interesting to me.
12	I	The instructional methods used in this module help my attention.
13	I	I enjoyed the instructional methods used in this module.
14	I	The instructional methods engaged me in the module.
15	I	I enjoyed completing the after-lab e-quiz.

Legends: Q: Question Number, C: Components, U: Usefulness, S: Success, I: Interest

RESULTS

An improvement in learners' performance in terms of average marks is observed in the experimental groups using spaced repetition learning with unlimited attempts (E1 and E2) and with limited three attempts (E3) in the initial study, as observed in Figure 2.

In the study, the result tends to contrast to the findings in literature studies that learning is improved through spaced repetition, the results from Figure 3, the study showed that there was no significant difference in the quiz scores between learners in EG and E1 (d = 0.29). The reason for the small effect size could be that only two questions per topic were repeated. Thus, the effort required by learners to recall previously taught topics was negligible and resulted in insignificant improvement in learners' academic performance. In comparing EG and E2, the repeated questions were increased from 2 questions to 3 questions. The effect size was medium (d = 0.32), an improvement over the effect size between EG and E1. This suggested that increasing the additional weekly questions from previous topics alone does not significantly contribute to a significant improvement in learner's academic performance. In order not to overwhelm learners with too many questions, the assessment method in E3 is identicial with E2 but the number of tries was limited to 3 instead of unlimited. The effect size had increased to large (d = 0.67). This suggests that restricting the maximum number of attempts for the quizzes together with a suitable number of additional questions from previous topics had a moderate impact on learners' academic performance. E3 methodology was deemed as appropriate for validation study to determine if this methodology was suitable for different levels of learners in learning key concepts in networking topics.

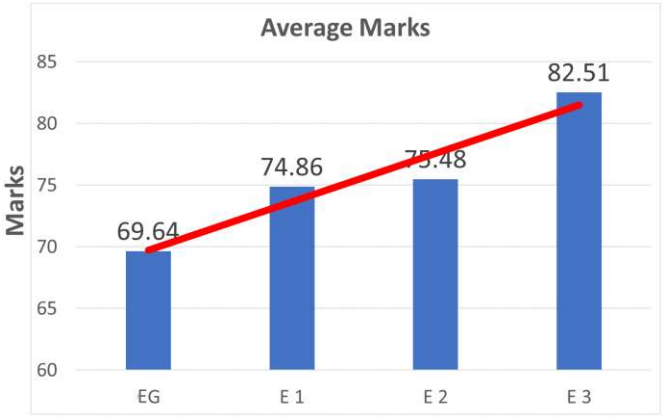

Figure 2. Average marks for all the experimental groups. The linear trendline is indicated as the red line.

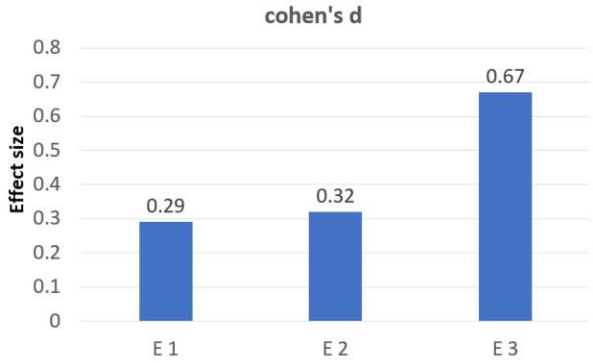

Figure 3. Cohen's d effect size for all experimental groups with respect to the control group

The performance of academically stronger learners (GPA≥3.0) in each of the experimental groups and the control group was further analyzed. The result was shown in Figure 4. Interestingly, we found that the effect size across the experimental group and the control group had significantly increased. This suggested that spaced repetition methodology had a greater influence on the academically stronger learners.

Interestingly the results show that with the same assessment methodology used, the average marks and the Cohen's effect size show the same outcome. This suggested the result could be repeated across learners studying networking modules.

Learners in the experimental groups responded positively to the spaced repetition methodology, as shown in Table 3. EG did not take the MUSIC survey as we are not making a comparison with the experimental group. The MUSIC survey is to determine if nudging will motivate our learners. They generally agreed that the after-lab e-quiz was useful, helpful, and interesting, and it helped to improve their learning. Learners remained interested in the module content, and instructional activities led them to believe they can succeed if they put forth the effort required. The increase in difficulty for the weekly e-quiz (from unlimited attempts to maximum three attempts) did not create any significant impact on learners' motivation.

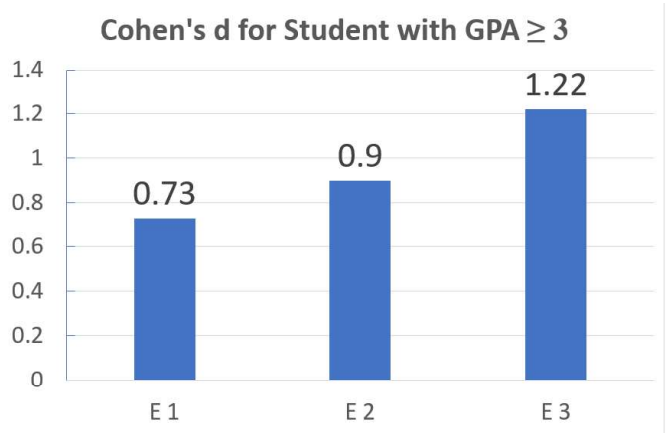

Figure 4. Cohen's d effect size for academic stronger learners (GPA≥3)

Table 3. MUSIC Survey result, taken from each experimental group at the end of the experiment

Component	Average MUSIC Score#		
	E1	E2	E3
Usefulness	5.24	5.25	4.52
Success	5.17	5.18	4.51
Interest	5.34	5.22	4.63

Likert-type scales: Strongly Agree (Score = 6), Strongly Disagree (Score = 1)

DISCUSSION

The results proved the effectiveness of using spaced repetition learning in improving learners' performance. This is in-line with other studies in virtually every standard experimental learning paradigm, with various traditional research material (Dempster, 1987a; Hintzman, 1974; Melton, 1970). However, it was interesting to see that spaced repetition was much more effective on learners who are academically stronger. Although this could be attributed to space effect, where two spaced presentations were about twice as effective as two massed presentations (Hintzman, 1974; Melton, 1970), and the difference between them increases as the frequency of repetition increases (Underwood, 1970). In this investigation, spaced presentations were employed, but the repetition frequency was the same for all learners and, therefore, may not have the same influence on the academically weaker learners as compared to the academically stronger learners. Learners who have a longer knowledge retention period generally performed better academically. This could suggest that some of the academically weaker learners could have forgotten before the questions were repeated and reviewed; therefore, the spaced repetition impact on academically weaker learners was limited.

An interesting finding was that learners responded positively, showing an improved motivation for them to learn the concepts in networking despite the increased workload required from them. The "nudging effect" in urging learners to revise on a weekly basis had influenced their behavior and motivated them to work.

CONCLUSION

There is a need to prevent "binge and purge" learning, which will not equip learners with the knowledge and skill for the real world. The learner with "binge and purge" learning may get a good grade, but they will not gain the knowledge. Continuous assessment at a bite-size level in "nudging" our learners to study on a regular basis could be the way forward for the learner to retain the knowledge as they progress. Based on our studies, the "nudging effect" of having weekly E-Quizzes in evaluating our learners' academic achievement could be more appropriate for learners to remember the key concepts for their future work. Incorporating spaced repetition learning into the weekly E-Quizzes showed great potential in improving learners' academic performance among the academically stronger learners. The hypothesis we have drawn is that the academically weaker students could have forgotten before the revision. Every learner has different knowledge retention periods, and it is logical to assume academically stronger learners have a longer knowledge retention period as compared with academically weaker learners. Further research will be necessary to understand the impact of the repetition frequency on the academically weaker learners for a holistic, practical learning system. In addition, further work to apply this study with other modules on upper-level learners would be carried out to validate the findings.

Learners generally perceive this methodology to be useful, interesting, and it helps them to improve their learning. Their motivation was also not affected by the additional effort that they must put in to remember the technical terms. In addition, learners' performances in the repetition tests provide useful feedback to learners learning progress and facilitate the lecturers to intervene and provide help to learners in a timely manner.

REFERENCES

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, 95(9–10), (pp. 1082-1095).

Bloom, K. C., & Shuell, T. J. (1981). Effects of massed and distributed practice on the learning and retention of second-language vocabulary. *The Journal of Educational Research*, 74(4), 245-248.

Brown, P. C., Roediger, H. L., & McDaniel, M. A. (2014). *Make it stick.* Harvard University Press.

Carroll, G. D., Choi, J. J., Laibson, D., Madrian, B. C. & Metrick, A. (2009). Optimal Defaults and Active Decisions. *Quarterly Journal of Economics*, 124(4), (pp. 1639-1674).

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences Lawrence Earlbaum Associates.* Hillsdale, NJ, (pp. 20-26).

Cohen, J (1992). "A power primer". *Psychological Bulletin*. 112 (1): 155–159

Damgaard, M. T. & Gravert, C. (2018). The Hidden Cost of Nudging: Experimental Evidence from Reminders in Fundraising. *Journal of Public Economics*, 157, pp. 15-26.

Dempster, E N. (1987a). Time and the production of classroom learning: Discerning implications from basic research. *Educational Psychologist*, 22, 1-21.

Escueta, M., Quan, V., Nickow, A. J. & Oreopoulos, P. (2017). *Education Technology: An Evidence-Based Review*. NBER Working Paper No. 23744.

French, R. & Oreopoulos, P. (2017). Behavioral Barriers Transitioning to College. *Labour Economics,* 47, (pp. 48-63).

Handel, B. R. (2013). Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts. *American Economic Review*, 103(7), 2643-2682.

Hintzman, D. L. (1974). Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology.* The Loyola Symposium, 77-99

Jabbar, H. (2011). *The Behavioral Economics of Education: New Directions for Research. Educational Researcher*, 40(9), (pp. 446-453).

Jones, B. D. (2009). Motivating students to engage in learning: The MUSIC model of academic motivation. *International Journal of Teaching and Learning in Higher Education*, 21(2), 272-285.

Jones, B. D., & Wilkins, J. L. (2013). Validity evidence for the use of a motivation inventory with middle school students. *In Poster presented at the annual meeting of the Society for the Study of Motivation, Washington*, DC.

Jones, B. D., & Wilkins, J. (2015). More validity evidence for the use of the MUSIC Model of Academic Motivation Inventory with upper-elementary and middle school students. *In Poster presented at the annual meeting of the Eastern Educational Research Association*, Sarasota, Florida.

Kang, S. H. (2016). Spaced Repetition Promotes Efficient and Effective Learning: Policy Implications for Instruction. *Policy Insights from the Behavioral and Brain Sciences*, 3(1), 12-19.

Kelley, P., & Whatson, T. (2013). Making long-term memories in minutes: a spaced learning pattern from memory research in education. *Frontiers in human neuroscience*, 7, 589.

Koch, A., Nafziger, J. & Nielsen, H. S. (2015). Behavioral Economics of Education. *Journal of Economic Behavior & Organization*, 115, (pp. 3-17).

Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), Practical aspects of memory, *London: Academic Press*, 625–632.

Lavecchia, A. M., Lui, H. & Oreopoulos, P. (2016). Behavioral Economics of Education: Progress and Possibilities. In: E. A. Hanushek, S. Machin & L. Woessmann, eds. *Handbook of the Economics of Education Vol. 5*. Elsevier, (pp. 1-74).

Mullainathan, S. & Shafir, E. (2013). Scarcity: Why Having Too Little Means So Much. *Time Books*, Henry Holt & Company LLC, New York, NY.

Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior*, 9, 596–606.

Schneider, W. (2014). *Memory development from early childhood through emerging adulthood*. Switzerland: Springer.

Spitzer, H. F. (1939). Studies in retention. *Journal of Educational Psychology*, 30, 641–657.

Thalheimer, W. (2006). *Spacing learning events over time: What the Research Says*. Retrieved November 31, from Work-learning: http://www.work-learning.com/catalog/

Underwood, B. J. (1970). A breakdown of the total-time law in freerecall learning. *Journal of Verbal Learning and Verbal Behavior*, 9, 573-580

BIOGRAPHICAL INFORMATION

Yew Fei Tang is a Senior Lecturer in the School of Engineering, Nanyang Polytechnic (NYP). His expertise covers artificial intelligence (AI) algorithms in the artificial intelligence group (AIG). He is the lead author for the research paper titled "Investigating Nudge Effects with Spaced Repetition Contemporary Learners Motivation and Performance in Learning Data Communication and Networking". The paper was accepted for publication and presented at ISATE 2019 in Japan. He will be pioneering using AI with spaced repetition to create a personalized education system within NYP.

Wilson Huan is currently a Development Engineer in the School of Engineering, Nanyang Polytechnic. His expertise covers cybersecurity, cloud, and data center technology and management in the Communications and Networks Group. He is one of the co-authors of the research paper titled "Investigating Nudge Effects with Spaced Repetition Contemporary Learners Motivation and Performance in Learning Data Communication and Networking."

Corresponding author

Dr. Tang Yew Fei
Nanyang Polytechnic (NYP)
180 Ang Mo Kio Ave 8,
Singapore 569830
tang_yew_fei@nyp.edu.sg

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.