Enhanced Learning and Retention of Medical Knowledge Using the Mobile Flash card Application Anki

Matthew Lu[1] · John H. Farhat[1] · Gary L. Beck Dallaghan[2]

Abstract

Introduction As medical schools condense the basic science phase of undergraduate medical education, it has become increasingly important to identify methods and tools that facilitate learning, mastery, and application of medical knowledge. One increasingly popular tool that promotes engagement with content is Anki, a web-based flash card system. Using Anki, medical students can access pre-made flash cards specifically tailored to prepare students for the United States Medical Licensing Exam (USMLE) Step 1 exam. The objective of this study was to identify Anki use and its association to USMLE Step 1 performance.

Methods In March 2020, medical students in years 2, 3, and 4 who had completed USMLE Step 1 were administered a survey to measure Anki usage. The survey was locally developed and was reviewed by survey experts on campus. Survey responses were paired with USMLE Step 1 results for analyses. Descriptive and inferential statistics were used for analysis.

Results Anki usage was associated with higher USMLE Step 1 scores. Additionally, amongst those who used Anki, those with more consistent use had higher USMLE Step 1 scores and higher perceived levels of knowledge retention.

Conclusions This research suggests that Anki is an effective educational tool that should be recommended to medical students alongside other evidenced-based study tools, such as the popular question bank USMLE World. Future research should attempt to identify a relationship between Anki usage and future clinical performance to demonstrate the implications that Anki has on clinical skills.

Keywords Anki · USMLE Step 1 · Spaced retrieval · Medical students · Flash cards

Introduction

One of the greatest challenges medical educators face is helping students master, retain, and apply basic science and medical knowledge [1]. With recent moves to shorten basic science phases of the curriculum, delivery of the content is concentrated and may impact retention and mastery of the content. Although attempts at longitudinal integration of

basic science topics have occurred, content continues to be delivered in organ system blocks. Based on a review study by Custers [2], approximately one-third of this knowledge is lost after one year and nearly half in two years. Therefore, medical students often have to work hard to relearn content for the United States Medical Licensing Exam (USMLE) Step 1.

One strategy to curtail forgetting material is retrieval practice [3]. Retrieval practice, also known as the testing effect, allows learners to revisit key learning points through a range of question types from basic identification questions to more sophisticated synthesis and application questions [4]. As a result, retrieval practice facilitates the learning of new material as well as integrating new knowledge with existing schemas [5, 6]. Furthermore, the spacing out of retrieval practice, known as spaced retrieval, has also been shown to enhance learning [7]. Retrieval attempts spaced over longer periods of time were initially thought to also increase forgetting [8]; however,

✉ Gary L. Beck Dallaghan
gary_beck_dallaghan@med.unc.edu

Matthew Lu
matthew_lu@med.unc.edu

John H. Farhat
john_farhat@med.unc.edu

[1] University of North Carolina School of Medicine, Chapel Hill, NC, USA

[2] University of North Carolina School of Medicine, 108 Taylor Hall, CB 7321, Chapel Hill, NC 27599, USA

additional studies have suggested that expanded retrieval time increases recall difficulty leading to enhanced learning [9, 10]. More recent studies have shown that retrieval practice impacts performance more than time intervals of retrieval [11].

One popular and effective form of spaced retrieval practice that medical students are encouraged to use are question banks and practice exams. A study conducted by one of the authors (GLBD) at a different institution found that completing the UWorld question bank was associated with higher USMLE Step 1 scores [12]. However, that particular study did not take into account other retrieval practices of the medical students.

Another popular form of spaced retrieval practice that medical students have been using for years is flash cards [13–15]. With the advent of digital flash card programs, creating cards has become much easier. In fact, students using the same program can now download pre-made, content specific flash card decks as well as create and share their own. Once created or acquired, students can access the cards via mobile electronic devices to test themselves anywhere.

One flash card program in particular, Anki, which is reputed to be used worldwide, has been adopted by many medical students at our institution (https://apps.ankiweb. net/). The most popular pre-made Anki decks for USMLE Step 1 preparation are subdivided into organ systems. When a student is learning about a certain organ system in their medical school curriculum, they begin encountering all of the Anki flash cards related to that specific organ system. When the curriculum moves to another organ system, students can either choose to continue encountering flash cards from the past organ systems or choose to just encounter flash cards from the new organ system. Whenever a student correctly answers a flash card, the time until they encounter that card again is increased, thus taking advantage of the spaced retrieval concept.

Several studies have investigated the relationship between flash card use and academic performance measures, such as the Rana et al.'s study that showed medical students found flash cards to be helpful in learning anatomy and reducing anxiety [16]. Flash card use associated with psychiatry subject examination performance showed no improvement in one study [17] while others demonstrated examination performance on USMLE Step 1 after a large number of cards were reviewed [13, 18]. Since Anki is primarily a tool to encourage spaced retrieval, we hypothesized that the benefit would be found not only in how many unique cards are seen, but also in how often those cards are reviewed. Thus, we sought to explore how our students used Anki to study, and if particular Anki usage habits were associated with USMLE Step 1 exam performance and knowledge retention.

Materials and Methods

Survey Creation

This was a retrospective cohort study using survey research with medical students. A literature search resulted in no previously developed surveys measuring Anki usage. Therefore, a novel survey was created by the authors. The survey was reviewed by experts at the University of North Carolina's Odom Institute for Research in Social Science before being administered. No pilot testing was performed before administering the survey.

Participants were asked to report all of the study tools primarily used during the pre-clinical phase of their medical education. If students indicated that they used Anki significantly or not, they were asked to confirm their response, which enabled us to group respondents into Anki using and non-Anki using groups. For this study, it was assumed that the non-Anki using group was synonymous with non-flash card using, unless the student specified that they used another flash card system.

For those who indicated they were Anki users, questions explored the nuances of Anki use. The survey items included rating scales as well as short answer responses. There were a total of twenty-two multiple choice questions, five Likert-scale questions, and three free response questions. The complete survey is available as Appendix A.

Survey Administration

The survey was administered to every 2nd-, 3rd-, and 4th-year medical student at the University of North Carolina School of Medicine in March 2020 using Qualtrics (Provo, UT). They were invited to participate in the survey via an email explaining the purpose of the study. Consent was implied if they completed the survey. Reminders to fill out the survey were sent out once a week. After two reminder emails, the survey was closed. Survey responses were confidential, and respondents were not offered any incentives for completing the survey.

USMLE Step 1 Score Data Retrieval

First-time USMLE Step 1 scores were obtained from the office of student affairs. Scores were matched to survey respondents. Prior to analysis, data was de-identified to ensure participants remained anonymous to the research team. This study was reviewed and approved by the institutional review board (IRB No. 19–2592).

Data Analysis

Descriptive and inferential statistics were used to analyze the data. Specifically, comparisons and associations were conducted using independent samples *t*-tests, analysis of variance, and regression for normally distributed measures. For data that did not meet criteria for normal distribution, nonparametric tests were used, such as chi-square and Wilcoxon signed-rank tests. Effect size was calculated using Cohen's standard interpretation that $0.2 =$ small, $0.5 =$ moderate and $0.8 =$ large [19]. Data was analyzed using IBM SPSS v 26 (Armonk, NY).

Results

The survey was sent to medical students in years 2 ($n = 189$), 3 ($n = 187$), and 4 ($n = 169$) for a total of 545. We received 245 surveys. Of those, 34 were incomplete and removed from final analysis, leaving a total of 201 surveys for analysis. There were 11 from year 2, 95 from year 3, 93 from year 4, and 2 others (on leave for dual degree) for an overall response rate of 36.9%. Students in the "other" category are taking a year off to complete another degree or do research. Per our grouping for Anki use, 132 used Anki significantly and 69 did not. Internal consistency was calculated using Cronbach's alpha, which was 0.65.

The Most Popular Study Resources Used

The most popular study resources that were used at this institution were UWorld, First Aid, Pathoma, Sketchy Pharmacology/Microbiology, Lectures, and Anki.

Nonresponse Bias

Given the response rate, we compared USMLE Step 1 scores of survey respondents versus those who did not respond. Participants who completed the survey had a mean USMLE Step 1 score of 238.82 (*SD* 15.44, range 188–268). Those who did not complete the survey had a mean of 229.89 (*SD* 18.85, range 152–269). The mean score difference between the two groups was significant ($t = -5.71$, $p = 0.001$, $d = 0.51$) with a moderate effect size. The nonrespondent group had a greater number of scores less than 185 which brought the mean down.

Flash Card Creation

Although 57.4% of respondents used lecture material to create Anki flash cards, at least 80% of participants used a variety of pre-made Anki flash card decks that are available to download online. These flash card decks included Zanki Step Decks (72.7%), Zanki Pharmacology (57.6%), Lolnotacop (21.2%), Pepper Micro (63.6%), and Light Year (21.2%). These pre-made flash card decks have been created from popular study resources such as First Aid and Pathoma (Zanki Step Decks), Sketchy Pharmacology and Microbiology (Zanki Pharmacology, Lolnotacop and Pepper Micro), and Boards and Beyond (Light Year).

Anki Use and Exam Performance

To determine if USMLE Step 1 scores were different between Anki use and minimal use/no use, an independent *t*-test was used using a level of significance of $p < 0.05$. The mean Anki user's score was significantly higher than that of nonusers (Anki Use $M = 241.10$, $SD = 13.17$ versus No Use: $M = 235.51$, $SD = 17.68$, $t = 2.53$, $p = 0.012$, $d = 0.38$). The effect size of this difference was small to medium.

Additionally, participants indicated how often they completed full reviews of all previously completed organ blocks using a scale of never ($n = 60$), sometimes ($n = 19$), half the time ($n = 8$), most of the time ($n = 18$), or always ($n = 15$). A comparison of USMLE Step 1 scores across groups based on their completion rates indicated a significant difference (median $= 3$, $H_4 = 10.31$, $p = 0.036$). Pairwise comparisons indicated the mean score for those reviewing material sometimes ($M = 236.32$) was significantly lower than those who reviewed most of the time ($M = 246.94$, $p = 0.010$) and always ($M = 246.7$, $p = 0.013$). Pairwise comparisons with "never" were not statistically significant. The comparisons can be seen in Fig. 1.

Self-reported Impact on Learning

Participants were asked to what degree they felt Anki use helped them learn basic science material as well as retain it over time. Of the 137 responses, 36 (26.3%) felt that Anki use helped them learn material "not at all" or "a little." Even more ($n = 101$, 73.7%) felt that Anki use helped them retain material "somewhat" or "a lot."

Several questions sought to characterize Anki use behaviors. Participants rated their Anki use to review during the summer break between years 1 and 2 with the scale: not at all, a little, somewhat, and a lot. There was a statistically significant difference in USMLE Step 1 scores (median $= 3$, $H_3 = 15.71$, $p = 0.001$), specifically for those who reviewed "a lot" ($M = 248.73$) compared to those who responded "not at all" ($M = 238.67$, $p = 0.001$) and "a little" ($M = 237.31$, $p = 0.000$).

All participants indicated the percentage of material they felt they needed to relearn during their dedicated study time for USMLE Step 1, using a scale of 0–25, 26–50, 51–75, or 76–100. At this institution, dedicated study time is approximately 8 weeks from the end of the
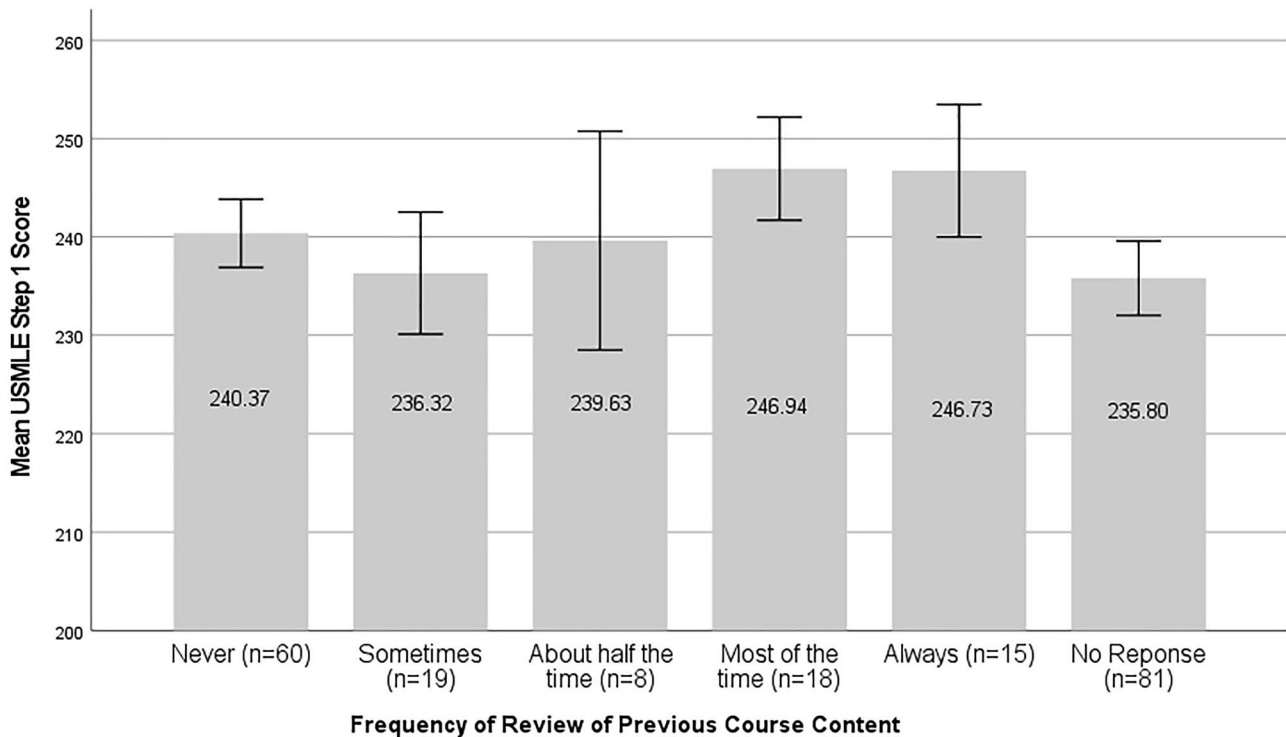
Fig. 1 Use of Anki to review previous material versus USMLE Step 1 scores. A comparison of USMLE Step 1 scores across groups based on their completion rates indicated a significant difference ($H_4 = 10.31$, $p = .036$). Pairwise comparisons indicated the mean score for those reviewing material sometimes was significantly lower than those who reviewed most of the time ($p = .010$) and always ($p = .013$)

final semester of pre-clinical courses to when students begin the required clinical year. A chi-square test indicated a significant association with self-reported summer reviewing ($\chi^2_{12} = 38.69$, $p = 0.000$). For those participants that did not review at all during the summer, 41 (33.6%) indicated they needed to relearn >26% of the material during dedicated USMLE Step 1 study time. In comparison, each individual who reviewed a lot during the summer ($n = 11$) felt they had to relearn less than 25% of the basic science material.

Similarly, consistency of Anki use was significantly associated with perceived knowledge retention rates ($\chi^2_{12} = 36.37$, $p = 0.000$). Because of response rates, individuals were grouped into those who "never," "sometimes," and "about half the time" completed full reviews (low users) and those who "most of the time" and "always" completed full reviews (high users). Thirty-nine percent of high Anki users felt they had to relearn less than 25% of the material for USMLE Step 1 compared to 7% of low Anki users. Only 12% of high Anki users felt like they had to relearn more than 51% of the material compared to 41% of low Anki users. Perceived knowledge retention rates between non-Anki users and low Anki users were similar. A Wilcoxon signed-ranks test indicated that those who did not use Anki had significantly lower perceived knowledge retention rates

than all Anki users grouped together ($z = -7.23$, $p = 0.000$). The full data can be seen in Fig. 2.

Discussion

We hypothesized that Anki usage would be associated with higher USMLE Step 1 scores and knowledge retention rates because it utilizes important aspects of learning theory found in literature, specifically active learning and spaced retrieval [3, 5, 20]. The results of this research supported this hypothesis. In addition, this research builds upon previous research [13] by showing that certain Anki usage habits were associated with higher USMLE Step 1 scores and knowledge retention rates, specifically the habit of consistently using Anki to review past material.

As explained previously, students using Anki could choose to continue encountering flash cards from previous organ systems even after finishing that organ system in their medical school curriculum. Students who chose to continue encountering flash cards from past organ systems were found to have higher USMLE Step 1 scores. Furthermore, these same students reported they needed to relearn less of the USMLE Step 1 material in preparation for the exam. This suggests that students who used Anki to review consistently had better
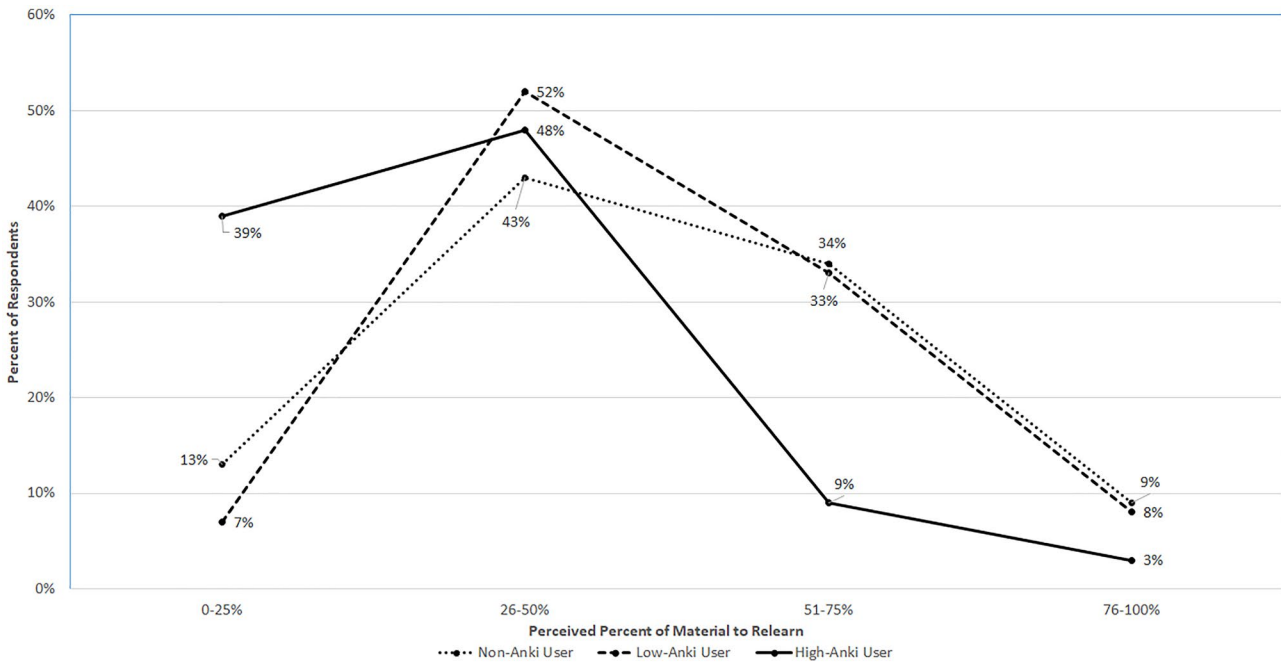
Fig. 2 Perceived percent of material to relearn based on Anki use. Low Anki users were those who "never," "sometimes," or "half of the time" completed full reviews of previous material with Anki. High

Anki users were those who "most of the time" or "always" completed full reviews of previous material with Anki

mastery and retention of the material [21–23]. These results are consistent with the learning theory of spaced repetition [3], as the students who chose to encounter flash cards from past organ systems were applying the concept of spaced repetition to a fuller degree than students who chose to only encounter flash cards from the current organ system. Therefore, how Anki is used is just as important as whether or not it is used. Ideally, Anki should be used to consistently review past material to prevent the time-consuming action of having to relearn material.

One main clinical implication of this research is that Anki is an effective studying tool that helps medical students to master the foundations of medical knowledge that they will then apply throughout their medical careers, as evidenced by higher USMLE Step 1 scores [21, 24–26]. Mastery of the fundamentals of medicine is not the only quality that a highly skilled physician requires, but it certainly is an essential one. While there is a lack of research exploring the direct relationship between Anki usage and future clinician performance, there is some literature that suggests a possible connection. Recent literature has shown a positive association between USMLE Step 1 and Step 2 scores and performance in clinical rotations during medical school, with USMLE Step 1 scores having the largest correlation with grades of honors on clinical clerkships [27]. Additionally, one meta-analysis showed that USMLE Step scores were some of the top predictors of performance in residency [28]. Therefore, since Anki is associated with higher USMLE Step 1 scores, then it is possible that it may also be associated with higher

performance in future clinical settings. Future research can be directed at exploring this possible relationship. One explanation for this is that Anki usage enables students to better master fundamental medical concepts that may be applied in the clinical setting. It is also possible that using Anki encourages self-directed learning, which is a crucial skill that carries over once the student enters the clinical realm [13].

An indirect implication of this research is that Anki usage can be an effective tool for students studying for the USMLE Step 2 Clinical Knowledge (CK) exam, which continues to be a high priority for medical students since it is still reported as a 3-digit score. While this study did not examine the relationship between Anki usage and USMLE Step 2 CK scores, research has shown that USMLE Step 1 scores are positively correlated with USMLE Step 2 CK scores [29]. Therefore, this suggests that if students were to use Anki to study for USMLE Step 1, their enhanced mastery of the material would eventually reflect in their USMLE Step 2 CK score. Anki use amongst residents in obstetrics and gynecology was associated with higher in-training examination scores [30], which further supports the rationale for our future work exploring the prevalence of Anki usage in preparation for the USMLE Step 2 CK and the relationship between Anki usage and these scores.

There are several limitations of this study. While the results of this study are consistent with results at other medical schools, it may not necessarily be generalizable to all other medical schools because the results are from a single institution. Additionally, the response rate was on

the lower side, which may affect the internal reliability. If students did not use Anki, they may have opted to not complete the survey. The lower response rate may have reflected hesitance completing the survey by students receiving a lower than anticipated score on USMLE Step 1 as evidenced by our analysis of nonresponders. Furthermore, since this study was observational in nature, it can observe correlation but not causality. While an attempt was made to reduce confounding variables, it is possible that there could have been other confounders present between the Anki and non-Anki groups, such as other study tools are timing of the examination. Finally, this study did not compare Anki to other flash card applications such as Quizlet or Firecracker, so while this study shows some of the benefits of Anki, it cannot conclude that Anki is superior to other flash card applications. Future research could also be directed at comparing popular flash card applications.

Conclusions

Overall, this study suggests that Anki usage was associated with higher USMLE Step 1 scores, which could be interpreted as a greater mastery of the fundamental medical knowledge that is tested on that standardized exam. Additionally, Anki usage was associated with a decreased perceived need to relearn material, which could indicate greater retention of the material. With the recent research that shows an association with USMLE Step 1 scores and clinical performance, Anki may not only be an effective studying tool, but also an effective tool in preparing medical students to become effective physicians.

Declarations

References

1. Schneid SD, Pashler H, Armour C. How much basic science content do second-year medical students remember from their first year? Med Teach. 2019;41(2):231–3.
2. Custers EJFM. Long-term retention of basic science knowledge: a review study. Adv Health Sci Educ. 2010;15:109–28.
3. Roediger HL III, Karpicke JD. The power of testing memory: basic research and implications for educational practice. Perspect Psychol Sci. 2006;1(3):181–210.
4. McSparron JI, Vanka A, Smith CC. Cognitive learning theory for clinical teaching. Clin Teach. 2019;16:96–100.
5. Racsmány M, Szõllõsi Á, Bencze D. Retrieval practice makes procedure from remembering: an automatization account of the testing effect. J Exp Psychol Learn Mem Cog. 2018;44(1):157–66.
6. Lee HS, Ahn D. Testing prepares students to learn better: the forward effect of testing in category learning. J Educ Psychol. 2018;110(2):203–17.
7. Greene RL. Repetition and spacing effects. In Roediger III HL (Ed.). Learning and memory: a comprehensive reference (Vol. 2): cognitive psychology of memory. Oxford: Elsevier; 2008. pp. 65–78.
8. Karpicke JD, Roediger HL III. Is expanding retrieval a superior method for learning text materials? Mem Cogn. 2010;38(1):116–24.
9. Gardiner JM, Craik FIM, Bleasdale FA. Retrieval difficulty and subsequent recall. Mem Cognit. 1973;1(3):213–6.
10. Pyc MA, Rawson KA. Examining the efficiency of schedules of distributed retrieval practice. Mem Cognit. 2007;35(8):1917–27.
11. Karpicke JD, Bauernschmidt A. Spaced retrieval: absolute spacing enhances learning regardless of relative spacing. J Exp Psychol Learn Mem Cogn. 2011;37(5):1250–7.
12. Parry S, Pachunka J, Beck Dallaghan GL. Factors predictive of performance on USMLE Step 1: do commercial study aids improve scores? Med Sci Educ. 2019;29(3):667–72. https://doi.org/10.1007/s40670-019-00722-4.
13. Deng F, Gluckstein JA, Larsen DP. Student-directed retrieval practice is a predictor of medical licensing examination performance. Perspect Med Educ. 2015;4:308–13.
14. Senzaki S, Hackathorn J, Appleby DC, Gurung RAR. Reinventing flashcards to increase student learning. Psychol Learn Teach. 2017;16(3):353–68.
15. Schmidmaier R, Ebersbach R, Schiller M, Hege I, Holzer M, Fischer MR. Using electronic flashcards to promote learning in medical students: retesting versus restudying. Med Educ. 2011;45(11):1101–10. https://doi.org/10.1111/j.1365-2923.2011.04043.x. PMID: 21988625.
16. Rana T, Laoteppitaks C, Zhang G, Troutman G, Chandra S. An investigation of Anki flashcards as a study tool among first year medical students learning anatomy. FASEB J. 2020;34(S1):1. https://doi.org/10.1096/fasdbj.2020.34.s1.09736.
17. Sun M, Tsai S, Engle DL, Holmer S. Spaced repetition flashcards for teaching medical students psychiatry. Med Sci Educ. 2021;31:1125–31.
18. Wu W, Garcia K, Chandrahas S, Siddiqui A, Baronia R, Ibrahim Y. Predictors of performance on USMLE Step 1. Southwest Respir Crit Care Chron. 2021;9(39):63–72. Available from: https://pulmonarychronicles.com/index.php/pulmonarychronicles/article/view/813. Accessed 25 Aug 2021.

19. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers; 1988.
20. Versteeg M, Hendriks RA, Thomas A, Ommering BWC, Steendijk P. Conceptualising spaced learning in health professions education: a scoping review. Med Educ. 2020;54:205–16.
21. Kerfoot BP, DeWolf WC, Masser BA, Church PA, Federman DD. Spaced education improves the retention of clinical knowledge by medical students: a randomised controlled trial. Med Educ. 2007;41:23–31.
22. Dobson JL. Effect of selected "desirable difficulty" learning strategies on the retention of physiology information. Adv Physiol Educ. 2011;35:378–83.
23. Kerfoot BP. Learning benefits of on-line spaced education persist for 2 years. J Urol. 2009;181:2671–3.
24. Boespflug A, Guerra J, Dalle S, Thomas L. Enhancement of customary dermoscopy education with spaced education e-learning: a prospective controlled trial. JAMA Dermatol. 2015;151:847–53.
25. Kerfoot BP. Adaptive spaced education improves learning efficiency: a randomized controlled trial. J Urol. 2010;183:678–81.
26. 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;211:331–7.
27. Gauer JL, Jackson JB. The association between United States Medical Licensing Examination scores and clinical performance in medical students. Adv Med Educ Pract. 2019;26(10):209–16.
28. Kenny S, McInnes M, Singh V. Associations between residency selection strategies and doctor performance: a meta-analysis. Med Educ. 2013;47(8):790–800.
29. Guiot HM, Franqui-Rivera H. Predicting performance on the United States Medical Licensing Examination Step 1 and Step 2 Clinical Knowledge using results from previous examinations. Adv Med Educ Pract. 2018;14(9):943–9.
30. Tsai S, Sun M, Asbury ML, Weber JM, Truong T, Deans E. Novel spaced repetition flashcard system for the in-training examination for obstetrics and gynecology. Med Sci Educ. 2021; ePub ahead of print. https://doi.org/10.1007/s40670-021-01344-5.