K-12 Center at ETS

Invitational Research Symposium on
Technology Enhanced Assessments

Integrating Assessment Within Instruction: A Look Forward

Neil T. Heffernan and Kenneth R. Koedinger

May 7–8, 2012

Integrating Assessment Within Instruction: A Look Forward

Neil T. Heffernan

Worcester Polytechnic Institute

Kenneth R. Koedinger

Carnegie Mellon University

Executive Summary

Teachers are in a bind: They are being asked to assess students to drive instruction, but every minute spent on assessments is a minute lost to instruction. Technologies that help teachers collect assessment information as students are engaged in problem solving, reasoning, and learning offer a powerful alternative. Technology enhanced assessment (TEA) promises many potential benefits. For example, they can be designed such that students can continue learning while they are being accurately assessed. In contrast to multiple-choice tests, TEAs also allow for assessment of richer constructs (e.g., you can ask a student to demonstrate the scientific method while interacting within a simulated lab). TEAs allow assessment results to be quickly communicated back to stakeholders (teachers, administrators, parents, and students) so they can take immediate action to adapt to students' needs. We report results showing a high correlation (R = 0.8) between a TEA system we developed and students' state test scores. Conventional wisdom suggests that when students are allowed to learn during an assessment, the assessment results will not be reliable (because it is harder to hit a moving target). However, we have evidence that data from student learning experiences, in an appropriately designed TEA, can enhance rather than reduce the reliability and validity of an assessment. Such assessments provide information not only about whether students get questions right or wrong, but also on how much instructional assistance students need before getting a question right. This extra information facilitates more accurate assessment. We also discuss findings indicating increased learning from the use of these TEAs by students and teachers. We end with a discussion of what the future classroom may look like using examples of changes we have observed in the practices of teachers who are using our TEA.

Introduction—How Assessment Has Challenged Teachers

Teachers know that using data on student learning can help students learn more, but every moment spent assessing students is a moment of instruction lost. And when teachers do collect data, what do they do with that data? A RAND Corporation study analyzed data-driven decision-making in school districts and noted how difficult it is to put into practice (Dembosky, Pane, Barney, & Christina, 2006). Teachers and administrators all agree that formative assessment (cf. Black & Wiliam, 1998) is a good thing, but teachers need support to put the data to good use—both support in gathering the data, so they can use the data, and coaching support on what to do with the data.

Teachers who use data to inform their teaching are more effective; however, educators often feel inadequate and threatened by the use and publication of achievement data (Earl & Katz, 2006; U.S. Department of Education, 2003). Making matters worse, educators are provided with little training in the use of data (Cizek, 2000; Creighton, 2001; Holcomb, 2004; Popham, 2008).

Initially, the data teachers are encouraged to use are primarily summative in nature, coming at the end of a course or school year. Summative assessments provide limited information that cannot be used easily by classroom teachers to inform their instruction. Additionally summative assessment data are rarely timely, often being released *months after* the assessment administration or after the students have moved on to another course of study. Increasingly, schools are turning to formative assessments to gain more frequent feedback on students' learning—they are seeking "biopsy" data that are still actionable and not "autopsy" data that are collected after the fact (Militello & Schweid, 2008). Formative assessments offer the potential to link curriculum, instruction, and assessment by providing timely feedback for teachers and administrators (Popham, 2008). What remain missing are diagnostic tools to inform teachers' everyday actions, tools that are locally controlled by the teacher.

In this paper, we use our technology enhanced assessment (TEA) system, ASSISTments™, as a frequent example. ASSISTments is a free web-based TEA that provides individualized tutoring to students and live assessment data to teachers. The fundamental idea, reflected in the name, is that of combining instructional *assist*ance with the collection of standards-based assess*ment* information. ASSISTments allows teachers to achieve these fundamental goals of education at the same time, without wasting student's time. More than 10,000 students across the United States and overseas are using ASSISTments in the current, 2011-12, school year. ASSISTments currently contains many middle school math and science learning and assessment tasks, but it is a general platform for authoring and delivery, which can be used for any content area (e.g., we have used it in graduate courses on artificial intelligence (AI) and educational research methods).

One of the most powerful aspects of ASSISTments is the instantaneous reports to teachers and students. Teachers can retrieve data on student performance instantly and use that data to inform their teaching. Reports provide teachers with the concepts and skills students need to solve problems, common wrong answers, options to individualize instruction, and ways to contact parents with a click of the mouse—all of this while giving feedback and help where needed to the student. Administrators and decision-makers also have access to up-to-the-minute data on student learning.

The next section highlights the importance of assessment at the level of the cognitive components, mental concepts and thinking skills, for guiding instructional decision-making. The following three sections discuss the assessment, learning, and research value of TEAs. Finally, we discuss the future of TEAs, including thoughts from teachers using cutting-edge TEAs now.

Cognitive Science Behind Standards

State and national standards documents, such as the Common Core State Standards, provide general descriptions of what students should achieve, but they do not specify in detail the components of knowledge (the concepts, procedures, principles, and thinking skills for each skill) that students must learn to meet these particular standards (cf. Koedinger, Corbett, & Perfetti, in press). For instance, consider number sense standards (5.NF) that indicate activities to be mastered, such as adding, subtracting, multiplying, and dividing fractions (e.g., “apply and extend previous understandings of multiplication and division to multiply and divide fractions”). Cognitive analysis of such *observable* activities typically reveals many potential *unobservable* concepts and skills that students may or may not learn as they become proficient in these standards (e.g., a deep fraction concept integrated with visual representations of quantity or a shallow representation of numbers separated by “/”; robust procedures for finding common multiples or equivalent fractions or shallow procedures, such as “invert and multiply”). Exactly what students are learning, that is, what changes in their minds, is largely unknown. But it is these unobservable components of knowledge and skill, not the activities, that are the critical outcome of education. They are what students’ minds carry or “transfer” outside the classroom, and they determine, and sometimes delimit, student future success in academics and in the workplace. Identifying the specific nature and the character (robustness, abstractness, reusability) of the knowledge behind standards achievement, that is, identifying exactly what students are learning, is a matter of scientific discovery.

Before describing how TEAs can help in such discovery, we provide an example of how cognitive analysis of data from carefully designed assessment tasks can reveal surprises about the nature of student learning and how to improve it. Take the three assessment tasks in Table 1. Heffernan and Koedinger (1998) found a result that contradicted the prevailing thinking that using variables was the critical part of learning algebra. If this were true, then the Symbolization problem (P1) would have been much harder than the Articulation problem (P2); but it was not. This was found in data produced by giving these items to students to solve. This set of results can be explained by hypothesizing that the real difficulty with word problems is not the comprehending of the English or the variables; it is about learning to write procedural “sentences” (with multiple arithmetic operators) in the “foreign language” of algebra symbols. These results are interesting because the current hypothesis about what caused difficulties was contradicted by the data. This research led to a new, more effective way to help students learn the important, but difficult, skills of expressing problems in mathematical symbols (Koedinger & McLaughlin, 2010).

Table 1. One of the Eight Cover Stories Used in the Studies With the Three Different Variants

Version	Sample answer	Sample question
P1 (Symbolizing) Accuracy 25%	800 – 40m	Ann is in a rowboat in a lake that is 2,400 yards wide. She is 800 yards from the dock. She then rows for "m" minutes back toward the dock. Ann rows at a speed of 40 yards per minute. Write an expression for Ann's distance from the dock.
P2 (Articulation) Accuracy 26%	800 – 40 * 3	Ann is in a rowboat in a lake that is 2,400 yards wide. She is 800 yards from the dock. She then rows for 3 minutes back toward the dock. Ann rows at a speed of 40 yards per minute. Write an expression for Ann's distance from the dock.
P3 (Computation) Accuracy 70%	680	Ann is in a rowboat in a lake that is 2,400 yards wide. She is 800 yards from the dock. She then rows for 3 minutes back toward the dock. Ann rows at a speed of 40 yards per minute. Compute Ann's distance from the dock.

So this element of research is important because, if we rely on the intuition of assessment designers, standard setters, and textbook writers, there is a risk of them getting it wrong. Nathan and Koedinger (2000) asked math teachers and education researchers which kinds of problems they expected to be the most difficult for students. They showed that both were unlikely to correctly predict that equations would be harder than story problems and that teachers who knew the most about the topic did the worst at predicting student difficulties. As shown in Figure 1, high school teachers, who have the most expertise in algebra, were the least likely to be correct. Their expertise with equations created a “blind spot” regarding seeing how difficult learning to use equations is for students.

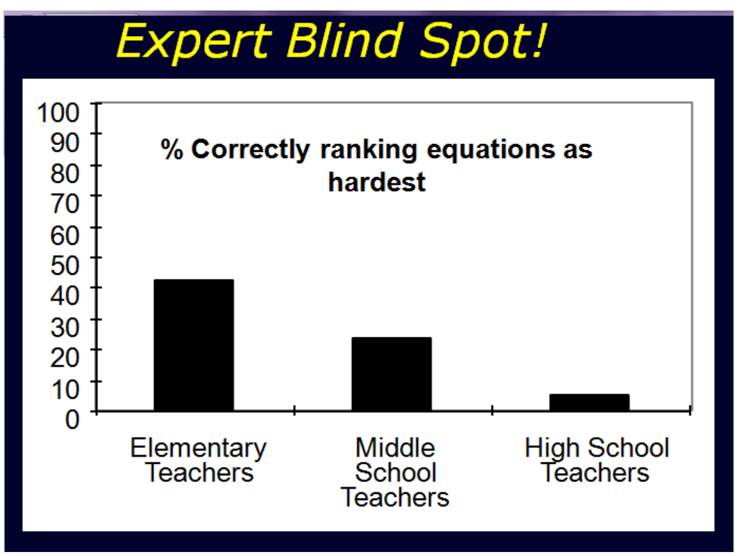


Figure 1. The more knowledge teachers had on a topic, the more likely they were to rank questions incorrectly as hard.

Experts, teachers, test designers, textbook writers, and educators all have these sorts of blind spots regarding what is hard for students to learn and in what pieces. Expectations about how students learn are part of our culture. For example, without hesitation, most teachers (and probably any given layperson) would say that story problems are harder to learn to do than numerical equations.

So why is this at all important? The way students *actually* learn should be reflected in the way we teach to the standards. We may have an idea of what the standards are that we want students to know, but it is an entirely different matter to understand what the knowledge components are that go into learning that standard (cf., Koedinger et al., in press). While working in ASSISTments, knowledge is tracked at a deeper and more detailed level, far deeper than specified in standards. ASSISTments tracks 147 knowledge components in mathematics from 4th to 12th grade, and these are mapped to the Common Core State Standards. For example, Common Core State Standard 8.G.9 involves cones, cylinders, and spheres. In ASSISTments, this is broken down into the six knowledge components related to the surface area and volumes of the three different shapes, as each can be assessed independently.

New Sources of Value in Technology Enhanced Assessments

The next three sections share the value of TEAs and how TEAs are able to provide more Assessment Value, Learning Value, and Research Value.

TEAs offer a profound ability to blend instruction and assessment at the same time. ASSISTments is one such web-based TEA that allows simultaneous assessment and learning opportunities. One of the “secret sauces” of blending instruction and assessment in ASSISTments is that, by providing instruction, we get more assessment information than we would without an instructional component. If a student gets a question right on the first try, that provides live assessment information, no different than a standard assessment in content (but much different in speed of availability). However, if a student gets an item wrong, we can use the amount of assistance the student needs as an extra indicator. Furthermore, that assistance helps the student learn. This works by giving students the tutoring they need to solve the problem. For example, some students need to see a worked example to get back on track, while others need to entirely relearn the material. How students perform on a question provides the basis for assessment report information for teachers. This reporting feature is discussed in a later section.

A web-based platform for assessment and instruction like ASSISTments isn’t unique; there are other platforms like it. This paper uses mostly examples from ASSISTments, but many other platforms do some of these things as well. , the ideas discussed are transferable. This makes what is being said about this one TEA relevant and applicable today. We will also discuss what makes ASSISTments special. For instance, some of the features that are particularly appealing to teachers were all developed with the aid of more than $13 million in funding from federal, state, and private grants.

Assessment Value of TEAs

There are two different value propositions for better assessment. First, we discuss the effective value of being able to assess students *during* instructional activities within a TEA and how assessment can be enhanced using the additional data available from these instructional activities (e.g., how much assistance a student needs). Second, we discuss how TEAs can expand what is assessed. The use of simulations and games is facilitating assessment in more complex and realistic contexts.

Tracking student learning within an assessment. The National Education Technology Plan (U.S. Department of Education, 2010) cited our work in predicting student knowledge by looking at students' use of tutoring as well as whether they were right or wrong on an item. We found that we could do a better job of assessing student knowledge by paying attention to how much assistance they needed. In this study (Feng, Heffernan, & Koedinger, 2009), data were collected from 2005-2007 to compare two different methods of performing and scoring online assessments. The first method was to give students practice Massachusetts Comprehensive Assessment System (MCAS) questions from released items and use their rate of success to estimate student proficiency (using Item Response Theory) and predict students' future MCAS scores with a simple regression. The second, more sophisticated method was to add to this simple model other variables from ASSISTments interaction, including the number of hints a student requested, the number of attempts needed to reach a correct answer, how many total problems a student worked on, and how many seconds it took students to complete a problem. These extra features turned out to be very predictive. This research demonstrated large statistically reliable and meaningful increases in accuracy in predicting MCAS scores from using these learning interaction variables above and beyond that predicted by proficiency estimates alone. In other words, we are capturing aspects of student achievement that go beyond cognitive measures of success on similar items (think practice tests), and these may include differences in student learning strategies, in work practices, or in effort. Figure 2 shows the scatter plots where we predicted MCAS scores (shown on the vertical axis) based on scores coming from a regression model using these enhanced features (shown on the horizontal axis). The predictions were quite accurate with correlations of $R > 0.8$.

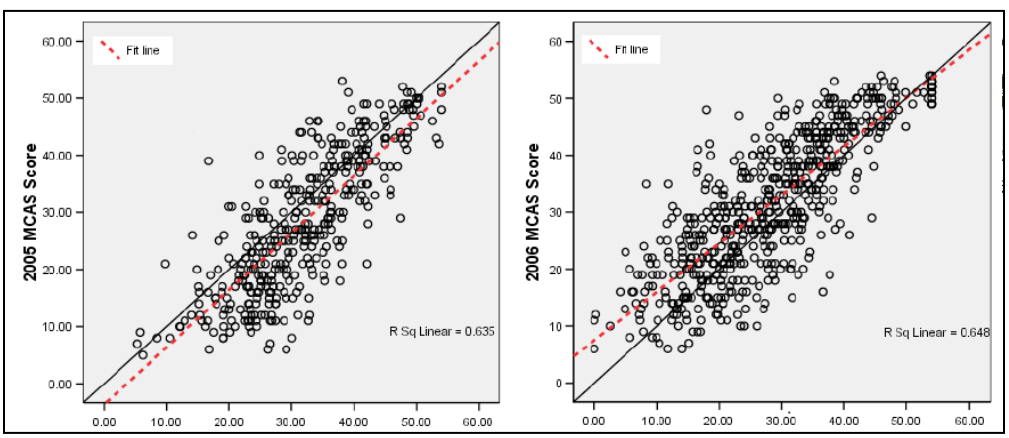


Figure 2. Scatter plots of MCAS score predictions.
MCAS = Massachusetts Comprehensive Assessment System.

Richer types of questions can be asked. One interesting example of a TEA is simulations and games. One type of simulation that works well inside a TEA is *microworlds*. Microworlds give students more authentic tasks that require them to manipulate the world to explore a topic.

A reenactment of Galileo's experiments with pendulums is a classic example of a microworld (see Figure 3). Students are presented with a pendulum, where they are able to measure periodicity and can manipulate the weight, how far back you pull the weight, and the length of the rod holding the weight. Students can manipulate the three variables to determine which of them will affect the period of the pendulum.

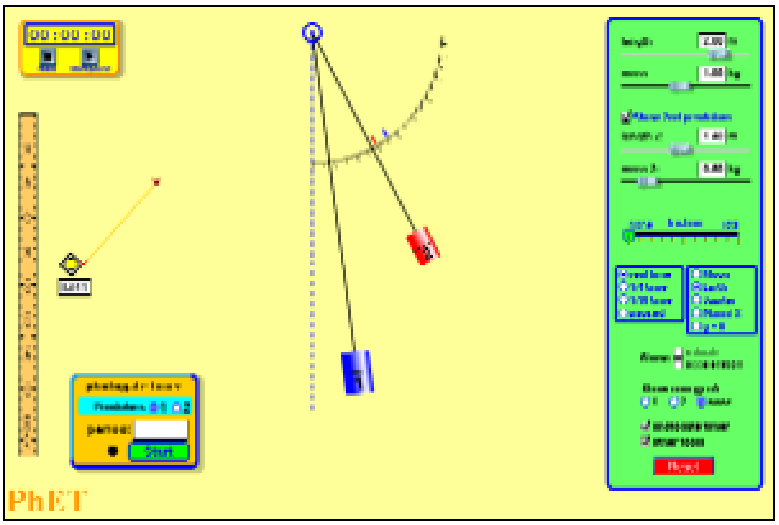


Figure 3. A pendulum microworld where students can manipulate the length of the string, weight, and period of the pendulum. Screen capture from PhET Interactive Simulations, University of Colorado. Retrieved from http://phet.colorado.edu/en/simulation/pendulum-lab

There are two uses of microworlds in assessment. The easy way to use them is to have students use the microworld as is and then follow that activity with traditional questions such as, "State what variables affect the period of the pendulum?" Students should say the length of string affects the period of the pendulum, but the mass does not.

A more detailed assessment approach is to monitor the actual actions students take in the microworld so the computer could say, "You said the length matters, but you never ran an unconfounded experiment that properly allowed you to conclude that."

ASSISTments is an ideal platform to either run microworlds in or to link out to microworlds. Student actions can be tracked and assessed for learning both the how-to of running a good experiment, as well as their understanding of the results. Figure 4 shows an image of a microworld inside ASSISTments (Sao Pedro, Gobert, Heffernan, & Beck, 2009). The data collected while students explore the microworld can be reported live to the teacher to be used to inform instruction.

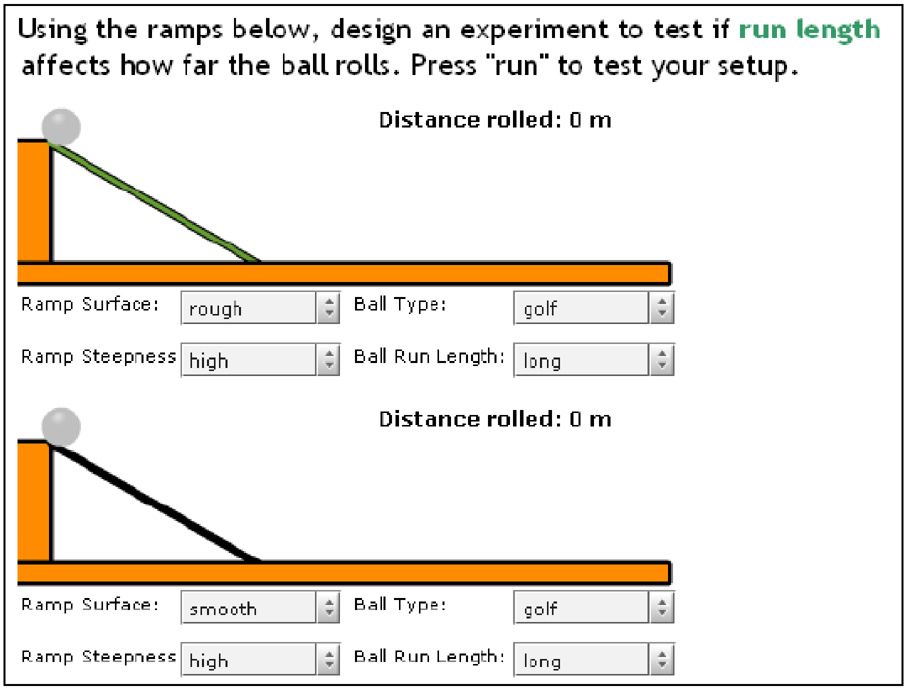


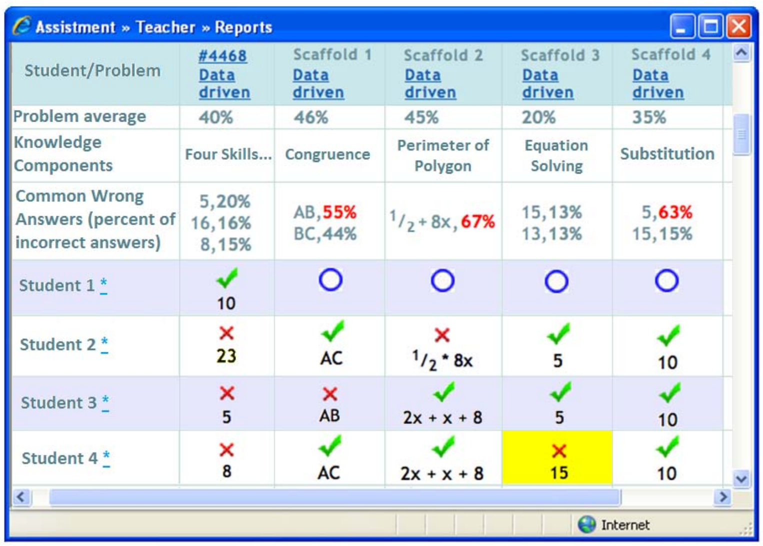
Figure 4. Typical ASSISTment question using the ramp environment.
The initial setup shown here is uncontrasted and confounded because the target variable, run length, is the same for each ramp, and ramp surface is not controlled.

While science is an easy-to-imagine use of microworlds, they are also used more broadly. Other research-based microworld TEAs are ScienceASSISTments, WISE, and PhET (phet.colorado.edu), and commercial versions include Gizmos from ExploreLearning and GeoGebra.org.

Learning Value of TEAs

Koedinger, McLaughlin, and Heffernan (2010) used a quasi-experimental design to evaluate whether student and teacher use of ASSISTments enhances student learning. The 7th grade math classrooms at three of four public middle schools, which had functioning computer labs, were assigned to ASSISTments use. The fourth school had a delay in the installation of its computer lab, so that school was unable to use ASSISTments and instead served as a quasi-experimental comparison group.

The data were collected from 1,240 7th graders in three treatment schools and the one comparison school. Posttest (7th grade year-end test) results indicated, after adjusting for the pretest (6th grade year-end test), that students in the treatment schools significantly outperformed students in the comparison school, and the difference was especially present for special education students. A usage analysis revealed that greater student use of ASSISTments is associated with greater learning, consistent with the hypothesis that it is useful as a tutoring system. We also found evidence consistent with the hypothesis that teachers adapt their whole class instruction based on overall student performance in ASSISTments. Namely, increased teacher use (i.e., having more students use the system more often) was associated with greater learning, even for students who made little or no direct use of ASSISTments. This result was indirect evidence that those students benefited from teachers adapting their whole-class instruction based on what they learned from ASSISTments use reports (see Figure 5). These results indicate potential for using technology to provide students instruction during assessment and to give teachers fast and continuous feedback on student progress.



Assistment » Teacher » Reports

Student/Problem	#4468 Data driven	Scaffold 1 Data driven	Scaffold 2 Data driven	Scaffold 3 Data driven	Scaffold 4 Data driven
Problem average	40%	46%	45%	20%	35%
Knowledge Components	Four Skills...	Congruence	Perimeter of Polygon	Equation Solving	Substitution
Common Wrong Answers (percent of incorrect answers)	5,20% 16,16% 8,15%	AB,55% BC,44%	1/2 + 8x, 67%	15,13% 13,13%	5,63% 15,15%
Student 1 *	✓ 10	○	○	○	○
Student 2 *	✗ 23	✓ AC	✗ 1/2 * 8x	✓ 5	✓ 10
Student 3 *	✗ 5	✗ AB	✓ 2x + x + 8	✓ 5	✓ 10
Student 4 *	✗ 8	✓ AC	✓ 2x + x + 8	✗ 15	✓ 10

Figure 5. This item report shows results for four students. The skills for each question, the common wrong answers, and each student's first submission are shown. Green check = student answered correctly on the first try; yellow highlight = student asked for every hint possible; red x = student answered the question incorrectly.

This study was undertaken over a full school year, so the duration helps to strengthen the results. It also compared the ASSISTments results to end of year, high-stakes tests, providing a measure of validity and pertinence.

These results also showed that there was a positive effect for special education students. These students appeared to benefit especially from ASSISTments. This makes sense—students who get patient, individualized help from the computer were more able to learn from that help.

We have also had the opportunity to evaluate the effectiveness of ASSISTments in true experiments with random assignment. The following studies focus on particular features.

Homework and immediate feedback. We compared, at the student level, two conditions: traditional homework where feedback was provided the next day and immediate feedback with students completing their homework via ASSISTments. The study involved 8th grade students who were randomly assigned to one of the two conditions, and the results demonstrated that students gained significantly more (effect size 0.40) with computer supported homework (Singh et al., 2011; cf., Mendicino, Razzaq, & Heffernan, 2009). This result has practical significance because it suggests an effective improvement over widely used paper-and-pencil homework. Other TEAs, not just ASSISTments, have the capability to provide that immediate feedback to students; these results suggest they should be used more broadly instead of current practices where students cannot get immediate feedback on their homework.

Mastery learning. In addition to immediate feedback, another benefit of TEAs is they can monitor student mastery of core standards. Mastery learning (e.g., Bloom, 1984) is a classic idea that a student needs just the right amount of practice to master a topic, and different students will need different amounts of time to master a topic. Heffernan and colleagues (2012) recently completed a study investigating the role of prerequisite mastery within the context of the Connected Mathematics curriculum. Connected Mathematics is one of the more popular curricula in the United States and has a strong following. Half the students in this study were assigned to a condition where they got a set of skill building problem sets that were prerequisites for the upcoming topic, and the other half were given problem sets that were unrelated. A mastery learning skill building set gives students practice on a topic until they are able to answer three correct questions in a row. We pretested and posttested students on the new topic of study. When we analyzed the work, students who had practiced the prerequisite skills learned more, with a 0.3 effect size. This means that students learned more when they had the prerequisite skills mastered. Remember, this work was done in one of the more well respected curricula, meaning that even when a curricula is well organized, many students fail to initially master or retain the knowledge—knowledge that is essential to learning a new topic.

Parent notification. Getting students to complete their homework leading to more learning seems like a simple idea, but teachers will tell you it is not that simple to accomplish. It was hypothesized that adding a parental notification component to ASSISTments would increase parental engagement in student learning and possibly student performance by allowing parents frequent access to the wealth of learning data collected by the software. An exploratory study was conducted at a local middle school, with promising results. In a follow-up to that study, automated e-mail messages were added to increase the strength of the intervention in response to our findings. We ran a randomized controlled experiment to test our hypothesis formally. Our results strongly indicate that parents felt more involved in their students' education after they began receiving these e-mails and that this involvement led to increased homework completion rates. No effect on student test score performance was detected; however, parent feedback regarding the intervention was positive (Broderick, DeNolf, Dufault, Heffernan, & Heffernan, 2011). When given good information, parents can help to ensure their children complete their homework. And an important benefit of ASSISTments and other TEAs is they make it easy to share this information with parents.

Research Value of TEAs

Should an administrator or teacher care if a TEA is useful for scientists to use for research? They should care immensely. A TEA that has design elements to help researchers set up and implement randomized experiments will help researchers discover the very best ways students, and various populations of students, learn. These research-based TEAs will pack the biggest punch—they will be able to be iteratively tested and improved with the most cutting-edge information.

ASSISTments is a test bed for learning. As a labile TEA, ASSISTments can bring the most recent and important cognitive science and computer science research to bear on every question students see, reports teachers see, and tests administrators give. Currently there are 147 experiments being run on ASSISTments. If the number 147 sounds familiar, it should—that is the number of different skills in the cognitive map. Every one of those skills is being tested to determine the best way to present that material or the order of presentation or one of a dozen other experiments currently running. Many are from within the ASSISTments team, but a growing number are developed and run from other researchers and teachers.

One of those experiments that holds great promise is "EdRank"—a way of ranking the educational content and efficacy of a web page. The goal is to customize learning for a specific student by using either the student's prior knowledge or other characteristics to select specific web pages to help students learn exactly the skills needed to achieve.

Games can be a viable method for students to learn as well as be assessed on their understanding of a topic. An example of this is the wildly successful, award-winning game Battleship Numberline (e.g., Lomas, Stamper, Muller, Patel, & Koedinger, in press). This game teaches numerical estimation, an important component of number sense. Students are asked to estimate the position of submarines on a number line, and a bomb is dropped in correct position. If their estimate was correct, the ship blows up, and they are awarded stars and points. Lessons are available for teachers to learn how to integrate this game into their teaching. Battleship Numberline is not just a fun game; it is used to perform experiments with hundreds of feature variations to determine which of these features most improve student learning and engagement.

Once again, this isn't a novel approach. Internet companies like Google, Amazon, and Zanga have been running randomized controlled experiments to optimize their products and sales. We as a society need to optimize the learning quality of our resources. What we are trying to optimize is not product use, but the learning value of different technologies, instructional strategies, or resources.

Implications for Policy-Makers

It is important as a nation to understand what works in education. There are myriad products on the market. How can we learn what works? The easiest way is to take policy stands that force data collection. It is important for policy-makers to choose not only something that teachers will use, but something that will iteratively improve through thoughtful, well-designed scientific studies and teacher feedback. This position was well expressed by James Pellegrino[1]:

> The combination of WPI [Worcester Polytechnic Institute]'s ASSISTments and the collaborative relationships with numerous schools and teachers provides an incredible test bed for designing and evaluating...instructional materials. There is no place else in the country where such a capability exists. (C. Heffernan, 2011)"

[1] It is worth noting other efforts to provide social and technical support for researchers and educators to work together in designing and rigorously evaluating innovative instruction, including the Pittsburgh Science of Learning Center's LearnLab (learnlab.org) and the Strategic Education Research Partnership (serpinstitute.org).

As the Michael & Susan Dell Foundation funding of the www.ed-fi.org project, principals and teachers need nonproprietary dashboards where they can, on a single screen, see data about their students, where different columns of data are coming from different vendors argues (Dell Foundation, 2011). No single vendor is incentivized to create dashboards that allow other companies' products to show up. Luckily, the Bill & Melinda Gates Foundation and others are trying, through the creation of the Shared Learning Collaborative (Gates Foundation, 2011), to take steps that support nonproprietary standards that will allow for data collections across proprietary systems. The Shared Learning Collaborative will help schools and states better understand what practices work best to enhance student learning. Nonpartisan clearinghouses that can look at these data can help policy-makers make better informed decisions. The drug industry requires products going through the approval process to collect and report such data, but after the drug is approved, they stop collecting the data. This is what developers and educators do as well, but our nation's students deserve better scientific study of the instruction they receive, whether that is traditional instruction or proposed innovation.

The Future

What is going to happen going forward? Schools, teachers, parents, and students are going to increasingly make use of technology for learning. As computers and handheld devices become more ubiquitous, teachers and school leaders will bring them into the classroom.

What will the benefits be? The adoption of this technology has the potential to lead to better learning, better formative assessment, and better integration of student assessment information from multiple sources, at both the micro level, from daily technology use, and at a macro level, from annual state testing. There is a potential for deeper learning, more individualized instruction, and better assessment of student progress and how to enhance it.

What will be challenges? Teachers and administrators will need training in how best to incorporate such systems into their curricula and, most importantly, how to best make use of the assessment data to guide policies in general and instructional practices in particular.

TEAs are important, but the decision to adopt them must go beyond whether they "look good" to teachers and administrators. If they are not adopted, they will not be successful. We have found that more than 80% of the teachers using ASSISTments found the website on their own and have not been given it from their principal or district leadership. Many teachers are eager for this type of technology.

A look into some ASSISTments classrooms provides a glimpse into what the future of technology could look like. Consider Bellingham, Massachusetts, math teacher, Barbara Delaney, who found out about ASSISTments at a math conference. She started using ASSISTments three years ago and quickly began to see the value of the endeavor, beginning with using nightly homework. Delaney had started using ASSISTments by herself with the assistance of staff at Worcester Polytechnic Institute in the form of workshops and class visits. In her second year, she recruited a second teacher and started using other features, such as parental notification and the automated reassessment and relearning system (ARRS). In her third year, a new principal came to her school from Grafton, a school already using ASSISTments. He understood the value of the program and granted Delaney the time to train all the math teachers in ASSISTments at a professional development day. Now, in the first year of this school-wide use, the teachers are using skill builders and nightly homework features and will be using more advanced features, such as ARRS and parent notification. The principal has plans to extend ASSISTments usage to science, English, and health classes and is planning to make more time available for teachers to work together and for Delaney to train them.

Delaney says that, other than her iPhone, there has been no other technology that has more profoundly affected her life as much as ASSISTments. She represents the champion that every school needs to get started on a new technology, because we will not be able to visit every school and start training the whole staff. We find that when we do a workshop at a school where there are no current users, often these schools fail to continue to use it or it just falls flat. We have found the most effective way to train schools to use these TEA environments is to get a champion and help that school's principal to support that champion to train the rest of the school.

From Delaney (personal communication, April 5, 2012):

> The success of ASSISTments in my school is because we have an administration that supports our work. The sharing of this assessment tool has provided us with a natural platform for developing a professional learning community. The other sixth grade teacher and I are in a continual daily discussion about our students' learning progression. We share common homework, common assessments, and common benchmarks. Having this data helps to launch discussion about our teaching practices and learning goals. We work on everything together with data to direct us. I wonder every day how I taught before I used ASSISTments. Our administration respects this important work and gives us professional development time to also have professional discussions on planning and assessments.

Summary and Classroom Changes

In summary, there are a lot of uses of TEAs. Below are some observations by Delaney and another champion teacher on how using a TEA has changed their teaching and their classroom.

Teachers report many changes in the classroom after using ASSISTments. One teacher, Kim Thienpont, a 7th grade math teacher from Grafton Middle School, describes the changes in her teaching like this (K. Thienpont, personal communication, August 30, 2010):

> In my 10 years of teaching, I can say that nothing has changed my delivery of instruction like ASSISTments. It promotes formative assessment at the daily level while providing differentiated instruction in a way that is not possible without technology.
>
> Before using ASSISTments, I would briefly check homework and give students an opportunity to ask questions. I really had no way to know how students were doing on their homework or where they were struggling. I noticed that without accountability, student effort was lacking. I began to score every assignment. While this provided accountability, the feedback was not timely enough to influence my instruction or to affect student learning. With the introduction of nightly homework with ASSISTments, students get feedback as they do their homework. This allows students to correct misconceptions at home and come in prepared for new lessons. I also have access to reports showing class performance on every assignment. I can decide, in the morning, based on actual data, if as a class we are ready to move forward and/or which specific question I need to review with my students. Common misconceptions are clear from the report and allow me to adjust my lessons to address the needs of my students. I honestly didn't even

know enough about formative assessment until I was encouraged to learn more because of ASSISTments. I now have a very different understanding of the purpose of homework, and I am able to adjust my instruction based on student performance.

Delaney has been amazed at the changes in her students:

ASSISTments has brought many changes to my classroom. First, we no longer begin class with correcting homework. My students come to class already knowing what they did well and with questions because of the answer feedback built into the ASSISTments system. (B. Delaney, personal communication, April 5, 2012)

The students have had time to reflect on their work, and Delaney has seen students move beyond seeing assessment as a grade and to seeing it as feedback. They use the feedback and seek to learn the material for the sake of learning, no longer just for the grade. “My students have really begun to take ownership of their own learning. It is an amazing transformation, and it couldn’t have happened without ASSISTments.”

Acknowledgments

We would like to thank Michele Decoteau for editorial assistance and the 23 funders (listed here http://www.webcitation.org/67MTL3EIs) of ASSISTments. We also thank the National Science Foundation (#SBE-0354420) for funding of the Pittsburgh Science of Learning Center.

References

Black, P., & Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. London: School of Education, King's College.

Bloom, B. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher, 13(6),* 4–16.

Broderick, Z., DeNolf, K., Dufault, J., Heffernan, N., & Heffernan, C. (2011). Increasing parent engagement in student learning using an intelligent tutoring system with automated messages. *Journal of Interactive Learning Research*, *20*(4), 467–490.

Cizek, G. J. (2000). Pockets of resistance in the assessment revolution. *Educational Measurement: Issues and Practice*, *19*(2), 16–33.

Creighton, T. B. (2001, April). Data analysis in administrators' hands: An oxymoron? *The School Administrator, 58,* 6–11.

Dembosky, J. W., Pane, J. F., Barney, H., & Christina, R. (2006). *Data Driven Decision making in Southwestern Pennsylvania School Districts*. (RAND Corporation Report No. WR-326-HE/GF). Retrieved from http://www.rand.org/pubs/working_papers/WR326

Earl, L. and Katz, S. (2006). *Leading schools in a data-rich world: Harnessing data for school improvement*. Thousand Oaks: CA, Corwin Press.

Feng, M., Heffernan, N. T., & Koedinger, K. R. (2009). Addressing the assessment challenge in an intelligent tutoring system that tutors as it assesses. *The Journal of User Modeling and User-Adapted Interaction*, *19*, 243–266.

Gates Foundation, Bill & Melinda. (2011, June). *Shared Learning Collaborative, LLC*. Retrieved from http://www.gatesfoundation.org/Grants-2011/Pages/Shared-Learning-Collaborative-LLC-OPP1041367.aspx

Heffernan, C. (2011). TestBed. Retrieved from http://teacherwiki.assistment.org/wiki/TestBed.

Heffernan, N., Heffernan, C., Dietz, K., Soffer, D., Pellegrino, J. W., Goldman, S. R., & Dailey, M. (2012, April). *Improving mathematical learning outcomes through automatic reassessment and relearning*. Presented at the annual meeting of American Educational Research Association, Vancouver, British Columbia.

Heffernan, N. T., & Koedinger, K. R. (1998)*. A developmental model for algebra symbolization: The results of a difficulty factors assessment.* In M. Gernsbacher & S. Derry (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 484–489). Hillsdale, NJ: Erlbaum.

Holcomb, E. L. (2004). *Getting excited about data: Combining people, passion, and proof to maximize student achievement* (2nd ed.). Thousand Oaks, CA: Corwin.

Koedinger, K. R., Corbett, A. C., & Perfetti, C. (in press). The Knowledge-Learning-Instruction (KLI) framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science*.

Koedinger, K. R., & McLaughlin, E. A. (2010). Seeing language learning inside the math: Cognitive analysis yields transfer. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 471–476). Austin, TX: Cognitive Science Society.

Koedinger, K. R., McLaughlin, E., & Heffernan, N. (2010). A quasi-experimental evaluation of an on-line formative assessment and tutoring system. *Journal of Educational Computing Research, 43(1)*, 489–510.

Lomas, D., Stamper, J., Muller, R., Patel, K., & Koedinger, K. R. (in press). The effects of adaptive sequencing algorithms on player engagement within an online game. In S. A. Cerri, W. J. Clancey, G. Papadourakis, & K. Panourgia (Eds.), *Proceedings of the International Conference on Intelligent Tutoring Systems, Lecture Notes in Computer Science, 7315*. Berlin, Germany: Springer.

Mendicino, M., Razzaq, L., & Heffernan, N. T. (2009). Comparison of traditional homework with computer supported homework. *Journal of Research on Technology in Education, 41*(3), 331–359.

Militello, M., & Schweid, J. (2008, October-November). *How formative assessment products can be used to improve teaching quality and student outcomes.* Paper presented at the 2008 Annual Convention of the University Council for Educational Administration, Orlando, FL.

Nathan, M. J., & Koedinger, K. R. (2000). Teachers' and researchers' beliefs of early algebra development. *Journal for Research in Mathematics Education*, *31*(2), 168–190.

Popham, W. J. (2008). *Transformative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.

Sao Pedro, M., Gobert, J., Heffernan, N., & Beck, J. (2009). Comparing pedagogical approaches for teaching the control of variables strategy. In N. A. Taathen & H. van Rjin (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 1294-1299). Austin, TX: Cognitive Science Society.

Singh, R., Saleem, M., Pradhan, P., Heffernan, C., Heffernan, N., Razzaq, L., & Dailey, M. (2011). Improving K-12 homework with computers. *Proceedings of the Artificial Intelligence in Education Conference*. (pp. 328-336). Berlin, Germany: Springer.

Susan and Michael Dell Foundation, (2011). *About Ed-Fi*. Retrieved from http://www.ed-fi.org/about/

United States Department of Education. (2003). *Using data to influence classroom decisions*. Washington, DC: US Government Printing Office.

United States Department of Education. (2010). *National Education Technology Plan*. Retrieved from http://www.ed.gov/technology/netp-2010.

K-12 Center at ETS

Invitational Research Symposium on

Technology Enhanced Assessments

The Center for K–12 Assessment & Performance Management at ETS creates timely events where conversations regarding new assessment challenges can take place, and publishes and disseminates the best thinking and research on the range of measurement issues facing national, state and local decision makers.