Spaced Repetition Learning as a Tool for Orthopedic Surgical Education: A Prospective Cohort Study on a Training Examination

Anton Lambers, MBBS, BMSc, *,[1] and Adrian J. Talia, MD, BBiomed, DipSurgAnat, MS †,[2]

*Department of Orthopaedics, Austin Hospital, Heidelberg, Victoria, Australia; and †Department of Orthopaedics, Northern Hospital, Epping, Victoria, Australia

OBJECTIVE: Most surgical residents will learn written information via passive, massed exposure of material. Spaced repetition learning is an alternative, more effective, and efficient method to this more traditional method of teaching. We hypothesized that this method in the lead up to an Orthopedic basic sciences examination for postgraduate trainees would provide an effective solution for time-poor surgical trainees.

DESIGN: Twelve participants were prospectively enrolled in the study and provided informed consent and then were sent a pooled set of 1400 practice questions in a flashcard format using the spaced repetition learning program Anki. Study habits were tracked via the program and recorded the evening prior to the examination taking place. The final examination scores were then recorded and tabulated. All data were anonymized. Pearson correlation coefficient and p values were calculated for time spent and number of cards reviewed using the program.

SETTING: Australian Orthopedic Surgical trainees prepared for basic science examination.

PARTICIPANTS: All trainees eligible to sit the examination were approached for participation. Twelve participants started and all completed the study.

RESULTS: There was a strong correlation ($R = 0.86$, $p < 0.001$) between time spent and hence number of flashcards reviewed using the Anki program and final examination scores.

CONCLUSIONS: Results from this study support our hypothesis that spaced repetition learning using a flashcard format is an effective study strategy for learning material for Orthopedic examinations. This method of learning can be applied to other areas of surgery such as clinical examination or procedural skills. (J Surg Ed 78:134–139. © 2020 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: Education, spaced repetition learning, SRL, orthopedics, learning, flashcards

COMPETENCIES: Medical Knowledge, Practice-Based Learning and Improvement, Interpersonal and Communication Skills

INTRODUCTION

Orthopedic training is difficult, with a great deal of knowledge necessary to be acquired in a finite amount of time. Often assessment drives learning, however, many trainees will lose the knowledge they learnt for their examinations soon after. Study usually occurs via massed presentation, a common yet inefficient way of learning.[1] This includes passive acquisition of knowledge such as reading textbooks, listening to lectures, and watching videos to cover the curriculum in an ordered manner in the lead up to an examination. Information is presented "en masse" and often not reviewed in any organized way thereafter. We propose an alternative, evidence-based approach to learning during orthopedics training which improves assessment results and promotes long term retention of information.

Spaced repetition learning (SRL) is an alternative way to learn and/or retain information. The concept of *Repetitio est mater studiorum* (i.e., repetition is the mother

Correspondence: Inquiries to Department of Orthopaedics, Northern Hospital, Cooper Street, Epping, VIC 3076, Australia; e-mail: ajtalia@gmail.com

[1] ORCID: 0000-0001-7796-6047.

[2] ORCID: 0000-0002-0386-7270.

 © 2020 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved. 1931-7204/$30.00
https://doi.org/10.1016/j.jsurg.2020.07.002

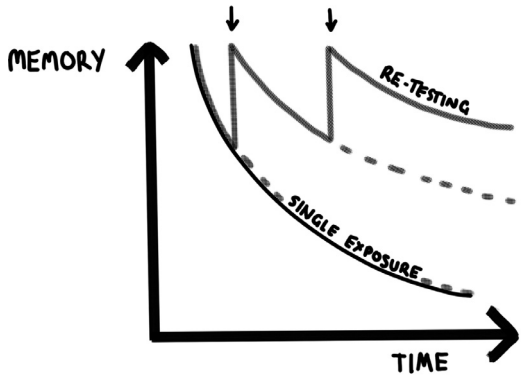

FIGURE 1. Ebbinghaus forgetting curve conceptualised.[2] This is a quantitative graph demonstrating the potential of spaced repetition for memory consolidation. The solid line demonstrates the usual course of memory consolidation after one massed exposure. The subsequent gray lines demonstrate potential memory consolidation after repeated exposures to the same material.

of studying) is a Latin phrase that acknowledges this. The modern concept of SRL has its genesis from works published in the late 19th Century by Hermann Ebbinghaus. He proposed the idea of a "forgetting curve" (Fig. 1) and noted that "frequent repetitions are indispensable in order to make possible the reproduction of a given content. Vocabularies, discourses, and poems of any length cannot be learned by a single repetition even with the greatest concentration of attention on the part of an individual of very great ability. By a sufficient number of repetitions their final mastery is ensured, and by additional later reproductions gain in assurance and ease is secured."[2] This is a seminal publication in cognitive psychology, but has failed to gain traction in mainstream education and certainly in surgical education programs. In addition, according to Edgar Dale's "cone of experience" learners retain more information of skills when they "do" a task rather than "hear, read, or observe" it.[3]

SRL enhances retention by addressing our poor ability to process and retain information when presented passively en masse,[4] and by forcing the trainee to *actively recall* information, rather than having it be passively presented to them. Using this technique, information is presented at varying time intervals depending on the trainee's interpretation of their ability to recall facts.[5] The benefits of this technique are evidence based, for not only memory retention such as in patients with Alzheimer's dementia, but for acquisition and honing of physical tasks, surgical skills, or sports performance.[5-9]

There is a paucity of evidence for this technique in medical education, and none relating to orthopedic training. Kerfoot et al. lead the way in this regard and have published widely on the subject in urological surgery learning. A 2009 randomized trial across multiple medical schools in Boston was designed to test the recall ability of 2 cohorts of students on their urology curriculum.[6] They were sent electronic learning email modules on either Benign Prostatic Hyperplasia or Prostate Cancer over a 3-month period where the content was repeated 3 times over increasing intervals, but the students in each cohort were not sent the modules for the other topic. Students who engaged in the modules were found to have dose-dependent increases in their end of term tests for those specific topics to which they had been exposed, these benefits can persist for up to 2 years in their studies.[7] Kerfoot et al. then took this concept and applied it to urology trainees throughout the country, the developed an online spaced education and testing program on diagnostic histopathology of the urological tract, skills that trainees may not be utilizing on a day-to-day basis, this randomized trial compared this exposure to material to a massed exposure of information. They demonstrated a 4-fold increase in testing scores in the longer term with spaced repetition ($p < 0.01$).[8] In addition, Kerfoot et al. have shown that the technique not only improved retention, but also learning efficiency, which is paramount for the time-poor orthopedic trainee.[9]

We designed a prospective cohort study to test the hypothesis that spaced repetition learning could lead to a higher examination pass rate on the Royal Australasian College of Surgeons Orthopedic Principles and Basic Sciences (OPBS) examination. This examination focuses on the basic science, pathology, and approaches relevant to orthopedic surgery.

MATERIALS AND METHODS

Twelve participants (orthopedic surgery trainees) sitting the OPBS examination were enrolled in the study. Each was provided with the same set of approximately 1400 practice questions based on the examination style and content. Each participant was tasked with creating questions on a particular topic which were then collated and redistributed to the group. The question style was with a single stem question with 5 true or false statements in order to accurately replicate the examination. Each participant provided written consent to participate in the study prior to the examination. Importantly, this decision was made prior to commencing studying and prior to participants being aware of their results. The participants used the electronic flashcard program Anki (ankisrs.net, Damien Elmes, GNU Open Source Software) via either their computer and/or tablet device as they wished as part of their preparation for the examination, there was no prescribed study regimen.

The Anki program automatically collects data on study characteristics as it is used. These statistics were collected from each of the participants the evening prior to the examination. Official Royal Australasian College of

Surgeons examination scores were collected by the authors from each of the study participants with their consent. Data collected included time spent on Anki program, time spent studying per day, number of reviews performed per day, total number of cards reviewed, and the maturity of studied cards. Examination score was provided as a percentage score which is how the exam is reported. The mark required to pass the OPBS examination is 70%, as predetermined by the Royal Australasian College of Surgeons. The Pearson correlation coefficient was calculated comparing time spent on the program versus examination score. All data collection was anonymized. Statistical analysis was performed using SPSS (SPSS, IBM Corporation, NY).

RESULTS

Responses

All 12 participants (100%) that agreed to participate prospectively provided a full set of both Anki and college examination result data. Candidates were spread across 3 Australian states (VIC, NSW, and WA) all sitting the same OPBS examination sitting in June of 2019.

Usage Pattern

Our cohort demonstrated different commitment to using the program, the information about the participants study habits was collected and is summarized in Table 1. The candidates started using Anki app an average of 69 days prior to the examination, and used the app for a mean of 34 days. Mean time spent on the app was 47 minutes per day and 30 hours for the duration of the study period (i.e., the entire lead up to the exam). On average candidates reviewed 3023 cards and saw two-thirds of the total question pool.

Appendix 1 collates individual data.

Relationship with Exam Score

Eleven of 12 participants (92%) passed the examination. The overall national pass rate was 67%. In our study cohort of 12, 1 participant (number 6) did not pass, whom had only undertaken 1 hour of study time on Anki. Of those that used the app for more than 1 hour, 100% passed the examination. One participant passed (participant number 8) but didn't end up using the app. Of those that used Anki the minimum usage time was 16 hours.

A strongly positive correlation coefficient between time spent on the Anki program and final examination score was found ($R = 0.86$, $p < 0.001$). This is demonstrated in Figure 2 and was again demonstrated when looking at total number of reviews versus exam score. Figure 3 examines card "maturity" (how well-learnt the card was) and illustrates the proportion of cards seen as well as seen and immature or seen and mature. Candidates who exposed themselves to more of the question deck scored higher on the examination.

TABLE 1. Study Habits of Cohort in Lead Up to Examination

	Mean	95% CI	Min	Max
Start day	69	61-77	50	91
Days app used	34	20-49	9	68
Reviews per study day	75	39-112	5	195
Time spent per study day (mins)	47	26-69	1	102
Time spent on app total (hours)	30	11-48	0	86
Proportion of total cards seen (%)	67	40-93	1	100
Number of total reviews	3023	1209-4838	62	7,539

The study habits of our cohort are detailed here. Statistics are presented as mean with 95% confidence intervals and range presented.

"Start day" refers to the number of days prior to the examination the candidate began using the program. "Days app used" refers to the number of days a candidate logged on and studied cards on the program. "Reviews per study day" is the number of cards seen on an individual day. "Time spend per study day" is the mean number of minutes per day a candidate was reviewing cards on the app, on the days they were logged on and reviewing cards. "Time spent on app" refers to number of hours using the app during the entire study period. "Proportion of cards seen" is a percentage of the cards actually seen by a candidate. "Number of total reviews" refers to the number of cards seen by a candidate during the entire study period, this included cards seen once, or multiple times depending on review schedule.

DISCUSSION

This to our knowledge is the first study directly correlating written examination scores in Orthopedics with

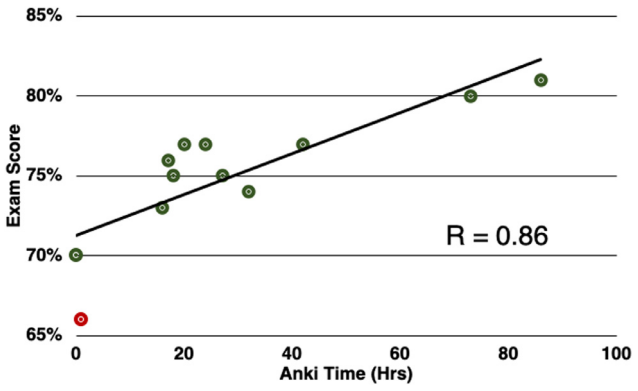

FIGURE 2. Scatter plot of examination score versus time spent on Anki program in hours. A line of best fit is drawn on the graph to mark the correlation. Pearson correlation coefficient was calculated at $R = 0.86$ with a $p < 0.001$.

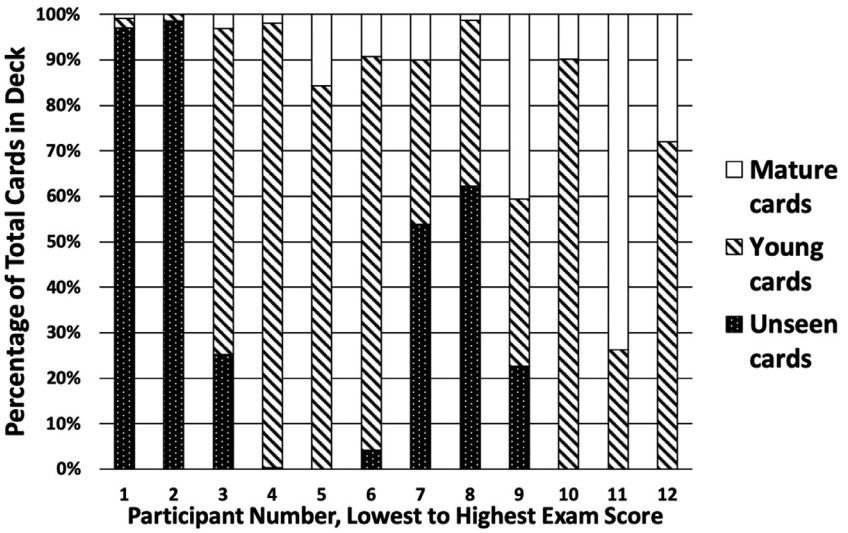

FIGURE 3. Stacked column chart of card reviews. Hundred percent is all cards in the flashcard deck. Unseen are cards that at time of exam still had not been seen by the candidate (black with dots). Young cards have been seen briefly but answered incorrectly or not yet committed to longer term memory as rated by the user (diagonal stripes). Mature cards have been seen and learnt and have a long time before being due again (clear white). Participants 1 to 12 are lined in order of lowest to highest percentage score on the exam.

Participants 1 and 2 (lowest exam scores) reviewed less than 10% of the entire question pool, whereas the candidates scoring the highest (10,11, and 12) reviewed the entire question pool. This would suggest that exposure to the entire bank of material correlated with a higher exam score.

time spent on a spaced repetition program to prove efficacy. Our hypothesis of more effective learning employing an SRL app such as Anki is subsequently proven correct. This is best reflected in the difference in examination pass rates between flashcard app users (100%) compared with all national exam candidates overall (67%). Only 1 study participant failed to pass the examination however they had used the app for only 1 hour of study. Of the participants that did use the program, more card reviews correlated strongly with a higher examination score illustrating that there is a "dose-dependent" relationship. The more cards were reviewed, the higher the trainees scored on the final examination. This dose-dependent effect mimics that found by Kerfoot et al.[9]

This data can be used to guide future study candidates. Knowing the "right amount" of study is rarely quantified and often based on myths and personal beliefs. In using data acquired prospectively through a flashcard app it has been possible to correlate study patterns with individual examination performance. This can guide people towards the time, effort and spread of study required to pass.

Despite being novel in Orthopedic learning, this concept has been widely published on in the psychology literature, dating back to Ebbinghaus in the late 19th century who coined the concept.[2] The education community has in part embraced this phenomenon with the design of school and university curricula.[10,11] The evidence is overwhelming that spaced repetition is an effective and efficient learning and education tool.[1,12] We chose to use the Anki program in our cohort as its algorithms take care of the spacing intervals and settings allow for cards to be studied in a random order. This type of mixed practice with expanding interval spacing has been shown to be effective for students when used with active rather than passive recall techniques improves long term memory or skill consolidation.[9,10]

In Orthopedic Surgery, there are no such trials as Kerfoot has published on for Urology. Studies on US residents training on laparoscopic simulators demonstrate that the learning decay effect is real not only for factual knowledge but also for procedural skills.[13] The American Board of Orthopaedic Surgery has acknowledged the benefits of spaced repetition for simulation-based tasks to master complex procedural skills such as diagnostic or therapeutic arthroscopy using simulation based training. Concerns have been raised about trainee exposure to procedures and experience secondary to the introduction of safe working hours in the UK and Europe, and duty hour requirements in the US.[14,15] Simulation of procedural skills will no doubt be a fundamental part of Orthopedic training moving forward, the American Board of Orthopaedic Surgery has acknowledged the need to make this efficient and useful for the next generation of surgeons. The authors feel strongly that spaced repetition learning should form a fundamental part of that curriculum.

This study is limited by its small numbers, narrow scope and lack of controls. The authors acknowledge that the time spent using flashcards for study is likely not the

only study that our cohort did for the examination. Informally most of the participants reported that it did indeed make up the majority of study time and for several of the candidates it replaced learning by reading. Given this was not a randomized trial there will be selection bias in that those who agreed to participate may self-select as those who are more confident of their performance. In having participants across several states of the country and large enough numbers (considering the small numbers of total exam candidates) we have attempted to mitigate this. Also, the linear relationship between time spent studying and exam score further suggests that differences in scores were more likely the result of flashcard app usage than individual differences in baseline intellect.

The merits of SRL have been proven to be overwhelming and using the built-in statistics of the app is the most pragmatic way of tracking study habits. The authors still feel there is significant advantage to this study method for written examinations in Orthopedic Surgery.

CONCLUSIONS

Spaced repetition learning is shown to be an effective method for learning basic sciences for Orthopedic surgical trainees. More time spent using a spaced repetition flashcard program in preparation resulted in higher examination scores on an Orthopedic basic sciences examination. This concept should be further studied and expanded to other parts of Orthopedic and other surgical specialty training such as procedural skills and clinical examination technique.

ACKNOWLEDGMENTS

Contribution to creating questions and/or providing exam result and usage data: Anton Lambers, Adrian Talia, Andrei Diacon, John McNicholl, Dharsh Musiienko, Jaye Yick, Sam Lynskey, Danie De Villiers, John Du, Nick Marson, and Shun Hayasaka.

APPENDIX 1. INDIVIDUAL DATA

Table A1

- Participant – participant number (random, deidentified)
- Days studied – total numbers of days app use occurred
- Total days – number of days regardless if app used from study commencing
- %Days – percentage of days studied since starting to use the app.
- Total reviews – total number of cards reviewed during the entire study duration
- Reviews/d (studied) – the average number of cards reviewed per day on the days the app was used
- Reviews/d (total) – the average number of cards reviewed per day over the entire study duration
- Review time – the amount of time spent reviewing cards on the Anki app during the entire study duration.

TABLE A1. Individual Data

Participant	Days Studied	Total Days	%Days	Total Reviews	Reviews/d (Studied)	Reviews/d (Total Days)	Review Time (Total; Hours)
1	18	70	25%	914	50	13	24
2	68	91	74%	6112	90	67	73
3	20	67	29%	1385	69	21	16
4	67	69	97%	7539	113	109	86
5	52	70	74%	2831	54	40	20
6	9	57	15%	103	11	2	1
7	30	66	45%	1833	61	28	17
8	12	70	17%	62	5	1	0
9	19	68	27%	3697	195	54	32
10	32	50	64%	1850	58	37	18
11	61	71	85%	3563	58	50	27
12	56	88	63%	6934	124	79	42

REFERENCES

1. Logan JM, Castel AD, Haber S, Viehman EJ. Metacognition and the spacing effect: the role of repetition, feedback, and instruction on judgments of learning for massed and spaced rehearsal. *Metacognit Learn*. 2012;7:175–195. https://doi.org/10.1007/s11409-012-9090-3.
2. Ebbinghaus H. Memory: a contribution to experimental psychology. *Ann Neurosci*. 2013;20:155–156. https://doi.org/10.5214/ans.0972.7531.200408.
3. Atesok K, Satava RM, Van Heest A, et al. Retention of Skills After Simulation-based Training in Orthopaedic Surgery. *J Am Acad Orthopaed Surg*. 2016;24:505–514. https://doi.org/10.5435/JAAOS-D-15-00440.
4. Greene RL. Repetition and spacing effects. *Learn Memory*. 2008;2:65–78.
5. Baddeley A. Human Memory: Theory and Practice, Revised Edition. Allyn & Bacon; 1997 Rev Sub edition.
6. Kerfoot BP, Brotschi E. Online spaced education to teach urology to medical students: a multi-institutional randomized trial. *Am J Surg*. 2009;197:89–95. https://doi.org/10.1016/j.amjsurg.2007.10.026.
7. Kerfoot BP. Learning benefits of on-line spaced education persist for 2 years. *J Urol*. 2009;181:2671–2673. https://doi.org/10.1016/j.juro.2009.02.024.
8. Kerfoot BP, Fu Y, Baker H, Connelly D, Ritchey ML, Genega EM. Online spaced education generates transfer and improves long-term retention of diagnostic skills: a randomized controlled trial. *J Am Coll Surg*. 2010;211:331–337. https://doi.org/10.1016/j.jamcollsurg.2010.04.023. e331.
9. Kerfoot BP. Adaptive spaced education improves learning efficiency: a randomized controlled trial. *J Urol*. 2010;183:678–681. https://doi.org/10.1016/j.juro.2009.10.005.
10. Nelson TO. Repetition and depth of processing. *J Verbal Learn Verbal Behav*. 1977;16:151–171.
11. Paivio A. Spacing of repetitions in the incidental and intentional free recall of pictures and words. *J Verbal Learn Verbal Behav*. 1974;13:497–511.
12. Brown PC, Roediger HL III, McDaniel MA. Make It Stick. Cambridge, Massachusetts, USA: Harvard University Press; 2014.
13. Edelman DA, Mattos MA, Bouwman DL. FLS skill retention (learning) in first year surgery residents. *J Surg Res*. 2010;163:24–28. https://doi.org/10.1016/j.jss.2010.03.057.
14. Inaparthy PK, Sayana MK, Maffulli N. Evolving trauma and orthopedics training in the UK. *J Surg Educ*. 2013;70:104–108. https://doi.org/10.1016/j.jsurg.2012.07.006.
15. Levine WN, Spang RC. ACGME duty hour requirements: perceptions and impact on resident training and patient care. *J Am Acad Orthop Surg*. 2014;22:535–544. https://doi.org/10.5435/JAAOS-22-09-535.