Views on resident education in the age of AI



Andrew Correll, MD
Chair, Resident-Fellow Member Committee

We are entering an exciting, new world where artificial intelligence (AI) has become one of the best ways to learn new information, capable of explaining difficult concepts in simpler terms almost instantly. I sometimes wish I'd had access to AI sooner because I think it would have greatly benefitted me, having gone through medical school in the time where Anki was first becoming very popular. Relevant to psychiatry residents and attendings alike, Anki is a program that implements spaced-repetition and active recall into learning by scheduling flashcard review at expanding intervals. Large premade decks, such as AnKing (~35,000 cards), are widely used in medical school.

One characteristic that Anki excels in is active recall, a learning paradigm that has been shown to be helpful. Thompson, Wenger, and Bartling (1978) performed a study where they had participants first hear a 40-word list and then get assigned to one of two conditions. They either heard the entire list three additional times or attempted to recall the words from memory three times without hearing the list again. Although hearing the list four times produced better recall after 20 minutes, that advantage disappeared after 48 hours. Measured against its own 20-minute performance, the repeated-exposure group had lost roughly 56% of what it initially recalled, compared with only about 13% lost by the repeated-recall group.[1] Flashcard usage in general tends to be associated with higher GPA and test scores.[2]

Spacing out this active recall is also helpful. In a group of 26,000 physicians who completed quarterly clinical questions, some were randomized to have six previously answered questions repeated either zero, one, or two times at different three-month intervals.[3] Repeating questions improved later recall and transfer to new clinical cases, with two repetitions outperforming one.[3] Two repetitions versus one in this study produced d = 0.43 for learning and d = 0.20 for knowledge transfer. Butler and Roediger (2008)[4] had students read prose passages and then either take no test or complete a multiple-choice test with no feedback, immediate feedback after each question, or delayed feedback after the entire test. On a cued-recall test one week later, students recalled 11%, 33%, 43%, and 54%, respectively, showing that retrieval improved retention and that feedback—especially after completing the test—strengthened it further. Their data are summarized below, adapted from a figure in a later review by Roediger and Butler:[5]

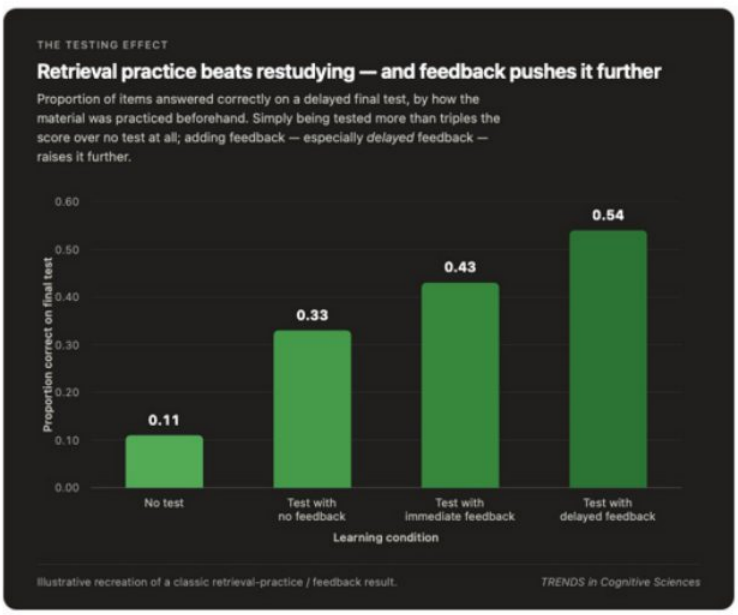

Figure 1 Adapted depiction of a figure of data from Roediger HL, Butler AC. The critical role of retrieval practice in long-term retention.Trends in Cognitive Sciences. 2011;15(1):20-27. doi:10.1016/j.tics.2010.09.003

At Wright State University's Boonshoft School of Medicine, classmates in my academic year ran an Anki study and found an association of Anki users

scoring significantly higher across four large summative exams in medical school.[6] Anki users scored about 6–7 percentage points higher on the three course exams and 10.7 points higher on the year-end Comprehensive Basic Science Examination (CBSE), an NBME-authored exam used to predict Step 1 performance. These results were after adjusting for baseline MCAT performance, which was higher in Anki users at the 73rd percentile vs 65th.[6] The raw CBSE difference was 12.9 points (72.2% vs. 59.3%). These findings are associational, and the baseline gap is itself a caution: students who choose Anki may differ in motivation and study habits in ways that statistical adjustment cannot fully capture. A similar result was found in a survey of medical students where daily Anki usage (vs. non-daily) was associated with a small increase in median Step 1 scores of 238 compared to 233.5 (used data from just before Step 1 became pass/fail in 2022), as well as improved sleep quality (though Step 2 scores were not correlated).[7] One major problem that I and other users have had with Anki, though, is that flashcards are intentionally written in short, high-yield sentences. For example, below is an Anki card:

Typical antipsychotics exert their effects by blocking the {{c1::D_2 dopamine}} receptor

The curly braces and c1 refer to a section the program will occlude and test the user on, with the user either hitting the space bar if they get the answer right or another key (often "j") if they cannot recall the information. The short format sometimes prevents an adequate context to be formed for cards that contain information a learner is seeing for the very first time, a problem that I and other medical students often dealt with when trudging through as much information each day as medical school requires. However, with the use of AI, the context can be instantly provided by a simple query and the explanation given to your exact level of understanding.

The Loma Linda program, which generously produces the PRITE Ninja review PDF that many psychiatry residents use, has now turned that content into Anki cards as well. Sometimes even this information, which is short form, can lack context. With the use of new tools such as Claude Code, the Wright State Department of Psychiatry is experimenting with making active recall and spaced repetition more accessible for PRITE study and board review. We are testing our own website for internal program use that quizzes our psychiatry residents on the last roughly twelve years of PRITE questions and tracks their performance over time, giving an estimated PRITE score based on percentage correct and past standard deviation statistics. This website could not have been made except for the use of advanced AI tools such as Claude Code that allow website features to be added in real time and with simple requests that any psychiatry resident, without coding training, could make. This method has also been extended to include email reminders prompting users to complete a set number of questions each day and in-class live cellphone polling of questions from previous tests to support group discussion and collaborative learning.

We also hope to adapt the user interface of this program and create something that can be shared publicly for use with generally available psychiatric material that will allow attendings and psychiatry residents alike to learn from the advancing artificial intelligence tools that continue to improve. While bugs have arisen in development, they are usually fixable by another prompt to AI, and such tools may help us become better physicians by creating resources that help us apply the latest psychiatric knowledge. As our department experiments with new spaced repetition (the website resurfaces past missed questions periodically) and

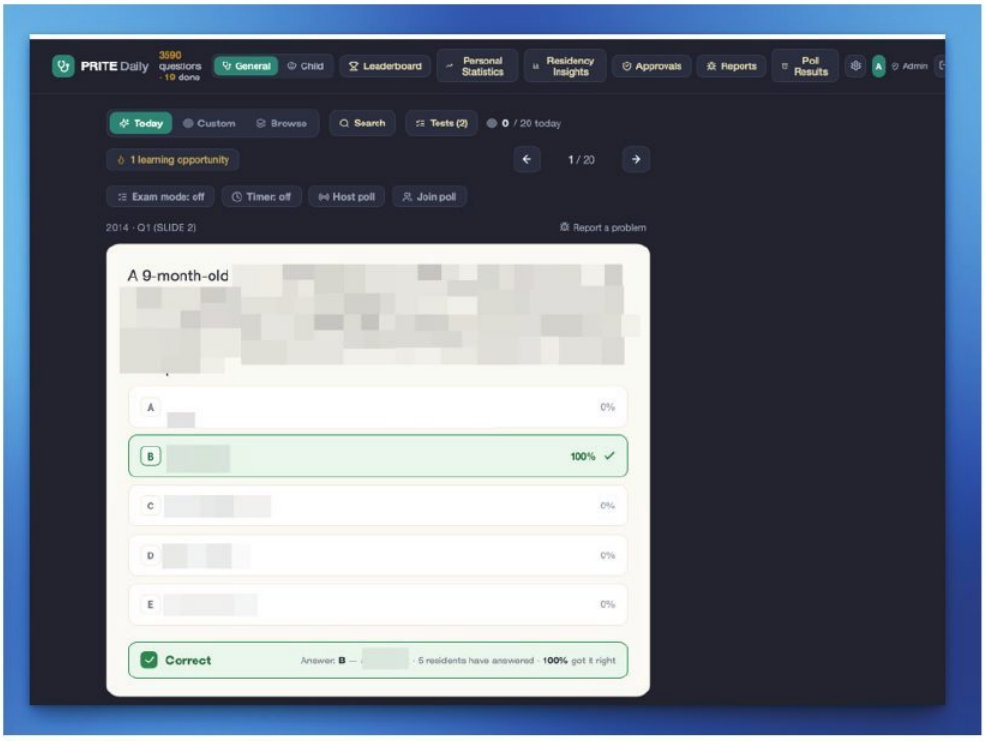

Figure 2 A screenshot (with the PRITE question protected) of the website's user interface and design, a model that other residencies could adapt for their own purposes.

active recall methods, we will closely monitor for hallucinated or inaccurate information as efforts are underway to safely let AI provide context for these concepts as well. Ethically, practice question scores likely will need to remain anonymous unless residents opt in to sharing, but residency-wide topics of weakness will likely prove useful for our residency-leadership to improve focus of didactics in needed areas.

As we move into a world that is itself moving fast, it is a bright future for the psychiatric resident who is tackling new treatments from lumateperone to Cobenfy, from TMS to Spravato. May we move forward mindful that incorrect information may occasionally be presented with alarming confidence by AI (not totally unlike many psychiatry residents at some point in training), while maintaining—as any therapist might encourage—optimism about what this technology could become.

References

