Optimizing Learning with Artificial Intelligence: Individual and Organizational Benefits

Amy M. Smith
Blank Slate Technologies
Arlington, VA
amy@blankslatetech.co

ABSTRACT

Learning science offers a wide range of techniques to enhance knowledge acquisition and long-term retention. The field is at a pivotal moment as technological advances shift it from broad recommendations to the use of machine learning and other AI tools that provide adaptive cognitive support. We present results from three research studies demonstrating the individual- and organization-level benefits of an AI-based learning platform in academic learning, military training, and a military-adjacent workplace. Findings across studies show that, as a result of using a platform that integrates machine learning with evidence-based learning techniques, individuals from a variety of organizations experience improved long-term knowledge retention in a fraction of the usual study time. Further, individuals demonstrate an improved ability to apply their knowledge flexibly to solve new problems (university students), pass skills-based tests (military students), and reduce mistakes made in the workplace (security guards). Additionally, organizations benefit from reduced training time and increased visibility into knowledge gaps, which the machine learning algorithm detects and mitigates. These findings highlight the transformative potential of AI to deliver the best practices from the learning sciences efficiently and effectively, thereby optimizing both individual and organizational performance.

ABOUT THE AUTHORS

Amy M. Smith is the Chief Scientific Officer at Blank Slate Technologies and holds a PhD in experimental psychology. Her research investigates cognitive approaches to memory optimization with a focus on using technology to strengthen memory against the effects of acute stress, neurodegenerative disease, and aging. Amy has a history of conducting behavioral research with the U.S. Army, U.S. Air Force, and in higher education. At Blank Slate, Amy spearheads all research efforts, including formal collaborations with academic and government institutions and case studies with the company's clients.

Optimizing Learning with Artificial Intelligence: Individual and Organizational Benefits

Amy M. Smith
Blank Slate Technologies
Arlington, VA
amy@blankslatetech.co

BACKGROUND

In both academic settings and professional training programs, the typical approach to accomplishing knowledge retention is to teach material, administer an assessment, and then presume the knowledge has been retained. However, decades of cognitive science research show that without timely review, newly learned information fades quickly from memory. In the late 19th century, psychologist Hermann Ebbinghaus conducted pioneering research on the topic, revealing a phenomenon now known as the forgetting curve (Figure 1; Ebbinghaus, 1913). His research, and replication studies conducted more recently (Murre & Dros, 2015), showed that we forget new information rapidly after learning it—as much as 50% within a day—unless we actively reinforce it. In a real-life applied setting, researchers similarly found that flight medics, who scored 92% on a knowledge test after their eight-hour annual training, scored just 65% on the same test six months later (Wynn & Black, 1998).

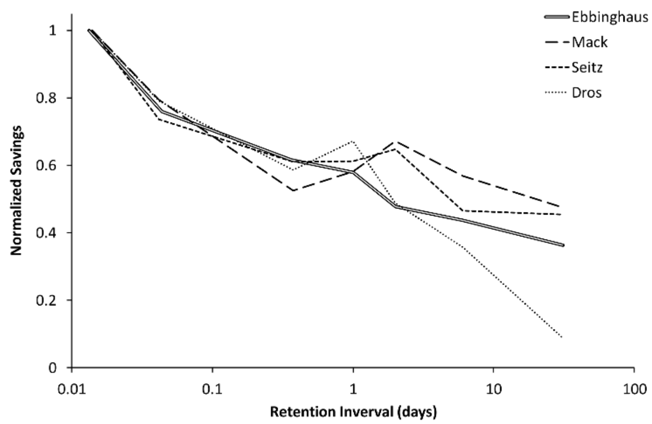

Figure 1. Rates of Forgetting (Calculated as Normalized Memory Savings on the y-axis) Over a 100 Day Interval (x-axis) for Hermann Ebbinghaus and Researchers Mack, Seitz, and Dros, Who Replicated Ebbinghaus' Experiment in 2015. Figure Reproduced from Mack & Seitz (2015) under a Creative Commons Attribution License.

Unfortunately, the normal process of forgetting can have dangerous consequences. In one analysis of 1,299 reports from the Aviation Safety Reporting System, researchers identified 75 incidents directly linked to memory errors, such as omitting checklist items (Nowinski et al., 2003). Tragically, two of these incidents resulted in crashes, claiming 308 lives. Similarly, in a study examining 188 adverse surgical events across three hospitals over six months, researchers found that 46 incidents were due to memory errors such as momentary lapses or complete omissions of essential training (Suliburk et al., 2019). These individual cases reflect a broader trend identified in a 2009 analysis of 90 studies across various industries, including manufacturing, healthcare, defense, and transportation (Christian et al., 2009). The analysis concluded that safety knowledge was the strongest predictor of safety performance in workplaces, underscoring the importance of reinforcing critical knowledge year-round and not just once annually.

Since Hermann Ebbinghaus' foundational work on memory, cognitive scientists have sought to fully understand the constraints of human memory and how to overcome them. In recent years, this body of research has increasingly come to be known as "learning science," reflecting its focus on identifying the most effective strategies for helping people retain information over the long term, spanning weeks, months, and even years.

Learning science research has shown that popular study methods—like re-reading, highlighting, rote memorization, and repetitive drills—are often inefficient and ineffective. Instead, strong evidence supports using three strategies together to create lasting memories: retrieval practice, interleaving, and spacing (Agarwal & Roediger, 2018; Brown et al., 2014; Carpenter et al., 2022). This means learners should practice actively recalling the information they need to learn (retrieval practice), vary the order in which they practice (interleaving), and revisit the material repeatedly and spread out overtime (spacing). A common way learners intuitively combine these techniques is by reviewing shuffled flashcards for several weeks leading up to a major exam. Researchers believe these learning strategies are so effective because they, both individually and in combination, require effortful recall, which strengthens each memory network to a greater degree than less effortful study techniques (e.g., Pyc & Rawson, 2009).

Despite its immense utility, spaced and interleaved retrieval practice (SIR) can be burdensome for learners, particularly when they have an immense amount of information to retain. Consider a stack of 1,000 flashcards: How should a learner prioritize which ones to review first? The stack of flashcards places a heavy burden on both the learner's time and metacognition, requiring them to either dedicate hours each day to tackling the stack or to create a system for determining which cards they do and do not know well and then binning them accordingly. The latter strategy depends on learners' metacognitive ability to accurately assess what they know and what they do not, an area where most individuals are not perfectly calibrated (Koriat & Bjork, 2005; Kruger & Dunning, 1999).

Recent advancements in artificial intelligence (AI) offer a solution to the burden of implementing the best learning techniques. Contemporary machine learning algorithms can now model individual forgetting patterns, optimizing review schedules to minimize the time commitment required for spaced-retrieval programs and remove the metacognitive burden on the learner. One example of this approach is a mobile platform, henceforth referred to as the *SIR intervention,* that leverages SIR and machine learning advancements to provide personalized long-term memory support (McHugh et al., 2021). This system employs a flashcard-style interface designed to aid users in retaining targeted information over extended periods. Notably, its machine learning algorithm monitors individual forgetting rates, delivering customized spacing schedules tailored to each user's unique memory retention needs. Independent research on the SIR intervention found that, relative to traditional flashcard review, the adaptive algorithm resulted in similar ceiling levels of memory retention but for about one-third of the time commitment (McHugh et al., 2021).

The present paper summarizes results from one experimental study and two case studies examining the efficacy and utility of using the AI-optimized SIR intervention to continuously support memory retention. All three studies used the platform previously described, though there are other products that offer similar memory support through SIR. The first study, an experiment conducted with undergraduate students, showed that optimized SIR may help teach learners new content in the absence of formal instruction, and that the knowledge gained can be successfully applied to answer new questions. The second study, a case study conducted with US Air Force students, shows how optimized SIR can both combat the forgetting curve that happens post-training and help people apply their knowledge in skills-based scenarios. The third study, a case study involving security guards at a data center, demonstrates how optimized SIR can identify and close knowledge gaps across high-performing teams, resulting in reduced workplace incidents. Together, these studies demonstrate an innovative and effective alternative to infrequent training events: A time-efficient SIR intervention that promotes year-round knowledge recall and resulting enhancements in performance.

STUDY 1

Background and Research Objectives

Lectures are a learning staple in workplaces and educational settings. However, as discussed above, lectures must be accompanied by additional knowledge reinforcement, such as a program of SIR, to prevent memory decay. In an experimental study conducted at Assumption University in Worcester, Massachusetts, we aimed to determine whether students could learn new information through a program of SIR in the absence of a lecture (Gordon et al., 2025). If successful, this approach would simultaneously accomplish knowledge acquisition and reduce the burden on educators to deliver lectures. We also aimed to determine whether the information acquired through the SIR intervention could be flexibly applied when answering novel questions pertaining to the newly learned topic.

Method and Results

We used G*Power (Faul et al., 2007) to estimate the minimum sample sizes needed to detect medium effects ($f = 0.15$) with 80% power and $\alpha = .05$. For the 2×5 repeated measures ANOVA on learning curves (assuming $r = .80$ among repeated measures), the required sample was 24 participants. For paired-samples t-tests on final test scores, the required sample was 27 participants.

Twenty-eight students from Assumption University participated in the four-week study; all participants were screened to ensure they were novices to the Psychology topics chosen as study stimuli. Participants attended an initial one-hour orientation session either in person or via video conference, during which they provided informed consent, completed demographic questionnaires, and were randomly assigned to watch one of two 27-min lecture videos. The videos were condensed versions of 50-min classroom lectures on (1) social/personality psychology and (2) psychological disorders. Afterward, participants downloaded the mobile application described above.

For the following four weeks, participants engaged in the SIR intervention every one to three days, per the machine learning algorithm, to answer questions about the lectured topic and the non-lectured topic. The question banks for each topic were 25 questions each, and questions comprehensively covered the material presented in each lecture. Participants completed 15 questions each session spending, on average, 9.05 (SD = 3.94) min per week on non-lectured questions and 7.92 (SD = 3.28) min per week on lectured questions. Learning curves comparing mastery of lectured versus non-lectured material are presented in Figure 2. Though on Day 1 of the intervention participants had higher accuracy on questions related to lectured content than non-lectured content ($t(27) = 4.53$, $p < .001$, $d = 0.75$), by Day 8, participants' accuracy on lectured versus non-lectured questions no longer differed ($t(27) = 1.91$, $p = .878$, $d = 0.08$).

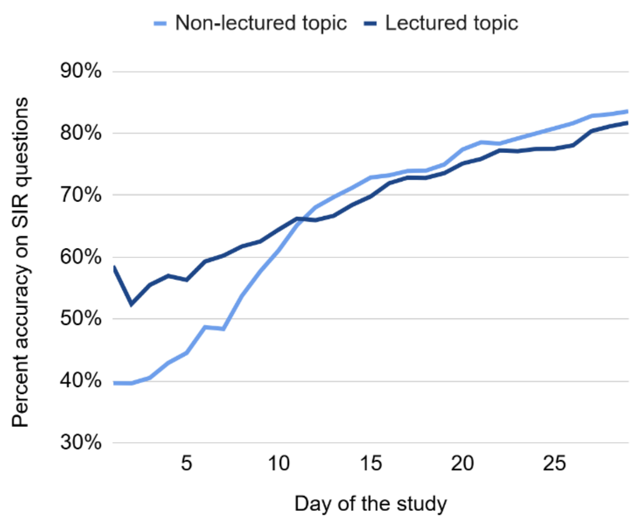

Figure 2. Learning Curves for Information that was Lectured and then Reinforced Via the SIR Intervention Versus Solely Learned Via the SIR Intervention.

The day after completing four weeks of the SIR intervention, participants were emailed a link to a memory test administered via Qualtrics. Participants first answered 50 multiple-choice questions that were rephrased versions of the 50 questions participants delivered via the SIR intervention: 25 re-phrased questions for each of the lectured and non-lectured topics. They then answered four open-ended questions, two per topic (lectured vs. non-lectured), in fewer than 250 words per answer. Both the multiple-choice and short-answer questions thus assessed application of learning rather than rote memorization of the content. As shown in Figure 3, on both tests, participants' accuracy did not differ for questions related to the lectured content versus the non-lectured content (multiple choice test: $t(27) = 0.47$, $p = .319$, $d = 0.09$; short-answer test: $t(27) = 1.20$, $p = .119$, $d = 0.23$).

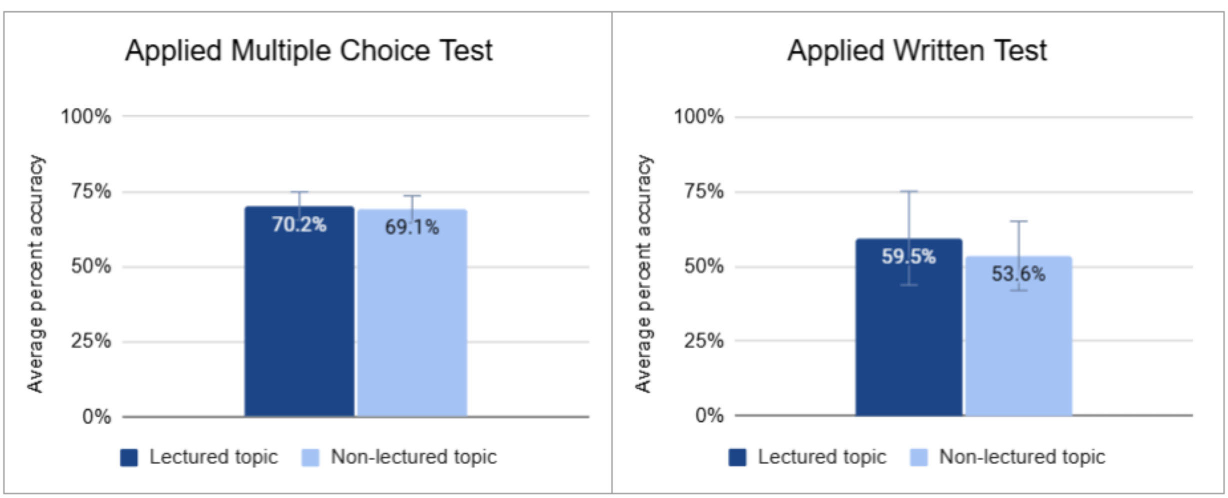

Figure 3. Participants' Performance on Applied Multiple-Choice and Written Tests After Using the SIR Intervention to Reinforce Knowledge for One Month. Error Bars Represent *SEM*.

Discussion

This study demonstrated that an optimized SIR intervention can impart foundational course content in the absence of a lecture. Although attending a preceding lecture initially boosted participants' performance on relevant questions during the SIR intervention, after one week this advantage disappeared and participants performed equally well on final test questions that probed at lectured and non-lectured content. This approach also saved time when considering that participants spent less total time learning new, non-lectured content via the SIR intervention ($M = 36.20$ min) than they would have spent simply watching the intended 50-min lecture in a classroom setting. Further, a similar prior study showed that when a lecture was the only content exposure students had, they remembered just 57% of the content one month later (McHugh et al., 2021), compared to the 69% memory retention we observed on the final multiple-choice test for non-lectured content. Thus, teaching new content solely via this program of SIR may be more time efficient and more effective for long-term retention than the conventional practice of delivering lectures.

This study was small-scale and warrants replication and extension with larger and more diverse populations before generalizations can be confidently made. Future researchers should consider replicating this methodology in applied educational settings, determining whether this method is similarly effective for more difficult subject matter, and examining memory durability over longer periods of time than just one month.

STUDY 2

Background and Research Objectives

In 2021, the 351st Special Warfare Training Squadron (351 SWTS) at Kirtland Air Force Base identified a critical need for knowledge support in their pararescue students. These students, training to rescue others in combat and high-risk situations, were demonstrating low pass rates on a medical skills assessment that was administered six weeks into the course. The assessment requires students to administer care under fire in a live simulation, something they learn how to do during their combat medicine training that precedes entry into the 351 SWTS. Thus, pararescue students were demonstrating knowledge deficiencies for information they had previously learned and mastered.

In 2022, the pararescue students began the SIR intervention previously described to reinforce and remaster the medical knowledge that was lapsing between their prior combat medicine training and their entry into the 351 SWTS. I aimed to determine whether the SIR intervention could close the knowledge gap that formed between courses and improve pass rates on their medical skills assessment.

Method and Results

The convenience sample included 89 male airmen enrolled in three separate cohorts of pararescue students at the 351 SWTS. These cohorts spanned the periods of January-May 2022, March-August 2022, and June-December 2022. Students were introduced to the SIR intervention during the week prior to the start of the course, with onboarding facilitated by the lead instructor. Engagement with the SIR intervention was strongly encouraged but not mandatory.

Out of the 89 students, 76 chose to engage with the SIR intervention regularly during the period of interest: the first six weeks of the course when students refresh on their combat medic training and then take a skills assessment. The SIR intervention delivered reminders via push notifications and/or emails every one to five days, depending on each student's customized spacing schedule. These reminders prompted them to review the curated set of questions selected by the machine learning algorithm. To reduce the time burden on the students, questions were capped at 15 per day. The question set was developed by a course instructor and comprised 52 pararescue-specific items targeting common knowledge gaps (e.g., "125mg of ____ should be taken twice daily beginning 24 hours prior to ascent to prevent acute mountain sickness."). Instructors kept records of student pass rates on the medical skills assessment administered during the sixth week of training, which served as the primary outcome measure.

As shown in Figure 4, in their first week of the SIR intervention, students demonstrated the knowledge deficiency that their instructors had noted: They averaged 72.65% (SD = 44.59%) accuracy on their assigned questions. By the sixth week of the course, during which they take their medical skills exam, students averaged 91.35% (SD = 28.12%) accuracy, a significant improvement from the first week ($t(1,780) = 8.09$, $p < .001$, $d = 0.45$). During these six weeks, students spent an average of 6.92 (SD = 5.76) minutes engaging with the SIR intervention per week.

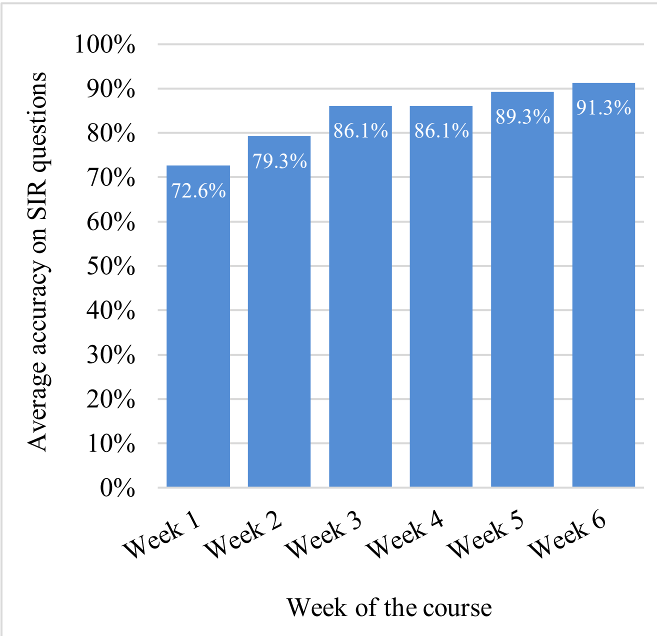

Figure 4. Knowledge Improvements for USAF Pararescue Students During a Six-Week Period of Using the SIR Intervention.

Of the 76 students who voluntarily engaged with the SIR intervention, 12 failed their first attempt at the skills assessment in the sixth week of the course, versus 7 of the 13 who did not engage with the intervention. Thus, students who used SIR to reinforce their combat medicine training were more likely than those who did not to pass their medical skills assessment during the sixth week of their course, as confirmed by a Fisher's exact test (log odds ratio: -1.80, 95% CI [-3.27, -0.38], $p = .005$).

Discussion

This study showcased the utility of an AI-driven spaced-retrieval program to identify and address knowledge deficiencies in military training. The SIR intervention identified substandard knowledge levels in students and then retrained their knowledge to above 90% average accuracy for fewer than seven minutes of weekly use per student, on average. Further, engagement with the program was associated with improved pass rates on a medical skills assessment that required application of knowledge in a real-life, care under fire simulation.

These results underscore two important points. First, training tools that incorporate machine learning and AI have great potential to improve human cognitive performance for minimal time investment. Second, strengthening memory in this way can result in flexible knowledge that can be applied to solve real-world problems. Specifically, because the medical skills assessment occurred within a high-fidelity simulation of battlefield conditions, these findings provide evidence that AI-optimized spaced retrieval can enhance performance in modeling and simulation (MODSIM) contexts by increasing the likelihood that critical procedural knowledge is accessible under realistic, high-stress conditions. These findings build upon those discussed in Study 1, in which students were able to answer new multiple-choice and essay questions after studying similar questions via an SIR program for a month. In this case, the knowledge built through this intervention demonstrates even more durability and flexibility, as students who engaged with the same SIR intervention were more successful on a hands-on test of their ability to provide care under fire.

This study was limited by a small sample size and a small question bank of just 52 questions. Further, more research is needed to determine a causal relationship between engagement with the SIR intervention and assessment pass rates, including research in which study participants are randomly assigned to the intervention or not prior to their assessment. Engagement with the SIR intervention was voluntary in this study because schoolhouse leadership questioned the ethics of withholding a helpful learning tool from students performing high-risk tasks.

STUDY 3

Background and Research Objectives

In January of 2023, a large security firm deployed the SIR intervention previously described. The firm employs cleared personnel for roles as armed guards and security escorts in high-security, privately-owned facilities and data centers.

The firm required that employees engage with the SIR program for three brief sessions each week to refresh their standard operating procedures. Their dedicated engagement offered a unique longitudinal research opportunity. Approximately a year after they started the SIR intervention, I analyzed their data to determine whether the full workforce – both people with relatively high and relatively low starting knowledge levels – had accomplished knowledge mastery. Further, I asked the firm's CEO whether he observed any operational changes since they started using the SIR platform.

Method and Results

Beginning in January of 2023, employees at the data-center security firm described above started engaging with the SIR intervention three times each work week. Company leadership assigned a question bank of between 100 and 250 total questions to each employee, depending on their role. As examples, screening officers answered questions about proper use of handheld metal detectors and the risk level associated with different equipment labels, among other topics. During each SIR session, employees had to answer six questions from their assigned question bank. These six were chosen for each person by the platform's machine-learning algorithm, which determines for each end-user the questions that are at highest risk of being forgotten and therefore need to be reviewed most urgently. Employees were required to complete three sessions per week and were penalized (e.g., deemed ineligible for promotion) if they did not complete at least one session per week.

Approximately a year after the security guards started the SIR intervention, I examined the progress of 70 of these employees who demonstrated consistent engagement. I conducted a descriptive analysis to examine knowledge changes for employees with different starting levels of mastery by grouping the 70 employees as "high knowledge" or "low knowledge," based on whether they performed higher or lower than 76% (their median score) the first time they answered their assigned SIR questions. This resulted in 35 employees categorized as "high" and 35 categorized as "low" knowledge. I did this anonymously and only for the purpose of this research; information about specific employees was not shared with company leadership.

For each of these groups, I calculated employees' average knowledge accuracy on each day of engagement with the SIR intervention, from the first day to the 336th day. That is, on each day, I calculated each person's proportion of accurate answers for all their most recent responses to each of their assigned questions. For example, an employee's accuracy score on their 45th day of use was calculated by searching for their most recent viewing of each of their questions and calculating their proportion correct for those questions. I converted proportions to percentages for ease of interpretation.

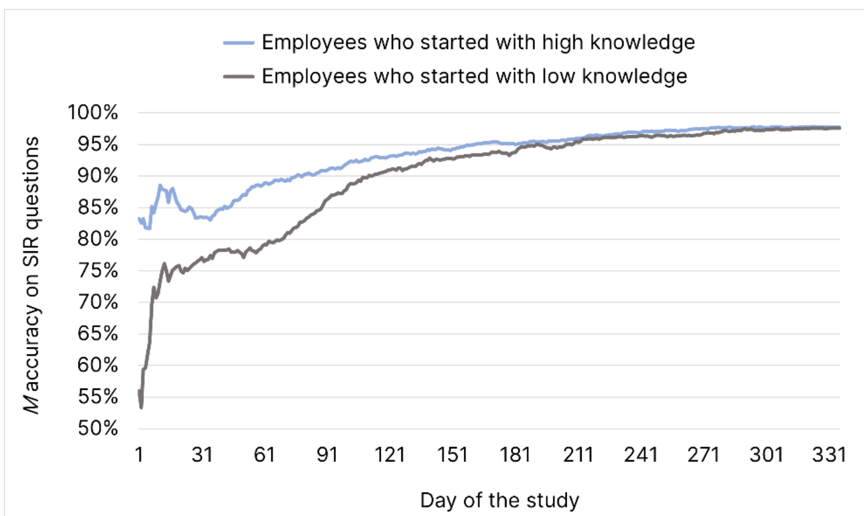

Figure 5. Knowledge Changes Over Time for Low and High Starting-Knowledge Employees Engaging with the SIR Intervention.

As shown in Figure 5, on the first day of the SIR intervention, low-knowledge employees (the lower half of the median split) averaged 55.95% (SD = 42.12%) accuracy and high-knowledge employees (the upper half of the median split) averaged 83.20% (SD = 33.34%) accuracy. These starting accuracy scores highlight two different knowledge gaps that are present in any workforce: the gap between employees who have different knowledge levels, and the gap between any employee and perfect 100% recall. As shown in Table 1, the gap between low- and high-knowledge employees closed shortly after the 180-day mark. With continued engagement thereafter, average employee knowledge levels continued to stay high and never decreased.

Table 1. Mean Knowledge Changes Over Time for Low Versus High Starting-Knowledge Employees Engaging with the SIR Intervention.

Day of the study	***M* question accuracy: Employees with low starting knowledge**	***M* question accuracy: Employees with high starting knowledge**	**Day of the study**	***M* question accuracy: Employees with low starting knowledge**	***M* question accuracy: Employees with high starting knowledge**
1	56.0%	83.2%	180	93.6%	95.1%
30	76.7%	83.6%	210	95.3%	95.8%
60	78.5%	88.4%	240	96.3%	96.9%
90	85.5%	90.8%	270	96.5%	97.4%
120	90.7%	92.8%	300	97.3%	97.7%
150	92.8%	94.1%	330	97.5%	97.8%

Last, I asked the CEO of the security firm if he could share any data regarding the number of reported workplace incidents that occurred before and one year after implementing the SIR intervention. Though specific data were confidential, he reported that the firm had experienced a 40% reduction in operational incidents since launching the intervention.

Discussion

This case study revealed that consistent use of an AI-optimized SIR tool can close knowledge gaps across a workforce and drive collective mastery, even among employees with initially low knowledge levels. Over the course of nearly a year, the 35 employees who began with below-median knowledge improved from an average accuracy of 56% to over 90%, achieving the same level of mastery as their higher-knowledge peers. This was accomplished with brief, focused review sessions three times a week, guided by the SIR platform's algorithm. Leadership at the firm reported a marked reduction in on-the-job errors during this period, suggesting that the knowledge gains observed translated into operational improvements. These findings emphasize the potential of AI-optimized SIR methods to both maintain knowledge and actively build and equalize it across a team, offering a scalable, sustainable solution to persistent training gaps in safety-critical environments.

Together with Studies 1 and 2, these results paint a comprehensive picture of the impact and flexibility of AI-optimized SIR interventions. While Study 1 demonstrated that students can learn entirely new content without lectures and still perform well on application-based assessments, and Study 2 showed that previously trained military personnel can recover lapsed knowledge to perform better in simulated combat scenarios, Study 3 extended these findings into the domain of professional operations. Importantly, Study 3 provides real-world evidence that SIR interventions can support large, diverse teams, and do so over the long term. Across all three studies, consistent patterns emerged: The SIR intervention reliably strengthened knowledge, reduced the time required to maintain it, and enabled the flexible application of that knowledge in test-based and operational settings alike.

GENERAL DISCUSSION

Every organization that requires annual re-training implicitly acknowledges a fundamental truth: memory decays over time. However, what is often overlooked is the speed and severity of that decline. As early as the 19th century, Hermann Ebbinghaus (1913) showed that memories fade at an alarming rate, and learning scientists have spent over a century developing interventions to combat this problem. Today, the scientific consensus is clear: Retention of important knowledge depends on regular, effortful recall through spaced and interleaved retrieval practice. To remember something long-term, we cannot simply learn it once. We must continually bring it back to mind across time and context. There is no final step in the learning process, only the ongoing work of remembering.

The three studies presented in this paper demonstrate how this principle can be operationalized at scale through AI-optimized SIR. Study 1 showed that undergraduate students who completed a program of optimized SIR to learn new information performed equally well on final applied tests for knowledge that was versus was not reinforced with a lecture. This finding suggests that, even in the absence of traditional instruction, optimized retrieval practice can produce strong learning outcomes. In Study 2, US Air Force pararescue students engaged with the SIR intervention to revive and reinforce previously taught medical content in preparation for a high-stakes skills assessment. The program helped them rapidly remaster their prior training, and students who used it were more likely to pass their assessment. Finally, Study 3 demonstrated that SIR can be deployed across a large workforce to close knowledge gaps and raise baseline expertise across diverse employees.

These results illustrate what is possible when organizations follow the wisdom of learning science. Rather than relying on infrequent training events followed by long periods of disuse, both the US Air Force and the security firm implemented SIR as a continuous reinforcement strategy. In doing so, they addressed the root problem—forgetting—and showed improved operational performance. These organizations exemplify a growing trend of translating decades of cognitive science into workflows that both improve individual- and team-level outcomes.

From a broader perspective, this research contributes to ongoing efforts to modernize how to train and sustain critical knowledge in high-stakes environments. Spaced and interleaved retrieval has long been the gold standard for promoting durable learning, but implementation challenges—such as time constraints, learner motivation, and metacognitive demands—have limited widespread adoption. AI-based SIR offers a solution by automating the individualization of study schedules and delivering effective practice in minimal time. In this way, AI can help resolve a key tension in organizational learning: the need to maximize retention without overburdening workers or instructors.

Last, the success of SIR-based interventions in simulation-heavy settings (e.g., Study 2's care-under-fire assessment) suggests compatibility with MODSIM environments. Integrating AI-driven retrieval with simulation-based training may enhance the realism and effectiveness of such programs by ensuring learners arrive cognitively prepared for immersive scenarios.

In summary, these findings make a case for integrating continuous SIR into organizational training, including MODSIM. Whether the goal is to teach new material, revive old content, or maintain readiness over time, SIR offers a flexible and evidence-based approach that works across contexts. As modern workplaces become increasingly complex and cognitively demanding, training practices must evolve accordingly. The studies reported here represent an important step in that direction, demonstrating how learning science can be brought to life in service of both individual growth and organizational success.

ACKNOWLEDGEMENTS

I would like to thank my colleagues Leamarie Gordon and Greg Hughes, who collaborated with me on Study 1. I would also like to thank the staff and students at the 351 SWTS at Kirtland AFB who made possible the work presented in Study 2, and leadership at the security firm who made possible the work presented in Study 3.

REFERENCES

Agarwal, & Roediger, H. L. (2018). Lessons for learning: How cognitive psychology informs classroom practice: Laboratory science and classroom observation reveal four simple strategies that can promote learning. (WHAT WE'VE LEARNED ABOUT LEARNING). *Phi Delta Kappan, 100*(4), 8–8.

Brown, P. C., Roediger, H. L., & McDaniel, M. A. (2014). Make it stick: The science of successful learning. Harvard University Press.

Carpenter, Pan, S. C., & Butler, A. C. (2022). The science of effective learning with spacing and retrieval practice. *Nature Reviews Psychology, 1*(9), 496–511. https://doi.org/10.1038/s44159-022-00089-1

Christian, M. S., Bradley, J. C., Wallace, J. C., & Burke, M. J. (2009). Workplace safety: A meta-analysis of the roles of person and situation factors. *Journal of Applied Psychology, 94*(5), 1103–1127. https://doi.org/10.1037/a0016172

Ebbinghaus, H. (1913). Memory: A contribution to experimental psychology (H. A. Ruger & C. E. Bussenius, Trans.). New York: Teachers College, Columbia University. (Original work published 1885)

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Gordon, L. T., Hughes, G. I., & Smith, A. M. (2025). Modernizing the Flipped Classroom: Replacing Lecture Time with Asynchronous Spaced Retrieval. *Teaching of Psychology*. https://doi.org/10.1177/00986283251325842

Koriat, A., & Bjork, R. A. (2005). Illusions of competence in monitoring one's knowledge during study. Journal of experimental psychology. *Learning, memory, and cognition, 31*(2), 187–194. https://doi.org/10.1037/0278-7393.31.2.187

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology, 77*(6), 1121–1134. https://doi.org/10.1037//0022-3514.77.6.1121

McHugh, D., Feinn, R., McIlvenna, J., & Trevithick, M. (2021). A random controlled trial to examine the efficacy of blank slate: A novel spaced retrieval tool with real-time learning analytics. *Education Sciences, 11*(3), 90. https://doi.org/10.3390/educsci11030090

Murre JMJ, Dros J (2015) Replication and Analysis of Ebbinghaus' Forgetting Curve. *PLoS ONE 10*(7): e0120644. https://doi.org/10.1371/journal.pone.0120644

Nowinski, J. L., Holbrook, J., & Dismukes, R. K. (2003). Human memory and cockpit operations: An ASRS study. In *Proceedings of the 12th International Symposium on Aviation Psychology* (pp. 888–893). Dayton, OH: Wright State University.

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language, 60*(4), 437–447. https://doi.org/10.1016/j.jml.2009.01.004

Suliburk, J. W., Buck, Q. M., Pirko, C. J., Massarweh, N. N., Barshes, N. R., Singh, H., & Rosengart, T. K. (2019). Analysis of human performance deficiencies associated with surgical adverse events. *JAMA Network Open, 2*(7), e198067. https://doi.org/10.1001/jamanetworkopen.2019.8067

Wynn, J. S., & Black, S. (1998). Evaluation of retention of safety and survival training content versus industry standard for training. *Air Medical Journal, 17*(4), 166–168. https://doi.org/10.1016/S1067-991X(98)90044-X