Enhanced STEM Subject Outcomes from the Use of Intelligent Tutors

Edward Harvey
Advanced Training & Learning Technology, LLC
Virginia Beach, VA
eharvey@atltgames.com

Emilie A. Reitz
Leidos
Norfolk, VA
eareitz@gmail.com

ABSTRACT

The military training and education community spends hundreds of hours a year on training warfighters not just for the job in front of them, but their next job, and contingencies foreseen and unforeseen. This training and education process requires a strong foundation of understanding Joint or Service procedures, and is dependent on each warfighter's area of specialization. There are highly technical skill-sets that can decay if unused or un-refreshed on a regular basis. Refresher training is often conducted as a standard process, untailored to each individual's unique skill-set. School age children are, in the end, no different in this learning challenge. A student's proficiency in foundational arithmetic and algebraic skills is critical for success in algebra and higher-level math courses.

Building on data from previous STEM-focused learning interventions, this paper will discuss a case study for applying diagnostic measures as a way to enhance and adapt future learning, focused on application of this technique for 7th and 8th grade foundational math; it will also propose integration techniques with current military data collection concepts such as xAPI. Application of this intervention process proved that use of the intelligent tutor makes for better grades, but also helps to develop and maintain those critical building blocks in the path to success in future STEM pursuits. The framework is currently populated with math content to create an Algebra Readiness Program, whose results will be discussed, but the framework is content agnostic so it can be used to develop effective teaching tools for other subject domains for which proficiency is dependent on learning and retaining declarative knowledge. As the STEM community puts these structured interventions to use, the extrapolation of these techniques to all stages of the military learning environment is crucial.

ABOUT THE AUTHORS

Ed Harvey is the President of ATLT, LLC, and resides in Virginia Beach, VA. ATLT develops educational programs based on serious gaming, simulation, and other advanced education and training technologies. His research interests focus on STEM education and the use of intelligent tutoring agents, automated assessments, and performance reporting to address the student math proficiency problem that is endemic in U.S. secondary schools, colleges, and universities. Ed recently set up an education non-profit, STEMprovement, which is organized to train schools how to use technology solutions to improve student performance on important math and science exams. He was the President and Director of Business Development for 22 years with BMH Associates, Inc., a Norfolk, VA training systems development company with 155 employees when BMH was acquired in 2006. Ed graduated from the Georgia Institute of Technology with a Bachelor of Chemical Engineering degree.

Emilie A. Reitz, M.A., is a Research Analyst at Leidos. She is currently supporting the Combined Fire Support Interoperability Division of Joint Staff J6, Deputy Director for C4 and Cyber Integration in Norfolk, Virginia. In this capacity, she is the data collection and analytical working group lead for Bold Quest. Her research focuses on integrating joint capabilities into modeling, simulation, and training, as performance enablers.

Enhanced STEM Subject Outcomes from the Use of Intelligent Tutors

Edward Harvey	**Emilie A. Reitz**
Advanced Training & Learning Technology, LLC	**Leidos**
Virginia Beach, VA	**Norfolk, VA**
eharvey@atltgames.com	eareitz@gmail.com

INTRODUCTION

During his or her lifetime, a single soldier could face multiple potential challenges relevant to the specific job or the hazards of the specific operating environment. Because of that high degree of risk, there is a sprawling education and training community preparing them to perform their skills and make decisions under often extreme circumstances. From recruiting to the end of the first six months, it costs an average of $58,000 to train and assess a single soldier for the U.S. Army (Niebuhr, Page, Cowan, Urban, Gubata, & Richard, 2013). This initial investment does not take into account that the military training and education community spends hundreds of hours a year on training each warfighter – not just for the job in front of them, but their next job, and contingencies foreseen and unforeseen.

This training and learning process requires a strong foundation of understanding Joint or Service procedures. Depending on each warfighter's area of specialization, there might be a highly technical skill-set which can quickly decay if unused or un-refreshed on a regular basis. This refreshing is often done in a blanket process, if it is offered at all, untailored to each individual's unique skill-set. The roots of this particular andragogic design are also in the beginning of any individuals' education process. School age children are, in the end, no different in their learning challenges from the soldiers some of them might eventually become. At all levels of the educational process, students of all ages face the decay of skills and knowledge that serves as the base on which new and increasingly more challenging topics are built.

While the experiment we performed focused on a student's proficiency in foundational arithmetic and algebraic skills, there is no reason that these gains shouldn't be extrapolated to other learning areas. Building on data from previous science, technology, engineering and mathematics (STEM) focused learning interventions (Harvey, Fuller, et al, 2015), this paper will discuss an experiment where we applied diagnostic measures as a way to enhance and adapt future learning to classes of 7th and 8th grade Algebra 1 students. It will additionally propose techniques for integration with current military data collection concepts such as xAPI.

KNOWLEDGE DECAY

Few people will disagree that practice makes perfect. The more a student practices a skill, the longer the student will maintain proficiency for that skill. Research shows that a student who practices enough will eventually master a skill (Bloom, 1985). Skill mastery is important because it indicates a base level of understanding that a student has achieved for the prerequisite knowledge necessary to learn subsequent information. How is mastery defined? Professional educators use a score of 90% on knowledge exams as the threshold for demonstrating mastery (Guskey, 2010). Echoing this standard, a score of 90% is the lowest score on Virginia state-mandated end-of-course (EOC) math exams a student must achieve to place into the highest "Advanced" proficiency category (Virginia Department of Education, 2017).

A driver for knowledge decay is that the amount of time it takes students to master a given skill varies from student to student. Student performance data on an important math exam, in this case the Algebra 1 EOC exam given to Virginia public school students, shows that 9% of Algebra 1 students (which is traditionally taken in grades 7, 8 or 9), place in the Advanced proficiency category with a score of at least 90% (Virginia Department of Education, 2017). This is a strong indication that at least nine out of ten students do not have not enough time to master Algebra 1.

Here is a simple example for how student success in Algebra 1 is dependent on their proficiency for a basic arithmetic skill of adding two fractions with unlike denominators. When given an Algebra 1 exam question that includes adding two fractions, in this case "1/7 + 1/2", too many students will enter "2/9" as the answer and get the exam problem wrong. They didn't get the problem wrong because they did not understand the algebraic concept being tested. They got the problem wrong because they lost their proficiency for a basic arithmetic skill of adding fractions. To correctly add "1/7 + 1/2"" they must know how to determine the Least Common Multiple (LCM) of 7 and 2, which is 14. (The correct answer to "1/7 + 1/2" is 9/14.)

Algebra 1 mastery is incredibly important since Algebra 1 represents the knowledge foundation necessary for success in the higher level math and science courses students take in high school that are prerequisites for college admission. Why do such a low percentage of students demonstrate Algebra 1 mastery?

The primary reasons include time constraints and the content delivery technique. There simply is not enough time in a school year for math teachers to dwell on each curriculum skill long enough for more than one-quarter of their students to master the skill. Teachers are required to move through a standardized curriculum at a pace necessary to cover every skill in the curriculum during the time available in the school year. This measurement of content taught on time is entirely independent of whether each student masters each skill or not. For the state of Virginia during the 2015-2016 school year, only 9% of students who took the Algebra 1 EOC exam, 13% of students who took the Geometry EOC exam, and 23% of students who took the Algebra 2 EOC exam received a score of at least 90% (Virginia Department of Education, 2017).

The predominant technique math teachers use in this time-constrained environment is to teach the same way to each student using conventional techniques of lecture to a passive group of students, practice using worksheets, and periodic assessments to measure student performance. This is essentially the same way math has been taught since math was invented (Allen, 2001). It remains in our classrooms as a set of techniques that has little integration of new techniques and technologies, despite the potential gains available by supplementing classroom learning with tutoring techniques and technology that have been demonstrated in the literature (VanLehn, 2011).

TESTING AN INTELLIGENT TUTOR INTERVENTION

The structure of the intervention we tested is simple, regardless of the solution used to reach that structure. It consisted of a pre-test to establish student understanding of the content. The pre-test was followed by remediation on areas of weakness. After remediation, students took a post-test to affirm knowledge newly re-mastered, prior to taking the EOC exam. We elected to solve this student math proficiency problem by developing a scalable, software-based training system that could be used to mitigate the effects of poor long-term retention to improve the quality of conventional in-class mathematics instruction in a non-disruptive manner, using techniques that have already proved to be effective for military procedural skills training.

Any solution would have to supplement in-class instruction, since we found from our research that due to time constraints, students are only given one opportunity in class to master math content. Those that "get it" the first time do well. Those who do not "get it" the first time fall behind, and until now there were no subsequent opportunities for them to catch back up.

There is so much material math teachers must teach during the school year there is little time available for an effective review prior to the EOC exam. Schools do provide remediation to the lowest performing students but that is mainly done to prevent those students from failing the EOC exam. Surprisingly, the EOC exam failure rate is the sole metric used to determine if a Virginia middle school meets the minimum accreditation standard for math[1]. So, schools are compelled to devote available remediation resources to students who are at-risk for failing the EOC exam, which leaves little, if any, resources to help the majority of students regain proficiency for any skills that decayed to prepare for the EOC exam. The main problem with the civilian education system is there is not enough time available in a school year for teachers to tailor content delivery to each student's specific needs. The result is too many students are unable to demonstrate proficiency on important math exams for much of what they were

---

[1] As long as less than 30% of students taking a math EOC exam fail the exam with a score less than 50%, a school is accredited in Virginia.

previously taught – amplifying the need for remediation to be built into the structure of the intervention to maximize student performance at essentially no cost in time or effort by schools.

There are multiple reasons for developing an intelligent tutoring system, whether it is used to address military training challenges or K-12 math comprehension challenges. The most obvious is the immediate ability to make a significant positive impact on student exam performance (Eagle, Corbett, Stamper, McLaren, Wagner, MacLaren, & Mitchell, 2016). One recent study of two computer-based tutoring systems (Carnegie Learning Algebra Cognitive Tutor, and ALEKS Algebra Course) revealed an effects size d=1.10 between the two systems in a series of controlled experiments. This effects size impact was additionally supported by positive student feedback about using the systems (Sabo, Atkinson, Barrus, Joseph, & Perez, 2013). And as mentioned above, there is not enough time available for today's students to master the concepts and skills being taught to them through use of conventional classroom teaching techniques. In math proficiency, U.S. students are ranked an unimpressive 38th out of 71 countries on the Program for International Student Assessment (PISA), which ranks math proficiency by country every three years (Desilver. 2017). With numbers like this, not enough U.S. students will be prepared to successfully pursue a college STEM degree when they graduate from high school (Shulock, & Callan, 2010).

With respect to the quality of civilian math instruction, the current status quo is pretty dire. Only one out of eleven Virginia public school students who take Algebra 1 place in the top "Advanced" proficiency category with a score of at least 90%, on the EOC exam. Only one out of four U.S. 12th grade students who took the quadrennial National Assessment of Education Progress (NAEP) exam in 2015 performed at or above the Proficient level for math. This level of performance is statistically the same as for 2005 and slightly lower than 2009. Math proficiency for six out of ten U.S. high school graduates who attend college has decayed enough that they must take non-credit developmental math courses that repeat math courses they completed in middle school before they can take the college-level math necessary to meet their degree requirements (The National Center for Public Policy and Higher Education and the Southern Regional Education Board., 2010). Seventy-eight percent of U.S. citizens (adults) cannot explain how to compute the interest paid on a loan, seventy-one percent cannot calculate miles per gallon on a trip, and fifty-eight percent cannot calculate a 10% tip for a lunch bill (Phillips, 2007).

METHODOLOGY

How is it possible to measure knowledge decay? In classrooms, teachers use quizzes and tests to measure student comprehension for recently taught material. Math teachers sum quiz and test results for the school year to generate each student's final course grade. The course grade most reflects student "short-term" proficiency. Students also take an EOC exam at the end of the school year that reflects how well they retained proficiency for the skills they were taught over the span of the current school year. And, in the case of math courses, EOC exams also test student aptitude for the foundational arithmetic skills they were taught in previous school years.

To test the intelligent tutor intervention, we were granted access to two sets of 7th and 8th grade students who had or were just completing their Algebra 1 course (Table 1). One group completed a pre-test after completion of their EOC exam, but did not receive the intervention. The second group completed a pre-test, the intervention, and the post-test prior to taking the EOC exam. Overall course grades for the year were made available for both groups. What follows is a discussion of the specific testing and intervention process, and the results of the comparison.

Table 1. Experimental Structure.

Control group (n=54)	Course Grades	EOC Test	Intelligent Tutor Pre-Test		
Experi-mental Group (n=33)	Course Grades	EOC Test	Intelligent Tutor Pre-Test	Intelligent Tutor Intervention	Intelligent Tutor Post-Test

Structure Behind the Interventions

The two primary system components are the subject content repository and the rules associated with delivery of the subject content. The subject content includes traditional video lessons and narrative text lessons, and participatory guided practice problems that emulate a human tutor.

Students who used the system to regain lost proficiency completed an assessment pre-test. The pre-test aligns to the curriculum at the skill level and is designed to predict EOC exam outcome. Pre-test results are used to generate a custom remediation plan that lists each skill lesson the student should complete to refresh his proficiency in the minimum amount of time. The intelligent tutor monitors student performance and automatically intervenes by directing the student to a set of learning assets upon detection of a skill deficiency. Skill-level learning assets include video lessons, narrative text lessons, and guided practice problems. The guided practice problems, which require active student participation, are the primary learning asset because it takes students less time to complete a guided practice problem to regain skill proficiency compared to watching a video or reading a narrative text lesson. Students can ask the intelligent tutor for help any time they are not confident they know the correct answer to an assessment question. Students who learn to correctly self-assess their skill level proficiency will ask for help only when necessary, and as a result, will seldom receive an intervention from the intelligent tutor. This mechanization allows each student to regain his proficiency in the minimum amount of time since students only need to spend time remediating their specific set of deficient skills.

The following flow chart is the best way to visualize the content-agnostic structure, or framework, for the system (Figure 1).
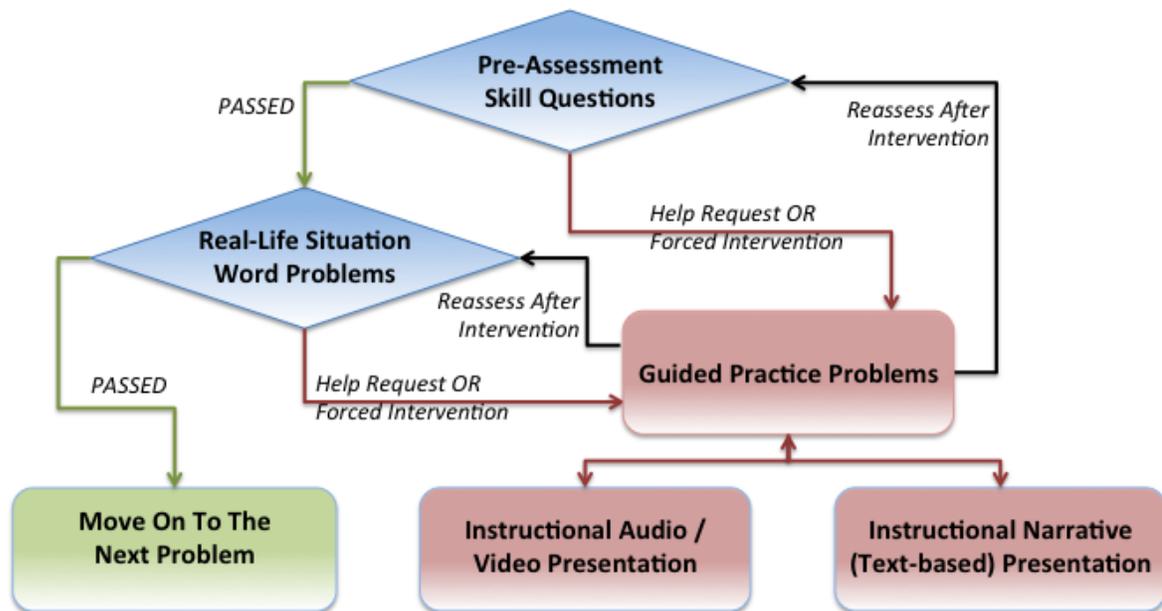
Figure 1. Intelligent Tutoring System Framework.

Students demonstrate their proficiency by answering a set of assessment questions for each skill lesson. There are 5,536 assessment questions for 130 arithmetic and algebraic skills in the Algebra Readiness Program. The reason there are approximately forty assessment questions for each skill, even though typically only five assessment questions are normally required to determine skill proficiency, is the majority of students typically use trial and error (guessing). Students quickly learn that trial and error doesn't work well, and the more effective and time-efficient learning strategy is to ask for help when they are not confident they know the correct answer to an assessment question.

In this remediation process, a student who correctly answers the assessment questions moves to the next skill listed in their remediation plan. A student can ask the intelligent tutor for help when he is not confident he can answer an assessment question correctly. He is presented with a list of learning assets he can use to refresh his proficiency for the skill before returning to the skill assessment questions. When a student receives a score of less than 80% during the skill assessment, the intelligent tutoring agent intervenes and presents a list of learning assets the student can use to refresh his proficiency for the skill.

Detailed performance data is generated and archived on a secure server as a student progresses through a curriculum. This data is processed and presented to the student and to authorized individuals in color-coded and tabular report formats to show student progress and performance. These reports are designed to show skill proficiency and completion time as well as the learning strategy a student uses to complete each skill. The following graphic is a sample progress/performance report for a group of college freshmen STEM degree students (Figure 2).

Color Code 0/0 Blue: Passed with 0 Tutor Interventions. 1/0 Green: Passed with NO More Than 2 Tutor Interventions. 3/0 Yellow: Passed with 3 or Mo[re]
Numbers in cell indicate the "number of tutoring interventions" / "number of help requests".

		Whole Numbers					Integers				Fractions & Decimals						
		Properties	OOO	Div/List Fac	LCM/GCF	Var/Expr	Intro	Add Like	Comb	Mult/Div	Frac ID/ Classify	Frac Convert/ Compare	Frac Add/Sub Like	Frac Add/Sub Unlike	Frac Mult/Div	Dec Add/Sub	Dec Mult/D
ID	**Name**	1.1	1.2	1.3	1.4	1.5	2.1	2.2	2.3	2.4	3.1	3.2	3.3	3.4	3.5	3.6	3.7
1418034	1418034	0/0	0/0	0/0	0/0	0/0	2/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0
1418012	1418012	0/1	0/2	0/0	0/1	2/1	1/0	2/1	1/0	0/2	2/1	2/1	1/0	1/0	2/0	0/0	2/0
1418086	1418086	1/1	0/0	0/0	0/0	0/0	1/0	4/0	1/0	0/0	4/1						
1418000	1418000	0/0	0/0	0/0	5/1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0	1/0	0/0	0/0	0/0
1418008	1418008	0/0	0/0	0/0	0/0	0/0	2/0	0/0	0/0	0/0	9/0	0/0	1/0	0/0	1/0	0/0	0/0
1418016	1418016	0/0	0/1	0/0	5/0	0/0	0/0	2/0	0/0	0/0	3/0	1/0	1/0	0/0	2/0	0/0	0/0
1418036	1418036	1/3	0/0	0/0	4/1	0/0	0/1	0/0	0/0	0/0	2/1	0/0	1/0	0/0	0/0	0/0	0/0
1418028	1418028	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0	5/0	0/0	2/0	0/0	1/0	0/0
1418032	1418032	1/1	0/0	0/1	2/3	0/0	1/0	0/0	0/0	0/1	0/2	0/2	1/2	3/0	1/4	0/0	0/0
1418040	1418040	4/0	1/0	1/0	1/0	3/0	0/1	0/0	0/0	0/0	5/0	0/0	0/0	0/0	0/0	0/0	0/0
1417978	1417978	1/0	1/0	0/0	0/0	0/1	0/3	0/0	0/0	0/0	6/0		2/0	0/0		0/0	0/0
1417982	1417982	0/0	0/0	0/0	1/0	0/0	3/0	3/0	0/0	0/0	16/0	0/0	1/0	0/0	0/0	0/0	1/0
1418113	1418113	7/1	0/0	0/0	1/0	0/0	2/0	2/0	0/0	0/0	1/0	0/0	2/0	4/0	1/0	0/0	0/0
1418042	1418042	14/1	1/0	0/0	4/0	0/0	5/0	2/0	0/0	0/0	13/0	3/0	1/0	2/0	0/0	0/0	0/0
Averages for Each Level																	
Average Num. Interventions / Help Requests		2/0	0/0	0/0	3/0	0/0	1/0	1/0	0/0	0/0	9/0	1/0	4/0	4/0	1/0	0/0	0/0
Average Time		00:04:46	00:04:27	00:07:39	00:09:07	00:04:47	00:04:09	00:05:20	00:04:07	00:02:23	00:07:13	00:04:42	00:08:18	00:08:25	00:08:38	00:04:38	00:03:5
Average Time In Tablet		00:00:24	00:00:01	00:00:09	00:01:01	00:00:02	00:00:06	00:00:05	00:00:00	00:00:03	00:00:30	00:00:18	00:00:12	00:00:05	00:00:24	00:00:00	00:00:2

Figure 2. Sample progress/performance report for a group of learners.

RESULTS

Table 2. Experimental Results.

	Grade Level	Course Grade (Avg)	Course Grade (Std.Dev)	EOC Exam (Avg)	EOC Exam (Std.Dev)	Intelligent Tutor Pre-Test (Avg)	Intelligent Tutor Pre-Test (Std.Dev)	Intelligent Tutor Post-Test (Avg)	Intelligent Tutor Post-Test (Std.Dev)
Control Group (n=54)	8th Grade	85.39	6.14	64.67	9.46	69.43	9.31		
Experimental Group (n=33)	7th Grade	89.55	6.21	85.59	9.52	83.90	9.01	95.48	4.17

The average course scores are presented in table 2, above, for both the control group and the group which received the intervention.

Figure 3 (below) is a quartile plot that compares the course grade, the EOC exam grade, and a diagnostic test score for the two classes of 8th grade Algebra 1 students (Control Group). The difference between the course grade and the EOC exam score provides a useful measure of the knowledge decay experienced over the course of a school year. A student may have demonstrated proficiency for that skill on a quiz completed close to the point of learning, but later in the school year when the student takes the EOC exam, his proficiency for that skill may have decayed. It is easy to see the course grade for this group of students does not predict the EOC exam score. The course grade indicates that one-quarter of the class should score high enough on the EOC exam to place into the Advanced Proficiency category with a score of at least 90%, when the highest score on the EOC exam is 84%. In fact, the median EOC exam score for these students is 21% lower than the median course grade.
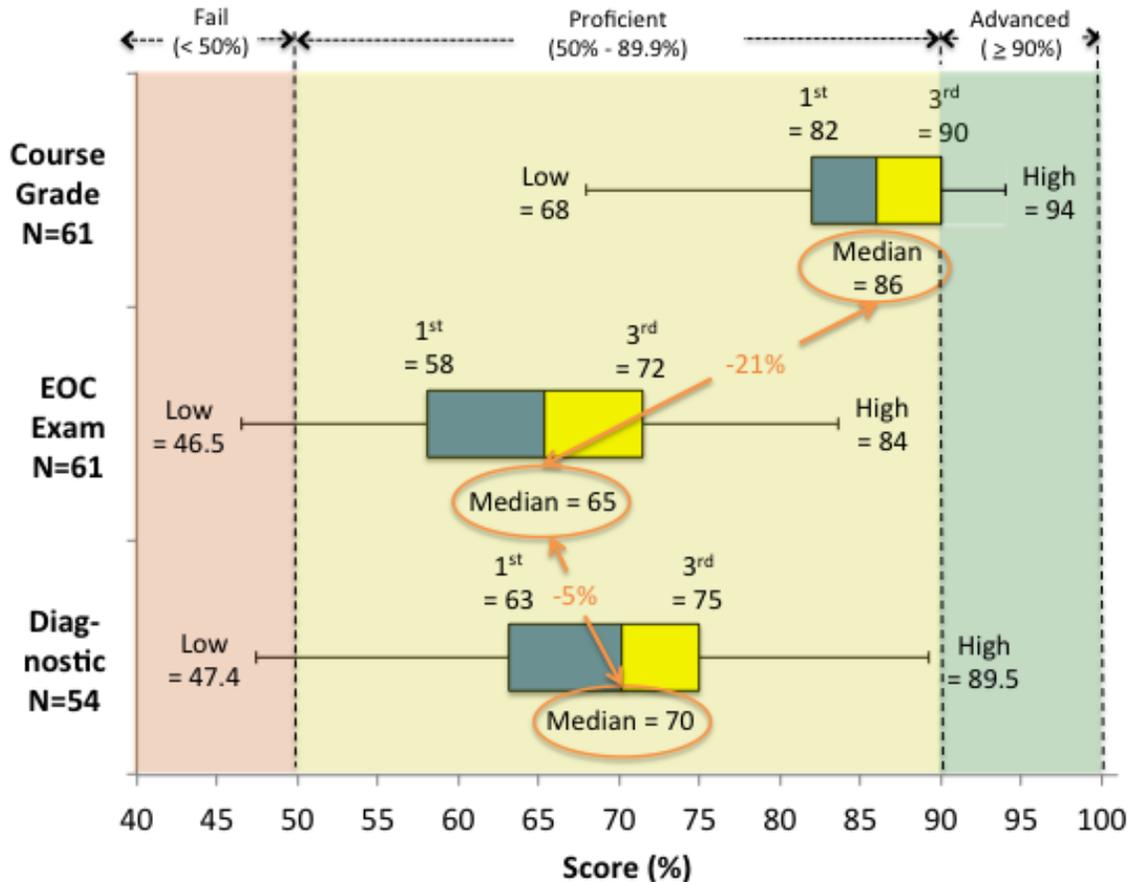
Figure 3. Control Group: Course Grade to EOC Exam Score Comparison with Diagnostic Test

A diagnostic test was given to this group of students to measure their current math proficiency, independent of their course grade and EOC exam score. The diagnostic test covered arithmetic and pre-algebraic concepts and skills that students must master to perform well in Algebra 1. The median score for the diagnostic test is within 5% of the median EOC exam score. In this case the diagnostic test was a much better predictor of student proficiency than the class course grade. This example is not an isolated case. Other course grade and EOC exam data available to us show the median EOC exam score is consistently 8 percentage points less than the median course grade for 7th grade Algebra 1 students, and 25 percentage points less for 8th grade Algebra 1 students.

Figure 4, is a quartile plot for our Experimental Group, a class of 7th grade Algebra 1 students, which includes the course grade, diagnostic test results, diagnostic post-test results, and EOC exam scores. The difference between this figure and Figure 3 is that this class of students took the diagnostic test and completed the intervention prior to taking the EOC exam. Their diagnostic test results were then used to generate a custom remediation plan every student in this class completed using the web app version of ATLT's Algebra Readiness Program. After completing their remediation plan, they completed a diagnostic post-test to verify they had successfully remediated each skill deficiency.
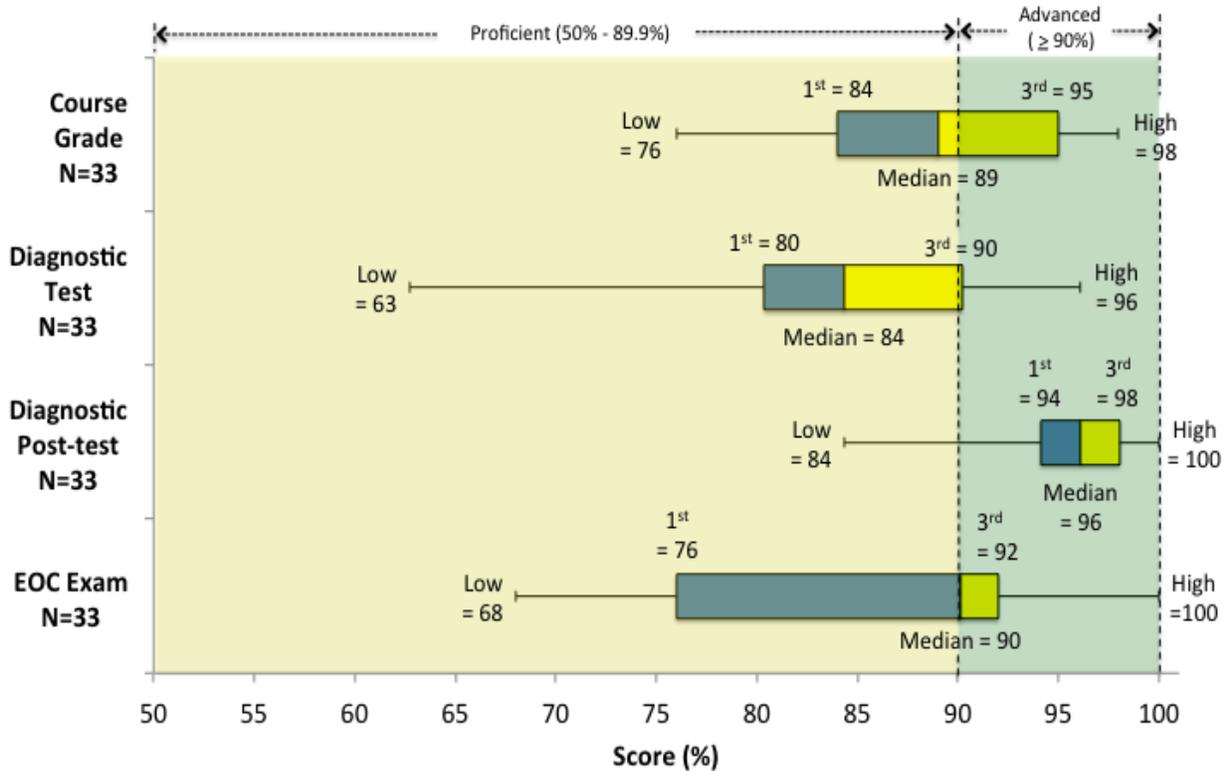
Figure 4. Experimental Group. Intervention to Mitigate Knowledge Decay.

The median course grade of 89% for this class indicated that almost one-half of the class would score in the top Advanced proficiency category on the EOC exam. However, the diagnostic test results indicated only one-quarter of the class would score in the Advanced proficiency category on the EOC exam. Students then completed their custom remediation plans that were generated based on diagnostic test results. The average remediation plan completion time was 61 minutes. This quick "reminder" where each student refreshed their proficiency for decayed skills that were indicated by their diagnostic test results led to a twelve percentage point improvement in the median score for the diagnostic post-test score over the diagnostic "pre-test" (for an Cohen's d of 1.65).

When we compare the course grades to the EOC exam scores, there is virtually no difference at the high end of the performance spectrum. The course grade predicted that almost one-half of this class would score in the Advanced category on the EOC exam, and after a structured, computer based intervention, one-half (52%) of the class did score in the Advanced category. This is a 48% increase in the number of students who placed into the Advanced proficiency category for the EOC exam compared to the other 7th grade Algebra 1 students who did not use the Program to prepare for taking the EOC exam. These results repeated for the 2016-17 school year when 74% more 7th grade Algebra 1 students who used the Program to prepare for the EOC exam placed into the Advanced proficiency category compared to the 7th grade Algebra 1 who did not use the Program. This example shows that it does not take much time to employ a very effective differentiated learning technique that supplements classroom instruction to allow each student to refresh his or her proficiency for any skills that might have decayed due to poor long-term retention. This minimizes, if not eliminates altogether, knowledge decay for the majority of students.

BRINGING IMPROVEMENTS TO A LARGER AUDIENCE

This is just one narrow use case performed as an experiment. In future iterations, we hope to extrapolate and leverage this system in more scenarios, as interest in intelligent tutoring techniques continues to grow in the educational and military arenas. The framework presented in this paper is currently populated with math content to create an Algebra Readiness Program, whose results are discussed, but the framework and software itself (Figure 1) is content agnostic. It can be used to develop effective teaching tools for other subject domains for which

proficiency is dependent on learning and retaining declarative knowledge. It can reduce the time necessary to train to standard, or quickly return warfighters to the standard they had previously attained.

As the STEM community puts these structured interventions to use, the extrapolation of these techniques to all stages of the military learning environment is crucial. Just as blended learning and more thorough integration of structured assessments at all stages of a warfighter's training cycle has been adopted from industry and civilian best practices, so too should the benefits of intelligent tutor-based interventions. In situations where students are deemed deficient from the start, interventions can and should be planned to increase student learning. Any intervention, regardless of system, should be well structured and tailored to provide multiple assessment points (Guskey, 2010). This technique aligns with ongoing Department of Defense (DOD) initiatives to improve the education and training environment, by encouraging adoption of the best technologies and important structural changes necessary to create ubiquitous learning environments (for more information on these efforts, read Folsom-Kovarik & Raybourn, 2016; or Schatz, Fautua, Stodd, & Reitz, 2015).

And, there are of course, a growing world of intelligent tutors in existence that narrowly focus on one crucial skillset. Different techniques exist for addressing these challenges; in this paper we presented just one set of data and one possible technique from a broad world of options (for more information on other systems, see Sottilare, & Boyce, 2016; Kulik, & Fletcher, 2016). Military systems such as the Army Research Lab's Generalized Intelligent Framework for Tutoring (GIFT) illuminate the potential for employing these techniques against challenging topics in the broader DOD environment (Sottilare, Brawner, Sinatra, & Johnston, 2017). The challenge will be to not create yet another stove-piped system, but to integrate these concepts into current and future learning management systems and larger data unifying protocols such as xAPI and Caliper (Haag, 2016). In the coming years, we hope to provide data exploring application of this system against a wider range of more conceptually-based topics, such as the application of *Joint Operation Planning,* or *Design and Planning.*

CONCLUSIONS

The 7th grade class who participated in this study achieved EOC exam scores that matched to their performance throughout the school year, rather than suffering the knowledge decay demonstrated by the 8th grade classes that participated in the study. A small thing, little over an hour of remediation to refresh the students on topics they'd forgotten, not only improved EOC exam scores but also prepared students to reapply that knowledge in their next math class.

The use of intelligent tutors is not "new" - very little in the education community is – but incremental design changes can demonstrate not just how using an intelligent tutor results better grades, but how it also helps to develop and maintain those critical building blocks on the path to success in future STEM pursuits. The current civilian education system is underperforming with respect to its success for teaching critical STEM content to students. It doesn't have to. Training soldiers and keeping them to a high degree of readiness is a high cost endeavor and manpower intensive. It doesn't have to be. By leveraging flexible assessment frameworks, and making them low-effort both for the educator and the student, we can reduce the time necessary and the costs associated with maintaining training readiness. But these systems must be treated as part of the educational process to mitigate the impact of poor long-term retention.

ACKNOWLEDGEMENTS

REFERENCES

Allen, J. (2001). Ancient Egyptian Science: A Source Book. Volume 3: Ancient Egyptian Mathematics by Marshall Clagett.

Desilver. D. (2017). U.*S. students' academic achievement still lags that of their peers in many other countries.* Retrieved May 2017 from: http://www.pewresearch.org/fact-tank/2017/02/15/u-s-students-internationally-math-science/

Eagle, M., Corbett, A., Stamper, J., McLaren, B. M., Wagner, A., MacLaren, B., & Mitchell, A. (2016, June). Estimating individual differences for student modeling in intelligent tutors from reading and pretest data. In *International Conference on Intelligent Tutoring Systems* (pp. 133-143). Springer International Publishing.

Folsom-Kovarik, J.T. & Raybourn, E.M. (2016). Total Learning Architecture (TLA) Enables Next-generation Learning via Meta-adaptation. *Proceedings of the I/ITSEC*. Arlington, VA: NTSA.

Guskey, T. R. (2010). *Lessons of mastery learning*. Educational leadership, 68(2), 52.

Haag, J. (2016). *ADL Experience API (xAPI) and IMS Caliper Discovery Review*. Retrieved May 2017, from https://www.adlnet.gov/adl-experience-api-and-ims-caliper-discovery-review

Harvey, E. Jr, Fuller, M., Harvey, E, III. (2015). The Secret for STEM Success: Employing Technology for Math Proficiency. *Proceedings of the I/ITSEC*. Arlington, VA: NTSA.

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: a meta-analytic review. Review of Educational Research, 86(1), 42-78.

National Assessment of Educational Progress. (2017) The Nation's Report Card. Retrieved May 26, 2017, from https://www.nationsreportcard.gov/dashboards/report_card.aspx

The National Center for Public Policy and Higher Education and the Southern Regional Education Board. (2010).*Beyond the Rhetoric, Improving College Readiness Through Coherent State Policy.* June 2010, http://www.highereducation.org/reports/college_readiness/CollegeReadiness.pdf

Niebuhr, D. W., Page, W. F., Cowan, D. N., Urban, N., Gubata, M. E., & Richard, P. (2013). Cost-effectiveness analysis of the US Army Assessment of Recruit Motivation and Strength (ARMS) program. *Military medicine*, *178*(10), 1102-1110.

Phillips, Gy W. (2007). Chance Favors the Prepared Mind: Mathematics and Science Indicators for Comparing States and Nations, American Institutes for Research: Washington, DC.

Sabo, K. E., Atkinson, R. K., Barrus, A. L., Joseph, S. S., & Perez, R. S. (2013). Searching for the two sigma advantage: Evaluating algebra intelligent tutors. Computers in Human Behavior, 29(4), 1833-1840.

Schatz, S., Fautua, D., Stodd, J. & Reitz, E. A. (2015). The Changing Face of Military Learning. In Doughty, R., Wells II, L., and Hailes, H. (Ed.), *Innovative Learning: A Key to National Security* (pp. 35-50). Fort Leavenworth, Kansas: The Army Press.

Shulock, N., & Callan, P. M. (2010). Beyond the rhetoric: Improving college readiness through coherent state policy. *San Jose, CA: National Center for Public Policy and Higher Education and Southern Regional Education Board*. http://www.highereducation.org/reports/college_readiness/CollegeReadiness.pdf

Sottilare, R., Brawner, K., Sinatra, A. & Johnston, J. (2017). An Updated Concept for a Generalized Intelligent Framework for Tutoring (GIFT). Orlando, FL: US Army Research Laboratory. May 2017. DOI: 10.13140/RG.2.2.12941.54244.

Sottilare, R. A., & Boyce, M. W. (2016, July). Elements of Adaptive Instruction for Training and Education. In *International Conference on Augmented Cognition* (pp. 85-89). Springer International Publishing.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. Educational Psychologist, 46(4), 197-221.

Virginia Department of Education. (2017). School Report Card. Retrieved April 15, 2017: https://p1pe.doe.virginia.gov/reportcard/