Using Technology and Evidence to Promote Cultures of Educational Innovation:

The Example of Science and Mathematics Education

April 2015

SRI Education™
A DIVISION OF SRI INTERNATIONAL

Authors

Barbara Means, Linda Shear, Jeremy Roschelle
Center for Technology in Learning, SRI Education

SRI Education™

Acknowledgements

SRI Educations thanks Stéphan Vincent-Laurent (OECD) for feedback on earlier drafts of this report.

Suggested Citation

Means, B., Shear, L., Roschelle, J. (2015). *Using Technology and Evidence to Promote Cultures of Educational Innovation: The Example of Science and Mathematics Education.* Menlo Park, CA: SRI International.

© Copyright 2015 SRI International.

Contents

Introduction

To increase the opportunities for students to learn the knowledge and skills that are valued in the 21st century, education must become more innovative. As numerous publications attest, societies value students who are skilled in solving open-ended problems, collaborating with others, and participating in innovation. Increasingly, students must be able to persevere in tackling complex, non-routine and long-term challenges. Societies also feel a pressing need for students to exercise these skills in combination with a deeper understanding of the STEM disciplines of science, technology, engineering, and mathematics. Our existing educational system, which was designed largely to address the learning needs of the industrial age, is not well-suited to these emerging needs. New approaches to education for the knowledge and innovation ages must be invented.

As we contemplate what these new approaches should be, technology has obvious appeal as an agent of transformation. People find different words to name this appeal. Some assert that technology can make education more "personalized" or deliver "anytime, anywhere" accessibility or make learning more "connected." Although the rhetoric varies, the essential appeal is consistent: a capability as flexible and powerful as today's information and communications technology (ICT) must be an important tool for inventing the future of learning (U.S. Department of Education, 2010).

We agree. And yet, as researchers who specialize in the evaluation of learning technology, we are strongly aware of the contravening impulse, an impulse to ask "where's the evidence that technology can improve learning?" Often the evidence is too hard to find, too thin, too weak, or too confusing to interpret. Broad concepts such as "personalization" or "anytime, anywhere" prove difficult to operationalize in terms of learning science theory and research methods, further compounding the challenge of gathering evidence of ICT's effectiveness. Moreover, technology enthusiasts are inclined to view the process of getting more evidence about the effectiveness of technology as too expensive, too slow, and too likely to yield the wrong kind of information.

These difficulties can lead to incongruity between technology investments, learning sciences research, and evidence-based decision making. The disconnections result in large pendulum swings. For example, as school leaders see the appeal of learning technologies, large educational technology investments are sometimes made, usually focusing on hardware and not always capitalizing on what we know about learning. Later, when evidence of improved student outcomes fails to materialize, those investments are characterized as failures. Not enough is learned in the oscillation between high expectations and disappointment as the pendulum swings between ICT optimism and results-oriented reviews of evidence of ICT impact.

Consequently, we argue that for new cultures of educational improvement and innovation to thrive, a *rapprochement* between technology enthusiasm and evidence is needed. Learning sciences research provides crucial theory and methods on which to base that resolution. Technology is one of innovation's most flexible raw materials. In turn, innovation requires theory to guide the process of shaping that material

© Copyright 2015 SRI International.

to new ends, and methods to gather evidence for iterative improvement. The old oscillations between high expectations and disappointment should be replaced by a disciplined, iterative evolution of connections among technology, learning theory, and research to yield progressively better evidence of progressively more important applications of technology.

To accomplish this, we argue for a rethinking of the relationships between learning technology and evidentiary arguments: a new evidence framework is needed. The framework should both fit the means of educational innovation and address its ends. For example, as educational innovation becomes more oriented to rapid iteration and rapid scale up, evidence gathering must occur more quickly and be capable of mining the patterns to be found in large data sets. Further, as educational innovation becomes oriented to 21st-century knowledge and skills, evidence must be collected on measures that matter today—a requirement that is challenging because many 21st-century concepts and skills are not well-captured by most existing assessments.

The path to this new evidence framework, we suggest, begins by thinking about technology and evidence from the perspective of different roles in the education enterprise. The forms of evidence that are most useful depend on a person's role in educational improvement and innovation. In this paper we consider four roles: researcher, developer, policymaker, and educator. Each faces basic evidentiary dilemmas: there is never enough perfect evidence available, and yet decisions must be made and actions taken. Each also has new opportunities for working with data-based evidence, many of which are made possible by ICT. Models for decision-making under conditions of uncertainty are needed, and these must vary with the role. If these models can bring the capabilities of technology and the discipline of analyzing and acting on evidence together as appropriate for each role, innovation can be accelerated.

Further, within the broader system, we can understand that many of the disappointments experienced with either evidence or technology come from trying to match technology as seen from the perspective of one role with evidence produced to meet the needs of another role. For example, the evidence that researchers collect will rarely answer most of the questions about technology that teachers have. By seeking to connect technology and evidence for each role, greater alignment can be achieved.

In this paper, we first describe the ways in which technology can provide the conditions that learning science has shown support meaningful learning. Next we discuss some broader considerations leading to the conceptualization of evidence frameworks. Then we examine each of the four roles in turn. For each role, we consider some of the basic opportunities available for working towards better support of 21st-century learning. We also consider one or two signature projects and how these projects leveraged an interplay of technology and evidence to address 21st-century learning goals. Throughout the paper, we exemplify our arguments with examples of projects or tools in science and mathematics education, but our argument applies in similar ways to other subject matters as well. We summarize some of the broad lessons learned and suggest the utility of sharing these broad lessons across roles. After discussing each role, our concluding section suggests broad trends in both use of technology and the use of evidence in education with the goal of further accelerating cultures of educational improvement and innovation.

© Copyright 2015 SRI International.

How Technology Can Improve Learning: Building on Research from the Learning Sciences

Imagine a broad policy effort to improve students' health through nutrition. Leading voices suggest that the next step in food is "personalized nutrition" and "anytime, anywhere healthy eating." Obviously, these are good ideas for today's students. Yet policy makers ask for evidence before proceeding with programs to put these ideas into practice, and evidence of the effects of personalized nutrition and anytime, anywhere healthy eating has been surprisingly hard to find. Should we abandon individualized nutrition programs due to the lack of evidence?

Obviously, no. Instead, we might do two things. First, we might want to become clear about the active ingredients that connect food to health through biological processes and make sure that the food in our nutrition program has these ingredients. Clearly, if there is no active connection between the nutrients in the foods labeled healthy and students' biological processes, the "personalized nutrition" program will not improve student health. Second, we might realize that health emerges from a broad complex of factors, including not only food but also exercise, sleep, and environmental factors. Thus, to improve health for students, we might want to test integrated interventions that combine these factors. Giving schools plenty of personalized, mobile nutritious snacks is more likely to increase student health in the context of such a comprehensive intervention. Thus, true innovations in student health would likely both (a) build on the science of nutrition as it relates to biological processes and (b) embed the nutritious foods in complete systems that address all the factors relating to healthy living.

Based on our long experience in evaluating the role of technology in improving learning, we argue for a parallel approach. First, we would like to be sure that a proposed educational application of technology is based on good science about learning processes – if there is not a good connection between learning processes and the proposed use of technology, it likely does not matter how "personalized" or "anytime, anywhere" the technology is. Second, learning emerges from a broad complex of factors, including not only technology but also curriculum, pedagogy, teacher professional development, assessment, and leadership. To improve students' academic learning, we prefer to test integrated interventions that align these factors, rather than trying to isolate the impact of a new technology alone. True innovations in learning, we argue, use technologies in ways that exemplify learning science principles and embed technologies in broader, integrated interventions.

Fortunately, recent decades of research in the learning sciences (Bransford, Brown, and Cocking, 2000; Sawyer, 2006; OECD, 2010b) have begun to illuminate how students learn difficult and important concepts and practices, and the ways in which technology can strengthen these learning processes. In some cases, the research presently is at the "active ingredient" level – establishing the connection between a new affordance of technology and successful learning processes. In other cases, both the active ingredient level and the comprehensive intervention level have been evaluated, and results on impacts are in.

© Copyright 2015 SRI International.

Technology Affordances and Learning Sciences Principles

The capabilities of ICT can be used to provide the conditions that learning sciences research shows promote deeper and longer-lasting learning. Research tells us that learning is enhanced when students do the following:

Engage their prior understandings and become active drivers of their own learning. Traditional behaviorist models of teaching and learning treat students as "empty vessels" to be filled with new knowledge by teachers and textbooks. In contrast, research demonstrates that student learning is strongly shaped by the preconceptions and experiences that students bring with them into the classroom (Bransford et al., 2000; Clement, 1982; DiSessa, 1983). Often these preconceptions are based on students' experiences in the everyday world, and include both ideas that are part of mature conceptions and ideas that are false. For example, children understand animals as living things, as distinguished from objects such as rocks, at a very early age. But at the same time, children often believe that self-initiated movement is a requisite for something to be alive, thus failing to recognize plants as part of the category of living things (Inagaki and Hatano, 2002). If these intuitive but incorrect ideas are allowed to persist, they can prevent students from understanding and retaining scientific explanations.

Good teaching therefore creates situations in which students have experiences that reveal the limitations of their intuitive ideas and supports students in developing more accurate ideas. Strong learning environments bring students' existing ideas out into the open and support students in building on and reorganizing their existing ideas into increasingly productive mental models (Minstrell, 1989; Smith, diSessa, and Roschelle, 1993). Responding to different mental models, these learning systems adjust learning paths to the needs and background of individual students (Corbett, Koedinger, and Hadley, 2001) and engage students as active agents in their own learning (Prince, 2004).

Technology is particularly adept at providing the range of representations, means of engagement, and opportunities for expression that are essential to universal designs for learning (CAST, 2011), enabling designs that are more flexible and effective for all students. For example, well-designed museum exhibits can evoke students' prior knowledge and stimulate the wonderment that drives inquiry (Roschelle, 1995). Sometimes, this works by changing the scale to make a phenomenon more visible; at a San Francisco science museum called the Exploratorium, the classic "giant guitar string" allows students to play with a "guitar string" that is 100 times larger than normal, making it easier for students to explore wave phenomena hands-on. A long tradition in technology-enhanced learning likewise provides students with probes and sensors which can be attached to graphing calculators or computers for instant data collection and display (Zucker et al., 2008). Probes allow students to directly explore phenomena (such as the turbidity of a local stream) and also dramatically speed up the process of going from data collection to interpreting a graph, allowing students to more directly and rapidly connect minds-on and hands-on science. Today, students are often immersed in scientific phenomena through game-like simulations,

© Copyright 2015 SRI International.

as well. In River City (Clarke and Dede, 2009), for example, students conduct an extended scientific investigation of the spread of disease in a virtual world designed to have biological, historical, and geographical realism. Experiences like River City can provide students with more authentic opportunities to play the role of a scientist.

Develop richly connected knowledge, not just isolated skills. Success in today's knowledge economy is based not just on what people know—facts, after all, are increasingly easy to look up on the Internet—but on how they are able to put that knowledge to work (Bereiter, 2002; Lesh and Doerr, 2003) to analyze complex emerging problems and create and communicate innovative solutions. Learning sciences research suggests that this application of knowledge relies on conceptual, not just factual, understanding in the form of higher-level principles and recognized patterns that can be transferred to new situations (Pellegrino and Hilton, 2012). For example, although a student who memorizes the names and locations of major cities in Europe can successfully fill them in on a map, a student who understands why people tend to settle near large bodies of water may be better prepared to investigate the growth of population centers in another part of the world (Bransford and Schwartz, 1999). This level of conceptual understanding, enabling application to new situations, is a prerequisite for the ability to solve the types of complex and non-routine problems that are a common feature of the 21st century workplace.

The process of building this rich conceptual understanding, often called "deeper learning" or "learning with understanding," can be promoted by a range of classroom strategies that include:

- The use of multiple representations that help students to consider complex ideas in multiple ways and see the connections among them (Ainsworth, 1999; Kaput, 1992);
- Instruction that emphasizes learning "big ideas" in depth rather than surveying a series of disconnected facts (Bruner, 1960; Harlen, 2010; Hiebert and Grouws, 2007; Schmidt, Wang, and McKnight, 2005); and
- Project-based approaches that allow students to investigate ideas in meaningful real-world contexts and to actively construct their understanding (Barron et al., 1998; Blumenfeld et al., 1991; Krajcik, Czerniak, and Berger, 2002).

Technology has been used to support all of these strategies. The hyperlinked Internet provides access to a much broader array of multi-modal materials and resources, including access to human experts who can support the development, understanding, and communication of complex ideas. Access to the Internet allows students to research broadly and deeply, and it can support the construction of meaning when students use it to observe and interact with interconnections among ideas. Specific features of technology relate to simulation, visualization, modeling, and representation. A picture is often worth a thousand words, and a moving picture can be especially helpful to scientists and to students who are trying to make sense of complex relationships in a phenomenon or data set. For example, in SimCalc, an example that will be discussed in more detail later, students learn how mathematicians analyze "rate of change"

© Copyright 2015 SRI International.

by investigating the relationship between graphs and simulated motions. The technology in SimCalc links graphs, tables, algebraic expressions, and simulated motions so that students can modify one of these representations and see the implications in other representations. By working across the representations through a set of carefully sequenced lessons, students build richly connected knowledge of a foundational idea that leads to algebra and more advanced mathematics.

Leverage social interactions to build knowledge together. In traditional models of teaching, learning is an individual enterprise, and looking at the ideas of other students is often considered “cheating.” In contrast, research emphasizes the value of collaborative learning for building deep knowledge (Bransford et al., 2000; Johnson and Johnson, 2009). In a learning community, students work together toward a goal of advancing shared knowledge on a topic in a way that helps each student learn (Brown and Campione, 1996; Scardamalia and Bereiter, 1994). When students collaborate or debate, their interactions require them to articulate their own ideas and evaluate, question, sharpen, or build on the ideas of others, thus deepening their individual and collective conceptual understanding (Fischer et al., 2013). In the process, students learn essential skills of listening, negotiating, and coming to agreement that will enable them to be effective collaborators in the future.

The emergence of online social interactions can support organized and focused student collaborations and community building, either within the classroom or across geographic distances and even borders. We know that students learn better when they explain to peers and have peer support for their own questions. However, it can be hard to organize effective collaborations in typical classrooms or lecture halls, and simply “working together” is no guarantee of better learning.

Fortunately, a well-developed theory of how technology can support collaborative learning is emerging (Kirschner and Erkins, 2012). Engagement with online learning environments can be designed to support presentation of ideas with peer interaction and feedback, the ability to co-construct documents and presentations online, as well as the ability to create homework help, study groups and challenge-based learning teams. Technology can provide an enduring record of the group’s thinking, allowing it to be reviewed and refined over time. For example, the Knowledge Forum environment provides a structure for students to articulate, link, and reflect on their own and each other’s ideas, encouraging the development of increasingly sophisticated conceptual frameworks that students have created on scientific topics such as human body systems and causes of pollution (Scardamalia and Bereiter, 2006). Technology can also offer ways to scaffold roles and coordination among team members such that tasks require both individual and collective accountability—important requirements of strong collaborative activities. For example, the electrostatics game “First Colony,” developed by researchers in Chile, utilizes multiple mice to allow each student to control a different astronaut who must collaborate with peers to move space crystals to a target location using electrical forces that the various players control. Researchers have found significant conceptual learning gains in physics content after using this game for just a single class period (Echeverría et al., 2012).

© Copyright 2015 SRI International.

These online social environments for learning also augment teacher capacity. Although teachers can help small groups, they can only visit one group at a time. Innovators have defined a range of technologies that can support small-group collaboration, ranging from simple "clickers" to elaborate supports for the interactive roles of each participant. For example, Peer Instruction (Crouch and Mazur, 2001) is a pedagogy that has transformed physics lecture halls, using handheld devices to collate all students' responses to a well-chosen question so that class discussion can directly engage students' collective ideas. The method has been shown to produce substantial improvements in student learning compared to a control classroom, as well as stronger attendance and engagement (Deslauriers, Schelew, and Wieman, 2011). Implementing the pedagogy, however, requires coordinated changes in the use of space, the content of the course, the role of the students and instructor, and training for the instructor. Likewise, students learn more about fractions with a handheld collaboration technology when content is carefully defined, behavioral training is provided to students, teachers are trained, and a new pedagogy is implemented (Roschelle et al., 2010).

Monitor the learning process and receive ample, useful feedback. A goal throughout child development is independence. Likewise, it is important to develop students' capacity for effective independent learning. Students can improve their ability to "self-regulate" – to monitor their own comprehension, direct their own learning efforts, and control their own activities (Butler and Winne, 1995; OECD, 2010b). Students who self-regulate can take on bigger challenges, practice on their own, develop deeper understandings, and manage their own success.

Feedback is important to self-regulation; research has demonstrated that the strongest learning takes place when students are responsible for their own learning and when they have the benefit of feedback that is productive toward that learning (Butler and Winne, 1995; Hattie, 2009). Feedback provided through an integrated technology system can supplement the capacity of teachers to provide appropriate feedback to each student in a class, and to provide faster feedback loops that can inform next steps for both students' learning and teachers' instructional decisions (Roschelle, Penuel, and Abrahamson, 2004). In mathematics, for example, ASSISTments is a computer-based system that uses a bank of math problems to analyze the patterns in students' responses, providing timely feedback and appropriate scaffolding tailored to each student. In addition, the system provides analytics for the teacher that describe individual student progress and common conceptual challenges, supporting teacher decisions about coaching for individual students and about the class discussion topics that will be most productive for the most students (Heffernan et al., 2012).

To use this feedback effectively, both students and teachers need support: students need to learn to use the feedback to regulate their own learning process, and teachers must be able to use the feedback to adapt. In one example of an integrated program, investigators used wirelessly connected calculators to give students and teachers more feedback in an Algebra course, and also provided support for students to become more self-regulating and for teachers to adapt instruction. A randomized controlled trial with 68 teachers and 1,128 students found this to be an effective combination of technology-provided feedback

© Copyright 2015 SRI International.

and supports for student and teacher use of feedback, with significant results in both mathematics learning and student self-efficacy (Pape et al., 2008).

Formative assessment can provide important and timely information to feed students' learning and shape teachers' instruction (Black and Wiliam, 1998; Fuchs and Fuchs, 1986). Typical assessments focus on right or wrong answers, and a technology-based assessment can be designed to give immediate simple feedback (right or wrong). But beyond the basics, technology-based feedback can include providing worked examples, modeling how to solve a problem and guiding a student through the steps. Other types of assessments can inform and further the type of deep conceptual learning described above. Technology can support this process by asking students to reason about many different situations and using each student's responses to diagnose the set of ideas that the student holds (Minstrell, 1989). Such assessments look at how students' thinking within a given domain is organized and identify patterns of responses that may suggest misconceptions or next steps on the path to understanding (Means, 2006). In turn, this makes possible customized instruction, guiding the student's construction of knowledge as it helps make their progress visible. Real-time formative assessment enabled by technology, for example, with tablets and software such as InkSurvey, is one example of how technology can help teachers better understand and respond to the knowledge and level of learning of their students (Gardner et al., 2013; Kohl et al., 2011).

Much research and development has focused on technology-based adaptive learning environments as well as cognitive tutors that augment (not replace) the teacher's ability to interact with individual students by providing responses and structured feedback in response to student input.

© Copyright 2015 SRI International.

Applying Learning Sciences Principles Leads to Student Gains

Technology-Enhanced Learning in Science (TELS) is a program designed to support students' deep understanding of core principles in the science curriculum. TELS modules offer interactive visualizations of scientific phenomena that are often impossible to observe directly (such as chemical reactions). Students explore these phenomena through the lens of current scientific issues (such as treatment options for cancer). Through the software, students are guided to generate and test predictions, explain their understandings, and engage in discussions with peers.

For example, in a 5-day unit called *Chemical Reactions* (Chiu and Linn, 2012), students conduct a series of activities about chemical processes related to the greenhouse effect. In one activity they work with an interactive visualization of hydrocarbon combustion reactions, manipulating variables to look at the relationship between ratios of reactant molecules and products, and they consider implications for carbon dioxide levels in the air. Students are prompted to conduct investigations and explain their understanding of related scientific processes. They use what they learn to write a letter to their congressperson about climate change and alternative fuels.

This program illustrates many of the learning sciences principles described above. Interacting with visualizations, generating and improving hypotheses, articulating explanations, and applying what they've learned to new situations ***engages students' prior understandings*** and gives them the opportunity to ***develop richly connected knowledge.*** Interactions with peers as they build this knowledge ***leverages social interactions to build knowledge together.*** Scaffolds in the software help students to ***manage and monitor their learning process.*** When enacted together, these elements build on each other to support a powerful and integrated learning experience.

Studies have shown that TELS students significantly outperform their counterparts using traditional curricula in a measure of students' integrated understanding of scientific phenomena (Linn, Lee, Tinker, Husic, and Chiu, 2006). Other studies have demonstrated how aspects of the program support students in becoming successful collaborators who are able to monitor their own learning, among other 21st century skills (Linn and Eylon, 2011).

© Copyright 2015 SRI International.

Rethinking the Approach to Evidence in an Age of Technology

Overall, learning sciences theory and research provide very solid grounding for the assertion that technology-based approaches can provide important, and at times transformative, support for deeper student learning, problem-solving, collaboration, and other essential 21st century capacities. We turn now to the question of what kinds of evidence best meet the needs of innovators who are developing these approaches or seeking to implement them in practice.

The conventional education research trajectory moves through stages from initial small-scale design research (a "proof-of-concept study") to studies of the effects of the fully developed intervention tried out in a conducive context with extensive support for implementation (an "efficacy study") to tests of effects of the intervention as implemented in a large number of more typical settings (an "effectiveness study" or randomized controlled trial). As applied to learning technology, this trajectory suggests that a researcher starts the R&D process by taking insights about learning based on theory and laboratory studies and then encapsulates those principles in a piece of learning technology or a protocol for using ICT to support learning. Then the researcher tries the learning innovation out in a few classrooms to test its usability before proceeding to the stage of efficacy studies. Following successful efficacy studies, the researcher seeks funding to try implementation, and evaluation of effectiveness, at scale with many classrooms. It is not unusual for this process to take ten years or more.

It is difficult to reconcile this staged research model with Web-based technology products that are in a state of constant change (U.S. Department of Education, 2013). By the time some version of the technology has been through the whole research trajectory, it will no longer be the version available for use. Just how limited this traditional approach to acquiring evidence of effectiveness is from the standpoint of those who want to make evidence-based decisions about learning technology can be seen in the What Works Clearinghouse (WWC) operated by the Institute of Education Sciences within the U.S. Department of Education. The WWC was established in 2002 specifically for the purpose of identifying educational interventions for which there is rigorous evidence of effectiveness. Once an intervention has been chosen for WWC review, a systematic, well-documented process for locating studies and judging their quality is implemented (U.S. Department of Education, Institute of Education Sciences, 2012). We consulted the WWC to find learning technology interventions with evidence of effectiveness in August 2012 and found intervention reports on just 45 digital learning interventions, of which 26 were found to have positive or potentially positive effects on at least one outcome. After a decade of operating with major funding, the WWC had identified just 26 learning technology applications that met their standard for demonstration or promise of effectiveness. Moreover, of these 26, 7 were no longer available.

Within the high-tech world, the prevailing approach to gathering evidence is much different. As more developers and investors from other technology-based industries have entered the field of learning technology development, they have brought an approach that operates on "Internet time." Instead of starting small, they seek to launch "minimally viable" online learning products at scale so that they

© Copyright 2015 SRI International.

can collect and then mine large amounts of user data for insights into how the user experience and learning effectiveness can be improved. Software is being designed in ways that permit developers to embed random-assignment experiments (called "A/B experiments" in the tech industry) in their systems and collect data from their large user base at an unprecedented pace, permitting more cycles of implementation, redesign, and improvement than would ever be possible for print-based instructional materials and interventions (U.S. Department of Education, 2013). These developers see no need to start with data on a small scale from a few users who may or may not be representative of their much larger target user base. Instead, they use huge user data sets to explore relationships between different user profiles, activities, and outcomes. Educational software development now often follows a model of starting at scale, and applying "agile" methods (Martin, 2003) to rapidly iterate based on data captured as students and teachers use the product or service.

We believe that learning technology R&D will benefit by combining earlier "staged" and newer "agile" approaches. Better technology-supported educational innovations can be created through rapid cycles of modification, testing, and refinement. But if a piece of learning software is intended for use as part of formal education, rather than as a learner- or parent-selected optional activity, learning technology developers do need to worry not just about what learners do within their online systems but also about external outcomes of using that product. In contrast to Web-based companies whose products are purely an online experience (Facebook, for example), developers of most learning technology products want to make claims that using their product will enhance an individual's functioning in settings other than the online system. Few education policymakers would be tempted to purchase an online system that improves performance only on some piece of software (for example, a mathematics game involving fractions) if it does not also have positive effects on external, longer-term outcomes, such as achievement test scores or performance on new tasks requiring the same understanding or skill. In order to know that a product produces these transferrable outcomes (i.e., "external validity"), learning technologies need to be subjected to research designs that collect data on outcomes outside of the software system—the kinds of designs that are the expertise of educational researchers.

We believe that conventional research, the high-tech industry, and educators all bring important perspectives, and that collaborations are most likely to be productive when they include researchers with a set of experimental tools, developers with data analytics tools, and educators with knowledge of practice. We illustrate the different evidence perspectives of these various roles and the ways in which different perspectives can be brought to bear for a specific purpose in the R&D examples described below.

© Copyright 2015 SRI International.

Role 1: Researcher: Testing learning concepts within technology-based learning environments

Learning technology researchers view digital learning environments as ideal places for testing out theories of learning and ideas about how to foster better learning, including the acquisition of 21st century skills. One long-standing example of sustained learning technology research is "Computer-Supported Collaborative Learning," a research community that held its first conference in 1995, has spread to include significant participation in Europe, Asia, South America and North America, and recently published an extensive set of theoretical syntheses (Kirschner and Erkens, 2012. See also the accompanying articles in the same special issue). Collaboration is widely recognized as a key 21st century skill (Trilling and Fadel, 2009). The field of Computer-Supported Collaborative Learning (CSCL) examines the interlinked roles for technology and learning processes in supporting students' development of collaboration skills in conjunction with student learning of content knowledge. Collaboration is seen as involving motivational, social, and cognitive aspects, centered on mutual contribution towards the construction of shared knowledge (Teasley and Roschelle, 1993). Technology contributes by enabling the behaviors and cognitions presumed to foster better collaboration—by providing tools for group work, enhancing students' joint attention to the key elements of knowledge work in the given domain and task, and guiding students via "scripts" which help organize their activities by having a distinct role for every student (Kirschner and Erkens, 2012). Basic principles of collaborative learning connect interpersonal and technological aspects of tasks and include integrating task and social dimensions in joint work (Barron, 2003), a focus around training students to ask for and to provide help to each other (Webb and Palinscar, 1996), and creating conditions where every student must contribute and is only rewarded when the group succeeds (Slavin, 1990).

Group Scribbles is an example of a technology that has catalyzed a program of research spanning research groups in the United States, Asia, and Europe, and that has yielded both theoretical insights and evidence of effectiveness in classrooms. Developed at SRI's Center for Technology in Learning, Group Scribbles is a network technology that supports classroom collaborative activities using text, sketches, and images. The technology provides electronic versions of common physical artifacts such as adhesive notes, whiteboards, pens and markers. Participants can scribble contributions on electronic notes and post them anonymously in a shared public space that becomes the object of discussion. Teachers can quickly configure Group Scribbles spaces for a short-term collaborative or group activity and, as the activity unfolds, alter the configuration on the fly.

Researchers at SRI developed Group Scribbles in order to have a relatively generic collaborative learning tool that could be reconfigured rapidly to support different activities and learning different subject matter content. This research developed a theoretical connection between theories of distributed computing and theories of learning, with the concept of "coordination" as a bridge (Patton, Tatar, and Dimitriadis, 2008). Just as computers can be seen as a set of distributed processors which coordinate by writing and reading messages from a common message board, students can be seen as distributed processors capable of coordinating by reading and writing messages on a white board. In both cases, it is desirable

© Copyright 2015 SRI International.

for the processors to be working independently at all times (that is, active learning for the students) and yet progressing towards common goals (understandings of subject matter for the students). The original SRI-based work with Group Scribbles established the viability of this analogy and that software based on the metaphor of students posting and taking “notes” from a public space could support a wide variety of collaborative learning activities in a variety of content areas (Roschelle et al., 2007).

Researchers at the University of Valladolid in Spain then introduced Group Scribbles into Spanish primary school classrooms and helped teachers turn their lesson ideas into Group Scribbles activities. The researchers documented teachers’ practices and the way those practices were affected by the introduction of the Group Scribbles technology. The researchers documented implementation challenges and how teachers responded to them, further developing theory around how teachers “orchestrate” collaborative learning in the classroom (Prieto et al., 2007).

In the United States, researchers found that teachers were not leveraging fully Group Scribbles’ capability to elicit student thinking and foster knowledge construction. Teachers often asked students to explain their ideas, but the teachers did most of the intellectual work of building on and connecting ideas and rarely actively engaged students in discussion of one another’s ideas. Further, the Group Scribbles developers had hoped that by having students externalize their thinking, the technology would provide teachers with information that teachers could use to adjust their instruction in response to students’ ideas. But in practice, such responsive instruction was relatively rare. To address these weaknesses identified in the early work, SRI researchers developed a set of classroom norms that would complement the implementation of Group Scribbles. They also developed a set of rules for “contingent pedagogy,” specifying what a teacher should do in response to different possible student thinking patterns (DeBarger et al., 2010).

SRI researchers conducted a field test to see whether student learning could be improved with a combination of norms, contingent pedagogy elicitation questions, and teacher training in adaptive instruction and discussion facilitation – thus implementing a kind of metacognitive pedagogy. Of the 19 middle school teachers in the study, 12 received the contingent pedagogies professional development and associated technology, and 7 were in the comparison group. The research showed a significant difference in learning gains between treatment and control classrooms, as measured by pre- and post-tests of target earth science content.

Concurrently, researchers in Singapore were investigating how to transform traditional classrooms into collaborative workspaces, and decided to scale up the use of Group Scribbles in science and Chinese language classes there. Along with partners in the United States and Taiwan, the Singapore team significantly upgraded the Group Scribbles software. They also developed an approach to teacher professional development and extensive data analysis tools for researchers. When specific opportunities to integrate Group Scribbles into the curriculum were designed and teachers had sufficient training, the Singapore team found significant positive effects of using the technology, both for content learning and for transforming the classrooms from quiet, individualistic settings to collaborative workspaces (Looi and Chen, 2010; Looi, Chen, and Ng, 2009). A theoretical perspective regarding how classroom communities can rapidly build group knowledge with Group Scribbles was another key outcome of the Singapore work (Looi, Chen, and Patton, 2010).

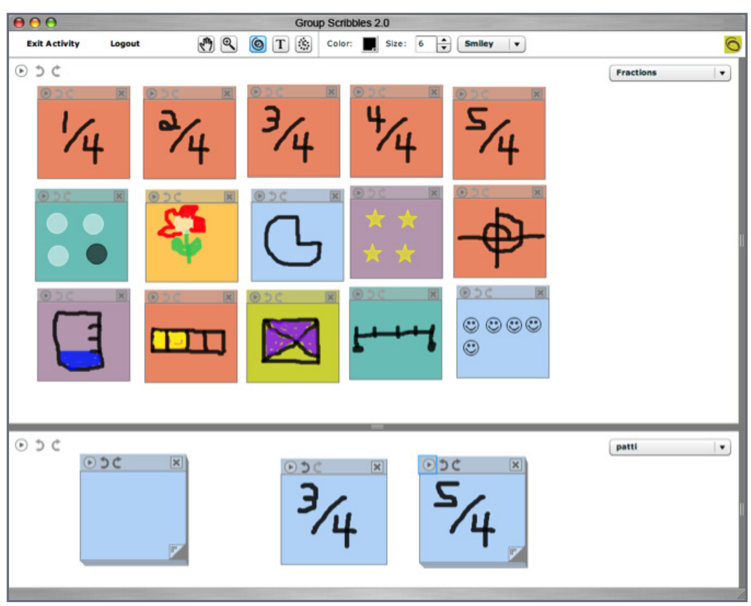
© Copyright 2015 SRI International.

In addition to the American, Spanish, and Singaporean research with Group Scribbles described above, the collective of researchers working with Group Scribbles has included scholars in Taiwan, Thailand, and the United Kingdom, who have pursued a wide variety of research questions and empirical activities. Key lessons learned include the theoretical value of coordination as a construct that structures the role of technology in collaborative learning, the design characteristics that enable a collaborative learning tool to work with many types of activities and subject matters, the necessary integration of teacher training and curricular focus in order to scale up effective use of a collaboration technology, the elaboration of an appropriate "contingent pedagogy," and principles for organizing classrooms to engage in rapid knowledge building.

The researchers working with Group Scribbles generally adopted a design-based, ethnographic approach that led to measurable improvements in teaching and learning. This collaborative model, sometimes referred to as a design-based implementation research approach (Penuel et al., 2011), leveraged international findings and allowed the gradual improvement in the design of the tool and accompanying resources and protocol uses, fostering better coordination of supports around the tool. This approach also helped surface and confirm important design tensions that arose from conflicting goals or values (Tatar, 2007), which also contributed substantively to an understanding of how to best orchestrate activities of teachers and learners across international settings.

More generally, researchers see learning technology as a medium for testing ideas about learning and instruction that apply across many different contexts and technology-based systems. They seek to set up contrasting conditions that allow them to isolate the effect of specific variables, such as training in classroom norms or making the work products of student groups available to the teacher. Researchers also investigate how theories of learning can be applied to explain learning with technology and explore how theories can be applied to new technological capabilities to enhance learning. Research insights accumulate over time and across research teams, gradually producing principles that can be used in the design of new learning technology innovations (Kali and Linn, 2010; Koedinger, Corbett, and Perfetti, 2012).

Figure 1: Representing Fractions Using Group Scribbles



© Copyright 2015 SRI International.

Posting Notes to a Whiteboard: A Powerful Metaphor for Classroom Collaborations

Group Scribbles was designed to be an easy-to-use technology for classroom collaboration which would also support a wide range of planned and improvised collaborative activities and a wide range of subject matter content. The essential metaphor in Group Scribbles is writing a small note and posting it to a public space. This familiar metaphor makes it easy for teachers and students to figure out how to use Group Scribbles in teaching and learning many different topics. The essential features for collaboration enabled by Group Scribbles are "taking" notes back from the public space to a private space and rearranging notes on private, small group, and public boards. The abilities to take items back and to organize items on multiple boards enable a wide range of learning activities.

For example, Figure 1 below shows a Group Scribbles activity in which a classroom is challenged to come up with a set of different ways to represent a set of fractions. Initially, the teacher posted only multiple copies of the numerical fractions (i.e., ¼). Using the Group Scribbles display on their device and a stylus, small groups of students "took" a fraction from the public board to their own group space. By taking a particular fraction, a group could indicate which fraction they intended to work on; collectively the "take" operation ensured that the classroom of students chose a variety of different fractions to work on, because there were not enough notes with "¼" for every group. Within a group, students then sketched representations of their fraction and posted these to the public space in the correct column. As students posted creative representations, other students were inspired to also be more creative.

The appropriate pedagogy for Group Scribbles is contingent on what students do. For example, a teacher might observe that students in a low-performing classroom made mistakes in representing particular fractions, such as 5/4. If this is what happens, a teacher might ask students to collectively identify which fractions were wrong (note that the posts do not have the students' names, so no one is embarrassed by making a mistake). Further, students can "take" mistakes off the board and fix them. In another classroom, students might get all the fractions right, and a teacher might lead a discussion of when a particular representation might be most appropriate. Is one representation best for "counting" type contexts and another best for "continuous variation" type contexts? Why might a number line be the least ambiguous representation of fractions greater than 1?

Although this example emphasizes mathematics, in other Group Scribbles classrooms students wrote Chinese characters, in an activity that focused on understanding which characters share similar common elements, called "radicals;" or posted hypotheses for what causes the temperature of the human body to rise, with each small group of students taking a cluster of related hypotheses and planning an investigation.

Unfortunately, adoption of Group Scribbles was initially slowed by the lack of availability of common classroom devices with a stylus; it is hard to "scribble" with a mouse or trackpad. However, as new tablet and mobile devices enter the classroom, implementing Group Scribbles at scale is becoming more feasible. When Group Scribbles is integrated into classrooms with appropriate designs for curricular activities and professional development for teachers, researchers have found impacts for both content learning and increased collaborative interaction among students (Looi and Chen, 2010; Looi, Chen, and Ng, 2009).

© Copyright 2015 SRI International.

Role 2: Developer: Observing usage to improve a learning technology

Learning technology developers seek to create software programs, learning systems, or platforms that will be used by students and educators. The technology industry approach to generating data for use in improving their products contrasts with conventional educational research in the nature and grain-size of the questions technology developers ask. Education researchers focus on generating evidence of the effectiveness of an innovation as a whole, comparing outcomes produced by the innovation to those produced by a different innovation or by “business as usual.” Technology developers start with the premise that having people use their product is a good thing, and design their data collections and analyses to explore questions about how to entice users to spend more time with it or how to achieve better outcomes on the learning measures that are internal to the system. A common industry practice is the use of A/B testing protocols (in which a random sample of users experience version A while other users experience version B) to test a range of variations of an intervention, exploring broad questions such as what types of “hints” work best for what kinds of learners or testing very fine-grained issues, often as subtle as alternative wordings for the hints shown at a specific point in concept acquisition.

An example of the way developers use A/B testing and the data generated when many learners use their product is provided by the Khan Academy. These mathematics resources have grown from a collection of a few hundred You Tube videos to the 3,500 math videos and 100,000 math practice problems available today (Murphy et al., 2014). Khan Academy software developers intentionally built their system to make it easy for them to run A/B tests on features they are considering changing. Because Khan Academy has so many users (10 million unique users a month in February 2014), any big effect of some variation in the way the software works can be detected in as little as an hour. More subtle differences can only be detected with more data points, but a week or so is the typical timeframe. An example of the kind of proposed change that Khan Academy evaluates through A/B testing was the change in their systems rule for deciding when a student has mastered the skill being developed by a set of problems. Before switching from their original rule that ten correct answers in a row signified mastery to a rule based on modeling the probability that a student’s next response would be correct based on the preceding pattern of correct and incorrect responses, the Khan Academy ran an A/B test. They knew that the new rule, which specified that a student had to have a 94% or better chance of success on the next trial to move on, would require some learners to have to spend significantly more time on a problem set. The A/B testing revealed that learners were willing to spend the extra time and that in fact learners earned more proficiency badges with the new rule than with the original.

In this way, developers can use A/B testing to gather strong evidence regarding the effects of proposed design changes. With online learning technologies, the only constraint on the frequency of changes is their potential to annoy existing users (and the size of this risk can be measured on a small number of users). Continuous improvement processes, with cycles of redesign, data collection, analysis, and refinement, can be implemented economically.

© Copyright 2015 SRI International.

What A/B testing does not reveal, however, is whether performance within the technology-based environment actually predicts how well a learner will do under other conditions, for example, when working on paper-and-pencil math problems or applying mathematical reasoning to everyday situations. Learning researchers are more likely than developers to concern themselves with learning measures external to the digital learning system.

Typically, developers and education researchers work separately, and sometimes developers are wary of education researchers who might perform what developers would consider unfair tests of their product's effectiveness compared to other instructional approaches. But developers and researchers can work together, and, as will be discussed later in this article, we believe that such collaborations often result in more effective learning technology products.

While learning technology researchers seek to test principles of learning and uncover generalizable design principles, technology developers focus on their particular product, striving to make it something people will use and value. By and large, they assume that use of their product is a good thing, and seek data about what parts of their product learners use and for how long, as well as about user perceptions of value. It is lesson common for developers to gather evidence that their product is more or less effective than an alternative approach. However, developers may try alternative versions of their own product, called "A/B testing," to see which gets the most favorable response. Driven by market considerations and cost consciousness, technology developers want information they can act on very quickly; by the time a decade of research has been conducted on their product, the market will have passed them by.

© Copyright 2015 SRI International.

Blended Learning Combines Roles for Technology and Teachers

Carnegie Learning offers a series of mathematics products aimed at improving students' abilities both to reason mathematically and to execute procedures correctly. A central element in most of Carnegie Learning's products is a computer-based tutor, which can track each student's mathematical progress and intervene whenever their work deviates from the expected trajectory. Students typically work alone on the tutor, using a computer either in a lab or a normal classroom setting. The tutor (a computer program employing artificial intelligence techniques) sequences mathematical problems for the student, and as the student works on a problem, the tutor follows every step. These steps are compared to an expert solution method, and when the student's step differs from the expert model, the tutor provides the student with feedback. The role of technology, therefore, is to track student learning, provide feedback to students, and to adapt to students' needs.

In typical classroom implementations of Carnegie Learning, two non-technology elements are also prominent. First, Carnegie Learning provides a textbook. According to the company, this book was designed based on research principles and specifically to address curriculum standards not covered by the tutor. The textbook is intended for use alongside the technology, with approximately 3 hours a week spent on activities described in the textbook and two hours spent using the computer-based tutor. Students may either go to a computer lab to use the tutor or rotate to a tutor station within their regular classroom. Second, Carnegie Learning provides extensive teacher training. This training emphasizes how to create a classroom that emphasizes growth in students' mathematical thinking. Collaborative learning principles and practices are the centerpiece of this teacher training. Consequently, Carnegie Learning blends technology-based instruction (emphasizing interactive feedback and adapting to individual students) with textbook and teacher-led instruction (emphasizing students' active engagement in mathematical thinking and collaborative learning). Although research underlying the technology component of Carnegie Learning's products has always been a very strong element of Carnegie Learning, studies have found a lot of variability in how blended implementations of all three components occurs. Not surprisingly, measurements of student outcomes have also been variable A recent randomized controlled trial of Carnegie's Algebra I Tutor found varied effects at the school level and no average effect across sites in the first year of implementation, but a significantly positive average effect in high school teachers' second year of using this software with their students (Pane et al., 2013).

© Copyright 2015 SRI International.

Role 3: Policymakers: Generating evidence for large-scale adoption

Education policymakers, such as administrators of local education authorities, differ from both researchers and technology developers in their relationship to educational technology. Their core responsibility is to serve the needs of the students under their authority. A technology that can help them better meet that purpose is a boon; one that does not serve that purpose is a distraction, and sometimes a costly one at that.

Education policymakers want to know which learning technologies are effective, but their concerns are more local than those of education researchers: What will work best in classrooms in the schools within my authority? What will help my region achieve more successful outcomes on the measures for which I am accountable? Because educators are ultimately responsible for making an innovation work in practice, they are also likely to look beyond generic data on product effectiveness to consider how well the product fits with their curriculum, current and desired pedagogies, and ICT capacities. They are less interested in evidence of effectiveness "on average" or under ideal conditions than they are with cases of positive outcomes in contexts similar to their own.

At the same time, studies of decision-making by local education authorities in the United States suggest that empirical research typically is not a major factor. Education leaders are more likely to choose innovations on the basis of the recommendation of peers. They appeal to research most often in cases where it supports a decision they have already made (Coburn, 2003). So we should not take the policymaker's quest for evidence for granted.

Although initiatives exist to encourage local authorities to weigh experimental research in their decisions, random-assignment experiments on technology-based innovations have some limitations. First, they can be expensive and time-consuming to implement. In addition, as Cronbach (1975) pointed out years ago, experimental designs often sacrifice elements of external validity (e.g., the sampling of the full range of contexts to which one would like to generalize) for internal validity (i.e., the ability to attribute differences in outcomes to the experimental manipulations). Although there is nothing in the essence of experimental designs that requires treating the experimental condition as a "black box," in practice there is a disturbing tendency to do so (Haertel and Means, 2003). When this is done, the research may show that the innovation did or did not have an impact on average, but provides little guidance for the policymaker who needs to understand how best to implement the innovation in his or her local schools to make sure that the necessary teacher training, infrastructure, and policies are in place to maximize the likelihood of successful outcomes in those specific contexts.

With sufficient research funding, it is possible to conduct studies that adhere to high standards of internal validity while paying attention to external validity and implementation issues as well. This is the approach that SRI has taken in a series of studies and implementations of mathematics units incorporating the SimCalc software, originally developed at the University of Massachusetts, Dartmouth. From the beginning, the goal of the SimCalc work has been "democratizing access to advanced mathematical concepts" so

© Copyright 2015 SRI International.

that more students could learn the foundational concepts underlying Algebra and Calculus at earlier ages and more deeply, in other words, giving them 21st century skills in the area of mathematics. SimCalc anchors student learning in familiar experiences of motion and engages students in activities in which they must use mathematics to analyze and control the motions of animated figures. The technology makes it possible for students to model more realistic motions using piecewise linear functions, a type of mathematical function that is commonly used in engineering but not typically introduced before higher education. In early studies, the SimCalc software was used with small numbers of low-income students, and at the end of the experience, those students made significant progress in understanding high-order mathematics concepts that are on the pathway to calculus (Kaput, 1997).

SRI researchers wanted to build on this early work, developing a comprehensive innovation that would include not only the software but associated curriculum units, teacher professional development, and the use of assessments aligned to mathematics skills and concepts that are addressed both by the SimCalc curriculum and by the usual program of instruction. They also wanted to show that the resulting innovation could be implemented by a wide variety of teachers working with students from diverse backgrounds, not just self-selected "adventurous" teachers. Between 2005 and 2007, SRI conducted a randomized controlled trial involving 95 seventh-grade math teachers and 1,621 students across all regions of the U.S. state of Texas. Seventh-grade students in classrooms using the integration of SimCalc curriculum, software and teacher professional development significantly outperformed students in classrooms using existing materials on a measure aligned both to Texas standards and to national and international achievement measures. What's more, equivalent advantages over their counterparts in control classrooms were found for girls and boys, Hispanic and non-Hispanic students, and higher- and lower-income students experiencing the SimCalc intervention.

Subsequently, SimCalc's effectiveness was demonstrated again in two quasi-experiments, involving an additional 538 seventh-grade and 825 eighth-grade students, also in Texas (Roschelle et al, 2010). Figure 2a shows SimCalc results as presented to policymakers. This representation illustrates the larger gains achieved by students experiencing SimCalc on a teacher-by-teacher basis. Policymakers find this kind of representation persuasive because it makes the amount of variation in gains across classrooms concrete, and gives them confidence that the vast majority of classrooms using SimCalc will learn proportionality and rational number concepts more deeply than their peers experiencing conventional instruction. These results led to a pilot and then a large-scale implementation in the United Kingdom, as well as a large-scale replication study in Florida. Students in U.K. classrooms using SimCalc obtained gains similar to those in the Texas classrooms, as shown in Figure 2b.

One of the hallmarks of the SimCalc research is its combination of a rigorous test of impacts, using what education research considers the "gold standard" methodology, and extensive studies of how the innovation is implemented by different teachers in different settings. For example, the research found that there were multiple "configurations" in which teachers could succeed with SimCalc. A teacher with strong mathematics content knowledge might lead her classroom in a different way than a teacher whose

© Copyright 2015 SRI International.

Figure 2a: Gains by Teacher

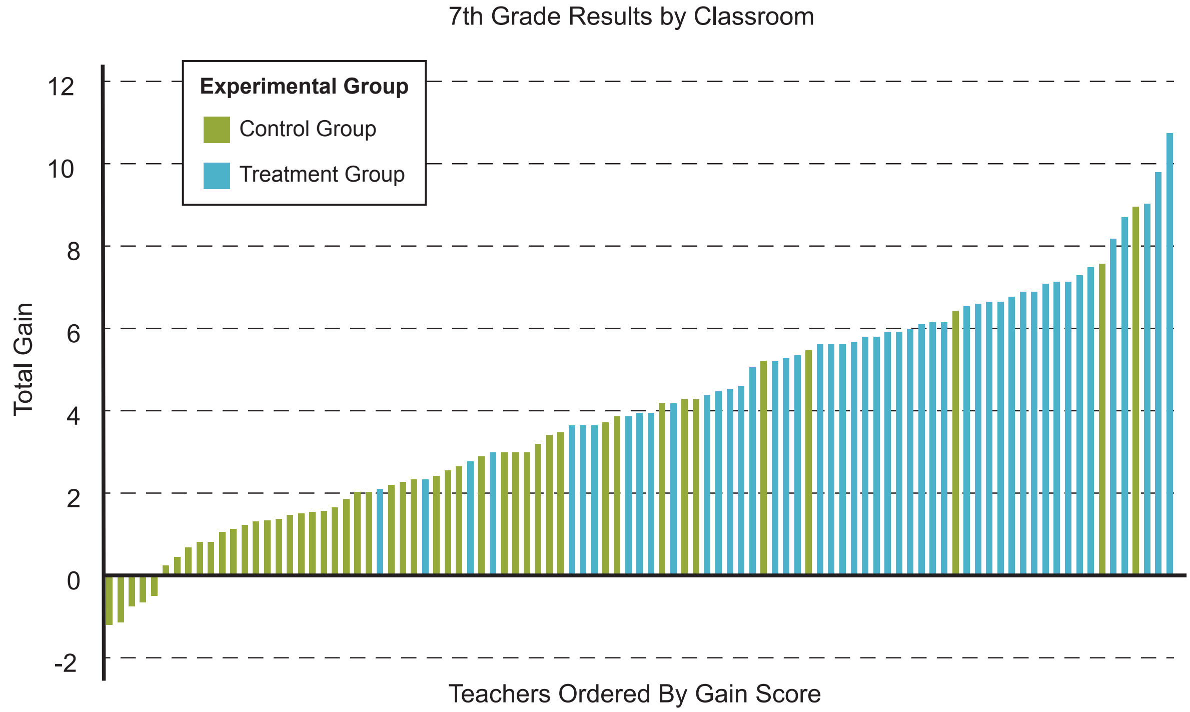
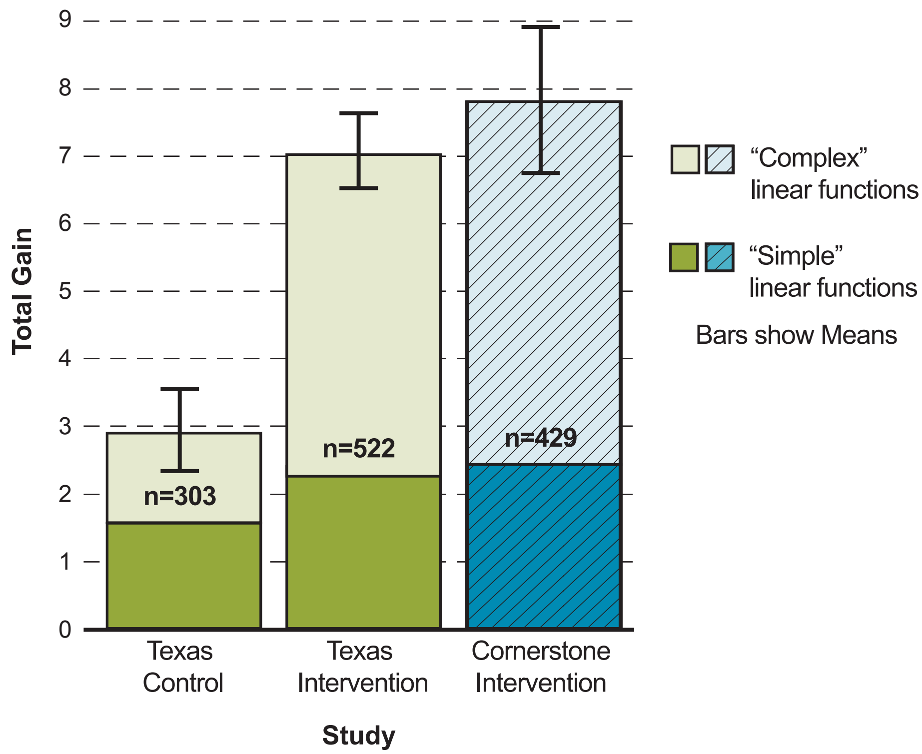


Figure 2b: Replicating Results



© Copyright 2015 SRI International.

strength lies in skills in maintaining the engagement and perseverance of struggling students (Empson et al., 2013). Conversely, there were a few critical factors among teachers whose classrooms did poorly with the intervention, such as taking too many liberties with the recommended timing and sequence of activities or failing to emphasize the big mathematical ideas in classroom discussions (Dunn, 2009). A recently published book presents the wealth of SimCalc outcome and implementation data and associated insights gained through many years of research in many settings (Roschelle and Hegedus, 2013). As noted above, this latter type of information is important to education policymakers responsible for planning and supporting the implementation of the technology-supported innovation.

Both researchers and policymakers look for evidence that a technology-supported intervention will have positive impacts, but many researchers are satisfied with statistically significant differences in average outcomes that coincide with predictions made from theory (with results portrayed as the difference between averages, as in Figure 2b). Policymakers, on the other hand, often worry about the robustness of effects—what proportion of classrooms will benefit from this intervention (that is, the kind of data displayed in Figure 2a). They may also be concerned with the cost of delivering the intervention at scale and whether the intervention might not work in certain settings or with particular populations. Policy makers often want to understand the support systems and intervention practices necessary to achieve consistently positive outcomes for all of the classrooms under their supervision, at reasonable cost.

© Copyright 2015 SRI International.

Achieving Robust Outcomes by Engineering a Curricular Activity System

As with all the examples cited in this paper, the results of the SimCalc experiments do not prove that "technology improves learning." Instead, the team has consistently preferred to view the positive effects that have been observed as products of the complete "curricular activity system" (Roschelle, Knudsen, and Hegedus, 2009). The system is "curricular" because the SimCalc approach deeply rethinks which mathematical concepts are emphasized and in what order they are presented. For example, it is common in the United States to restrict the topic of "proportions" to the formula "a/b = c/d." The SimCalc authors realized, however, that the topic of proportions could also be an ideal jumping off point for Algebra, if the function "f(x) = kx" was also introduced, where "k" is the constant of proportionality (or the rate of change). A component of the curricular activity system engineering, therefore, is to reconsider how concepts build upon one another in a curriculum. This engineering is made possible because technology gives students new ways to experience concepts that otherwise might be too difficult. For example, students can experience the role of the constant "k" by making a graph steeper and seeing the resulting increase in how fast a simulated soccer player runs across a field.

The SimCalc system emphasizes "activities" and not lesson plans or sets of exercises as the fundamental unit of design. An activity sets up goals for students and ways that students can act upon the technology to make progress on their goals. For example, in SimCalc, an activity can be to make a "race" that ends in a tie by manipulating a graph that is linked to the racers' movements. In a well-designed activity, students do not need the teacher or the computer to determine whether they are right or wrong; by pressing "play" they can cause the computer to run the race according to their graph, and they can see for themselves whether it was a tie or not. Of course, "designing a race" is not a mathematical learning goal – however, activities are chosen and sequenced to carefully align with learning goals and to evoke and work on student misconceptions. For example, a race that involves a position versus time graph that goes both up and down is often interpreted wrongly by students as a runner with faster and slower segments during the race, but experience in trying to make a successful race can teach students that an up-and-down position graph means "going forward, then going backward."

Finally, the object of design is a "system" and not just the technology. The system includes technology, but also the necessary curriculum workbooks, teacher guides, teacher workshops, and assessments, and all these elements are carefully aligned to reinforce each other. When technology alone is inserted into a classroom, only a few teachers are typically able to figure out how to deeply connect the technology, their government-mandated teaching goals, existing textbooks, and their own pedagogical approaches. In general, raw technologies rarely produce consistent effects in variable settings; only when a complete system is engineered do results become replicable across varying conditions. Thus, although the SimCalc story began with developing a technology based on learning science principles, the eventual robustness that the intervention achieved relied on engineering a complete curricular activity system around the technologies' unique affordances for learning.

© Copyright 2015 SRI International.

Role 4: Teachers and School Leaders: Generating evidence for peer adoption

Teachers and school leaders face decisions similar to those of education policymakers but on a more micro scale. Rather than making major technology purchases, school leaders and teachers typically make decisions about which of the digital learning resources available to them to use in lessons and how much of their students' and their own time to devote to learning technology use.

The task of deciding what technologies and applications to use has become more challenging as the range of choices expands daily. Especially with the advent of cloud-based computing and the dizzying array of open educational resources available for free, educators have more choices than they have time to investigate (OECD, 2009; Hylen et al., 2012). These resources range from textbooks available online to online courses to supplementary materials such as individual lessons, simulations, games, and interactive graphics.

Both nonprofit organizations and for-profit companies have identified educators' need for support in this area, and there are many portals and collections of digital education resources in the United States and in many other OECD countries (see box on the right). While such collections provide teachers and school leaders with a single site where they can review and organize many digital learning resources, educators still need to decide what they will use from such a site. Certainly, educators want high-quality resources that will help their

A Sampling of Collections of Digital Education Resources

- *Curriki*, one of the first players in this arena, was set up to allow users to share education resources with the goal of lowering economic and geographical barriers to learning. Resources on Curriki have been reviewed for quality by an in-house team.
- *BetterLesson* is a curriculum-sharing platform available to individual teachers for free and to local education authorities for a subscription fee. This platform contains 300,000 teacher-created lessons that can be browsed using key words and tools that can be used to create collections and "playlists" of favorite lessons.
- *Gooru* is a nonprofit organization offering students and teachers access to over 50,000 educational resources in math and science for grades 5 through 12. Gooru staff review open educational resources, selecting those they judge to be of high quality and related to California's science content standards and the U.S. Common Core State Standards in mathematics.
- *LearnZillion* is a startup company with a platform that provides free access to 2,000 teacher-created lessons incorporating video explanations, assessments, and progress reporting.
- In Korea, the *EDUNET* platform supports the distribution and use of high quality digital contents: its quality is assured by KERIS, which directly manages the development of digital content by setting standards, models and guidelines.
- In the Netherlands, *Wikiwijs* combines a resource bank for open educational resources, referrals to other digital educational resources and a collaborative space where teachers develop and personalise open educational resources (OER). The quality of content is ensured through peer-review, reputation effects, and referrals from trusted third parties (such as teacher training institutions).
- In the Dutch language space, *KlasCement* also constitutes a popular web portal for teachers (and by teachers).
- In France, *Sesamath* is a popular exchange platform and repository of resources for mathematics teachers: it claims to have received about 14 million visits in 2012 (Avvisati et al., 2013).

© Copyright 2015 SRI International.

students learn, but many of the new offerings have not been tested for their effects on student learning. By the time rigorous experimental tests of the learning outcomes of one of the digital learning resources have been conducted and made available to the public, there are sure to be many newer, potentially more innovative options available. The National Research Council (NRC) in the U.S. encountered this difficulty when it set out to examine the research supporting the hypothesis that online games could foster learning of academic content. The NRC applied its usual approach of convening a panel of research experts, having the panel invite presentations from researchers and developers of serious games, examining the available research on the impacts of learning games, and writing a book-length report (NRC, 2011). Once it was published, the report was met with immediate criticism that the games the NRC had studied were "old-school" and "out-of-date."

Educators who want to use digital learning resources to foster educational innovation are often going to be making choices without research evidence about learning impacts. Like policymakers, educators are inclined to place high value on the opinions of their peers about the effectiveness of an educational innovation. Testimonials from fellow teachers about the effectiveness of a technology-based innovation are, from the educator's perspective, a form of evidence. Increasingly, the Internet is broadening access to peer judgments of effectiveness through online product reviews and ratings.

In addition to aggregating teachers' perceptions of effectiveness, online reviews of technology products and innovations are addressing other aspects of the innovations that educators consider in selecting which ones to try. For example, they often indicate whether the resource:

- addresses a particular learning goal or educational standard;
- can be used by all students including those with vision or hearing disabilities;
- reflects a particular theory of learning;
- engages students;
- is easy to use or requires preparatory training;
- can be used with the technology infrastructure currently available;
- fits within the time available for teaching the learning content it covers;
- costs money to acquire or use over time.

The *Curriki* platform described above invites users to rate and comment upon the resources it contains. Similarly, *Classroom Window* includes a "report card" review template that asks people who have used a learning technology product to provide numerical ratings or select a description from a drop-down menu for specific product features. It also invites reviewers to give a ranking for how effective they thought the product was for different kinds of students. Reviews are aggregated in much the same way that book ratings are aggregated and reviews provided by companies such as *Amazon*. *EdShelf* solicits user reviews of learning resources in terms of ease of use, student engagement, and learning effectiveness. *Edmodo* is designed to support communities of teachers interested in the same learning resources. Teachers can engage in online discussions of their learning resource choices and the issues they are encountering in trying to implement them with students.

© Copyright 2015 SRI International.

Clearly, the number of online platforms and repositories designed to help educators locate and select learning resources is large and growing. These platforms are helping to make insights from practice available to a much larger community of educators.

What the platforms have not had, at least thus far, is evidence of the learning impacts of the digital resources that would stand up to scrutiny. Some of the platforms, such as *PowerMyLearning* offered by the nonprofit organization CFY, aspire to fill this gap by running online experiments in which students are given different learning resources targeting the same learning objective and results of online assessments are used to identify the most effective product. But rapid running of online tests of the effectiveness of new digital learning resources and posting results for educator use is still some time off. So is the identification of the best learning and teaching resources given information about teachers' students as gathered by longitudinal information systems (OECD, 2010a). It also remains to be seen whether teachers would choose to use this kind of information in preference to peer reviews even if it were available. An investigation of teachers' use of the research summaries published by the What Works Clearinghouse, for example, found that fewer than 5% of United States teachers had ever gone to this resource to look for evidence of effectiveness (U.S. Government Accounting Office, 2010).

Teachers and school leaders often use forms of evidence that are weaker than those considered to be valid by researchers, but the importance of teachers' perceptions should not be underrated. Often technologies can have the theoretical potential for positive impacts, but implementation is too time-consuming, difficult, or mis-aligned for practical realization of that potential. When large numbers of ratings from fellow teachers working with similar kinds of students are available, a teacher can be pretty sure that any major flaws or implementation difficulties of a learning technology product will have been spotted and commented upon by multiple people. At the same time, there would be value in getting more of the research-based learning technologies with good evidence of effectiveness onto the platforms and into the digital repositories that teachers turn to most frequently.

Design Collaboratives Instead of Ratings?

Despite many efforts to develop a "rating service" with the singular authority and utility of the ratings available for cars, home appliances, or restaurants, there is still no similarly successful ratings service for educational technologies. It may be that ratings work well when products are more standardized into categories and when the typical applications of such products has also become conventional. It appears that the use of technology in schools is still too variable and the identity of products is still too much in flux for this kind of rating system to find a large audience. The best answer to many obvious questions, such as "should a region invest in electronic whiteboards or tablets for every student?" is still "it depends." It depends on what content will be available for use with the whiteboards or tablets, how it will be integrated with everyday choices about learning materials, how teachers will be trained, and how outcomes will be measured (see for example Avvisati et al., 2013).

Tension may also arise because people tend to see educational technology as an intervention when its true characteristics are often more infrastructural. "Should a region build wider roads or a more extensive rail

© Copyright 2015 SRI International.

system?" is another infrastructural question to which the only reasonable answer is "it depends." Once a decision is made to build a rail system, however, much is then known about the next implementation steps that will enable the rail system to be successfully used to produce desired economic and social outcomes. Educational technology, argued James Kaput (the inventor of SimCalc), is likewise infrastructural: it produces the possibility of moving students' learning along new and more efficient pathways, but not the certainty.

The "ratings" analogy may also go astray by taking too conventional a view of educational publishers as "producers" and teachers and students as "consumers." Presently much energetic activity takes a more complex view. In the "Digital Promise" program (http://www.digitalpromise.org/), for example, schools with like needs for innovation are forming cooperatives to increase their power to advocate for the development of matching educational technology solutions and are participating in innovation process. Likewise, schools in the "open educational resources" movement are moving to self-publish open-source alternatives to traditional content offered by publishers (http://www.oercommons.org/). Although there is obviously an upfront investment, open educational resources can be thereafter used for free, and anyone can improve upon them.

Thus collaborative partnerships to develop local innovations may become more useful than ratings of products developed elsewhere. A STEM education example of a collaborative partnership is the Pathways Project initiated by the Carnegie Foundation for the Advancement of Teaching. Carnegie researchers are working with educators from two- and four-year colleges in the United States on the problem of getting students who come to college without the preparation needed for a college level mathematics course into and successfully through a college-level mathematics course. The stated goal was for twice as many students to earn college math credit within one year. The colleges that Carnegie convened agreed to collaborate with other colleges and with researchers and developers to analyze their developmental mathematics failure rates and then redesign their approach to developmental mathematics through collaborative development of new courses and associated polices, followed by analyzing system data and feedback from early implementations and refinement of the courses. In this analysis the collaborators recognized that many students were not successfully transitioning between courses in the developmental math sequence, and that existing course designs did not sufficiently engage students. They agreed to replace the series of separate courses with a single two-semester course emphasizing real-world problems.

Two new developmental math courses were developed by the Pathways Project: Statway (emphasizing statistics) and Quantway (emphasizing quantitative reasoning. The process used in developing Statway illustrates the iterative nature of co-design by educators and researchers. The first version of Statway was produced by a small group of academic researchers and curriculum developers. A number of two-year college faculty reviewed the course materials and tried some of the modules with their classes in fall 2010. Their reflections on this experience, along with input from a larger group of researchers and course designers, led to the conclusion that the materials needed a major revision before being implemented on a wider scale. A team of mathematics faculty from multiple colleges came to Carnegie to work together on redesigning the course, and the result was Statway Version 1.5, which was tried out by all of the colleges in the project in school year 2011-12. In the first year of Statway implementation, three times as many students earned a college math credit in one-third the time compared with historical averages at the participating colleges (Strother, Van Campen, and Grunow, 2013).

© Copyright 2015 SRI International.

A *Rapprochement*

The above discussion highlights the different goals, methods, and timeframes for seeking evidence about value of learning technology, and the different perspectives on evidence based on the roles of researcher, developer, policymaker, and educator. These differences make it hard to summarize what is or is not known about learning technology's value and complicate collaborations among stakeholders in learning technology. Nevertheless, we see increasing numbers of instances in which people in different roles have come together to design evidence-gathering activities that serve all their needs.

One of the best examples of researchers and commercial developers working together is the collaboration between academic researchers at Carnegie-Mellon University and the developers at Carnegie Learning, a for-profit company started in 1998 with the express purpose of creating and distributing tutoring systems based on the learning theory of Carnegie-Mellon's John Anderson. In 2011, Carnegie Learning products were used by over 600,000 students in 44 states in the United States (Gabriel and Rachtel, 2011).

Since its founding, Carnegie Learning has continued to draw on academic research and to maintain close ties with the university. An example of research-developer collaboration around Carnegie Learning products is the *Geometry Cognitive Tutor*. Researchers at Carnegie-Mellon were interested in applying research-based principles for designing multimedia to exercises within the Geometry Cognitive Tutor that requires students to use their knowledge of geometric theorems to calculate angle measures within a diagram. In initial program versions, the diagram was static: students used it to find angle relationships and values that they then entered into a separate table, along with the relevant theorems (Aleven and Koedinger, 2002).

Researchers (Butcher and Aleven, 2008) were concerned that the separate presentation of the table and the diagram might add extraneous cognitive load, which could in turn hinder student learning. Butcher and Aleven identified underlying psychological principles (the split-attention effect (Kalyuga, Chandler, and Sweller, 1999) and the contiguity principle (Mayer, 1989)), and used them in a series of design experiments that resulted in an updated version of the software. The new version used an interactive diagram into which students could enter their calculated angles and reasoning without need for a separate table. The researchers used an A/B test to compare student performance between the "table interaction" and "diagram interaction" versions of the *Geometry Cognitive Tutor*. The diagram interaction version was shown to be more effective in tests of transfer, which asked a new type of question that leveraged the learned content without being directly instructed (e.g., whether the diagram could be used to compute a particular angle). In delayed posttests, the diagram interaction condition was also shown to be more effective at promoting student retention of how to compute angle values using geometric theorems.

© Copyright 2015 SRI International.

Based on these results, Carnegie Learning implemented the diagram interaction version as a standard feature of their commercial ***Geometry Cognitive Tutor*** product, allowing them to continue data analysis with a much larger user base. In comparison to previously collected data from users of the older version, users of the new diagram-interaction version reached mastery more quickly, a result that was stronger for difficult problems (Hausmann and Vuong, 2012).

This example illustrates the benefits that can be derived from a combination of academic research and commercial-quality development and distribution. Researchers applied psychological principles to learning software redesign and tested the effects of their redesign on a modest scale, taking advantage of the fact that they were working with a digital learning system to conduct A/B testing relatively rapidly. Based on the findings from this research, Carnegie Learning implemented the redesign within its operational system and mined the data from many users receiving either the old or the new version of the angle-measure diagrams.

Researchers, policymakers, and educators (teachers and school leaders) also are finding ways to work together around the implementation of learning technology. Design-based research, in which learning technology researchers collaborate with classroom teachers to design technology-based interventions that test learning principles and can be implemented in real classrooms, with their many constraints in terms of time, place, and teacher focus, has a long tradition (Sandoval & Bell, 2004). More recently, learning scientists have articulated a collaborative research model they call "design-based implementation research" (DBIR, Penuel et al, 2011). Under this model, researchers and educators collaborate as equal partners in a joint effort to address a "persistent problem of practice." Under the DBIR model, the goal is to produce measureable improvement in educational outcomes, a purpose that is much more aligned with the educator's role than has been the theory testing that has been the sole objective of much academic research. DBIR researchers do bring theoretical concepts to bear in framing and addressing problems with their educator-collaborators, but theory is not allowed to overshadow practical concerns or the wisdom of practice. In such collaborations, as illustrated by the Pathways Project described in the Design Collaboratives box, educators participate in more rigorously designed research and have access to better data about implementation and outcomes within their schools because they have the support of professional researchers. Multiple iterations of design, implementation, data collection and analysis and redesign are central to DBIR (Fishman et al., 2013).

© Copyright 2015 SRI International.

Conclusions

As this paper has described, the intuitive appeal of technology to support deeper STEM learning has strong support from both research and practice. Building on established research principles, well-designed technology tools can be a powerful enabler of learning environments in which students drive and manage their own learning, build on their intuitive ideas to develop rich and integrated understandings of the big ideas of a domain, build knowledge through productive collaboration with peers, and receive actionable personalized feedback toward their academic progress. These outcomes are not an automatic result of adding computers to learning, as is too often assumed when adoption decisions are made. Successful results for student learning require attention to learning principles as programs are designed and implemented. They also require attention to a comprehensive system of supports, including curriculum, pedagogy, professional development, assessment, and leadership. Evidence from research can help to make clear what designs and conditions are linked to strong outcomes for students.

The above descriptions of selected technology-supported systems for learning STEM content illustrate how such well-designed uses of technology also support students' development of the kinds of skills called for in 21st century workplaces and increasingly in STEM standards: such innovations can give students opportunities to collaborate more deeply and with more diverse teams, engage with more complex real-world problems, communicate to authentic audiences, and learn to regulate their own learning. Debate persists about whether these skills should be taught on their own or within subject areas, and if the latter, whether adding a focus on 21st century skills distracts from the core requirements of subject matter learning. We believe that 21st century skills can be both core enablers and results of the strongest kind of subject matter learning, if these opportunities are explicitly designed and supported to do so. Once again, evidence can help us to sort out these complex issues.

Technology is also rapidly transforming both the volume and the character of the available data on which decisions might be based. Instead of a future in which relevant educational research narrows to emphasize only randomized controlled trials, it seems much more likely that access to data is expanding. In this expansion, we expect sound uses of data to be differentiated by roles. Researchers will make uses of data that advance theory and yield insights. Developers will make uses of data that improve products. Policy makers will seek reliable data on which to base broad policies on resource use, and randomized controlled experiments and like evaluation methodologies will still be important for establishing both internal and external validity. Teachers, parents, and other stakeholders in local school settings will seek ways to make rational decisions, given imperfect information and local variation in needs, capacities, and goals.

As the availability of data and evidence unfolds, educators are encouraged to think about their grounds for choosing learning technologies in the context of two factors: (1) their confidence that the technology-based innovation can produce better student outcomes than the status quo and (2) the level of risk involved, in terms of costs, loss of valuable learning time, and disruption of existing processes and

© Copyright 2015 SRI International.

practices (U.S. Department of Education, 2013). In some cases, confidence in the effectiveness of a learning resource is high and risk is low—for example, when thinking about making a math game that has been found to boost students' multiplication skills and can be used after school (and thus does not displace normal instruction). In such cases, deciding to implement the innovation is an easy choice. In many cases, however, the choice is less clear-cut. Educators might have reasonably high confidence that a learning resource will produce good outcomes, say because teachers in another school provided testimonials of its positive impacts, but the risk of implementing it is high, perhaps because it requires an expensive upgrade in the school's technology infrastructure or because the software purchase price is high. Whenever the level of risk is high, educators should devise strategies to manage risk and to engage in data collection that will reduce their uncertainty about whether or not the innovation will produce desired outcomes. One strategy for increasing confidence in a technology-based innovation's effectiveness and reducing risk is to try it out with a limited number of students on a pilot basis, carefully observing implementation and results.

It should be noted, however, that avoiding all risk is not only impossible but also incompatible with the goal of educational innovation and improvement. Schools and classrooms will never be dramatically transformed without some degree of risk. We encourage educators to manage this risk by adopting the kind of iterative improvement processes used by technology developers to the enterprise of implementing technology-based educational innovations as well. In today's world, schools and other educational stakeholders can be collaborators in the design and continuous improvement of technology-enhanced and research-based educational innovations, not just the consumers of a select group of innovations that have been proven effective elsewhere. We believe that it is only through such collaborations that the gap between STEM education research and practice can be bridged through jointly envisioned development and implementation of innovations.

© Copyright 2015 SRI International.

References

Ainsworth, S. (1999), "The functions of multiple representations", ***Computers & Education,*** Vol. 33, No. 2, pp. 131-152.

Aleven, V. and K. Koedinger (2002), "An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor", ***Cognitive Science,*** Vol. 26, pp.147-179.

Avvisati, F., S. Hennessy, R. B. Kozma, and S. Vincent-Lancrin (2013), "Review of the Italian Strategy for Digital Schools", ***OECD Education Working Papers,*** No. 90.

Barron, B. (2003). "When smart groups fail", ***Journal of the Learning Sciences,*** Vol. 12, No. 3.

Barron, B. J., D. L. Schwartz, N. J. Vye, A. Moore, A. Petrosino, L. Zech, and J. D. Bransford (1998), "Doing with understanding: Lessons from research on problem-and project-based learning", ***Journal of the Learning Sciences,*** Vol. 7, No. 3-4, pp. 271-311.

Bereiter, C. (2002), ***Education and mind in the knowledge age,*** Lawrence Erlbaum Associates, Mahwah, NJ.

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

Blumenfeld, P. C., E. Soloway, R.W. Marx, J.S. Krajcik, M. Guzdial and A. Palincsar (1991), "Motivating project-based learning: Sustaining the doing, supporting the learning", ***Educational psychologist,*** Vol 26, No. 3-4, pp. 369-398.

Bransford, J. D., A. L. Brown, and R. R. Cocking (Eds.) (2000), ***How people learn: Brain, mind, experience, and school,*** National Academies Press, Washington, DC.

Bransford, J. D. and D. L. Schwartz (1999), "Rethinking transfer: A simple proposal with multiple implications", ***Review of Research in Education,*** Vol. 24, pp. 61-100.

Brown, A. L. and J.C. Campione (1996), "Psychological theory and the design of innovative learning environments: On procedures, principles, and systems" in L. Schauble & R. Glaser (eds.), ***Innovations in learning: New environments for education,*** Lawrence Erlbaum Associates, Mahwah, NJ, pp. 289-325.

Bruner, J. S. (1960), ***The process of education,*** Harvard University Press, Cambridge, MA.

Butcher, K. R. and V. Aleven (2008), "Diagram interaction during intelligent tutoring in geometry: Support for knowledge retention and deep understanding", in B. C. Love, K. McRae, and V. M. Sloutsky (eds.), ***Proceedings of the 30th Annual Conference of the Cognitive Science Society*** (pp. 1736-1741).

Butler, D. L. and P.H. Winne (1995), "Feedback and self-regulated learning: A theoretical synthesis", ***Review of Educational Research,*** Vol. 65, pp. 245-281.

CAST (2011), ***Universal Design for Learning Guidelines version 2.0,*** Wakefield, MA.

Chiu, J. L. and M. C. Linn (2012), "The role of self-monitoring in learning chemistry with dynamic visualizations", ***Metacognition in Science Education,*** pp. 133-163.

Clarke, J., and C. Dede (2009), "Design for scalability: A case study of the River City curriculum", ***Journal of Science Education and Technology,*** Vol. 18, No. 4, pp. 353-365.

© Copyright 2015 SRI International.

Clement, J. (1982), "Students' preconceptions in introductory mechanics", *American Journal of Physics,* Vol. 50, No. 1, pp. 66-71.

Coburn, C. E. (2003). "Rethinking scale: Moving beyond numbers to deep and lasting change", *Educational Researcher,* Vol. 32. No. 6, pp. 3-12.

Corbett, A. T., K. R. Koedinger, and W. H. Hadley (2001), "Cognitive Tutors: From the research classroom to all classrooms" in P. S. Goodman (ed.), *Technology Enhanced Learning: Opportunities for Change,* Lawrence Erlbaum Associates, Mahwah, NJ, pp. 235-263.

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology, *American Psychologist,* 30(2), 116-127.

Crouch, C. H. and E. Mazur (2001), "Peer instruction: Ten years of experience and results", American *Journal of Physics,* Vol. 69, p. 970.

DeBarger, A. H., W. R. Penuel, C. J. Harris, and P. Schank (2010), "Teaching routines to enhance collaboration using classroom network technology" in F. Pozzi and D. Persico (eds.), *Techniques for fostering collaboration in online learning communities: Theoretical and practical perspectives, Information Science Reference,* Hershey, PA, pp. 224-244.

Deslauriers, L., E. Schelew, and C. Wieman (2011), "Improved learning in a large-enrollment physics class", *Science,* Vol. 332, No. 6031, pp. 862-864.

diSessa, A. A. (1983), "Phenomenology and the evolution of intuition" in D. Gentner and A. L. Stevens (eds.), *Mental models,* Erlbaum, Hillsdale, NJ.

Dunn, M. (2009), "Investigating variation in teaching with technology-rich intervention: What matters in teaching and teacher training at scale?", *unpublished Ph.D. dissertation,* Rutgers, Newark, NJ.

Echeverría, A., M. Améstica, F. Gil, M. Nussbaum, E. Barrios, and S. Leclerc (2012), "Exploring different technological platforms for supporting co-located collaborative games in the classroom", *Computers in Human Behavior,* Vol. 28, No. 4, pp. 1170–1177.

Empson, S. B., S. Greenstein, L. Maldonado and J. Roschelle (2013), "Scaling up innovative mathematics in the middle grades: Case studies of 'good enough' enactments", in *The SimCalc Vision and Contributions,* Springer, Netherlands, pp. 251-269.

Fischer, F., I. Kollar, K. Stegmann, and C. Wecker (2013), "Toward a script theory of guidance in computer-supported collaborative learning", *Educational Psychologist,* Vol. 48, No. 1, pp. 56-66.

Fishman, B., Penuel, W. R., Allen, A., & Cheng, B. H. (2013). Design-based implementation research: Theories, methods, and exemplars. *National Society for the Study of Education Yearbook* (vol. 2). New York: Columbia University.

Fuchs, L. S. and D. Fuchs (1986), "Effects of systematic formative evaluation: A meta-analysis", *Exceptional Children,* Vol. 53, No. 3, pp. 199-208.

Gabriel, T. and M. Richtel (October 8, 2011), "Inflating the software report card", *The New York Times.*

Gardner, T.Q., S.E. Kowalski and F.V. Kowalski (2012), "Interactive simulations coupled with real time formative assessment to enhance student learning", in *ASEE 2012 Conference and Exposition Proceedings,* San Antonio, Texas.

© Copyright 2015 SRI International.

Haertel, G. and B. Means (Eds.), Evaluating educational technology: Effective research designs for improving learning, Teachers College Press, New York.

Harlen, W. (Ed.) (2010), Principles and bid ideas of science education, *Association for Science Education,* Herts.

Hattie, J. (2009), *Visible learning: A synthesis of over 800 meta-analyses relating to achievement,* Routledge, Oxfordshire, England.

Hausmann, R., and A. Vuong (2012), "Testing the split attention effect on learning in a natural educational setting using an intelligent tutoring system for geometry" in N. Miyake, D. Peebles, and R. P. Cooper (eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society, Cognitive Science Society,* Austin, TX, retrieved from http://mindmodeling.org/cogsci2012/papers/0087/index.html

Heffernan, N. T., C. L. Heffernan, M. B. Decoteau, and M. Militello (2012), "Effective and meaningful use of educational technology" in C. Dede and J. Richards (eds.), *Digital Teaching Platforms: Customizing Classroom Learning for Each Student,* Teachers College Press, New York, pp. 88-102.

Hiebert, J. and D. A. Grouws (2007), "The effects of classroom mathematics teaching on students' learning", in F. K. Lester (ed.), *Second handbook of research on mathematics teaching and learning, Information Age,* Greenwich, pp. 371-404.

Hylen, J, D. Van Damme, F. Mulder, and S. D'Antoni (2012), "Open Educational Resources: Analysis of Responses to the OECD Country Questionnaire", OECD Education Working Papers, No. 76, OECD Publishing. http://dx.doi.org/10.1787/5k990rjhvtlv-en

Inagaki, K. and G. Hatano (2002), *Young children's naive thinking about the biological world,* Psychology Press, New York.

Johnson, D. W. and R. T. Johnson (2009), "An educational psychology success story: Social interdependence theory and cooperative learning", Educational Researcher, Vol. 38, No. 5, pp. 365-379.

Kali, Y. and M. Linn (2010), Curriculum design as subject matter: Science, in B. McGraw, E. Baker, and P. Peterson (Eds.), *International Encyclopedia of Education,* Third Edition, Elsevier, Oxford.

Kalyuga, S., P. Chandler, and J. Sweller (1999), "Managing split-attention and redundancy in multimedia instruction", *Applied Cognitive Psychology,* Vol. 13, No. 4, pp. 351-371.

Kaput, J. (1992), "Technology and mathematics education" in D. Grouws (ed.), *A handbook of research on mathematics teaching and learning,* Macmillan, New York, pp. 515-556.

Kaput, J. (1997). "Rethinking calculus: learning and thinking", *The American Mathematical Monthly,* Vol. 104, No. 8, pp. 731–737.

Kirschner, P. A. and G. Erkens (2012), "Towards a framework for CSCL research", *Educational Psychologist,* Vol. 48, No. 1, pp. 1-8.

Koedinger, K. R., A. T. Corbett, and C. Perfetti, (2012), The Knowledge-Learning Instruction (KLI) framework: Bridging the science-practice chasm to enhance robust student learning, *Cognitive Science,* Vol. 36, No. 5).

© Copyright 2015 SRI International.

Kohl, P.B., H.V. Kuo, S. Kowalski, and F. Kowalski (2011), "Promoting and assessing creativity and innovation in physics undergraduates", ***Proceedings of the 2011 Physics Education Research Conference,*** August 2011, Omaha, Nebraska.

Krajcik, J. S., C. Czerniak, and C. Berger (2002), Teaching science in elementary and middle school classrooms: A project-based approach, McGraw-Hill, Boston.

Lesh, R. A. and H. M. Doerr (Eds.) (2003), ***Beyond constructivism: Models & modeling perspectives on mathematics problem solving, learning & teaching,*** Lawrence Erlbaum Associates, Hillsdale, NJ.

Linn, M. C. and B. S. Eylon (2011), ***Science Learning and Instruction: Taking Advantage of Technology to Promote Knowledge Integration,*** Routledge, Taylor & Francis Group, Florence, Kentucky.

Linn, M. C., H. S. Lee, R. Tinker, F. Husic, and J. L. Chiu (2006), "Teaching and assessing knowledge integration in science", ***Science,*** Vol. 313, No. 5790, pp. 1049-1050.

Looi, C. K., W. Chen, and F-K Ng (2009), Collaborative activities enabled by GroupScribbles (GS): An exploratory study of learning effectiveness, ***Computers & Education,*** doi:10.1016/j.compedu.2009.07.003.

Looi, C. K. and W. L. Chen (2010), Community-based Individual Knowledge Construction in the Classroom: A Process-Oriented Account, ***Journal of Computer-Assisted Learning,*** doi: 10.1111/j.1365-2729.2010.00349.

Looi, C. K., W. Chen, and C. Patton (2010), "Principles and enactment of rapid collaborative knowledge building in classrooms", ***Educational Technology Magazine.***

Martin, R. (2003), ***Agile software development: Principles, patterns, and practices,*** Prentice Hall, Upper Saddle River, NJ.

Mayer, R. E. (1989), "Systematic thinking fostered by illustrations in scientific text", ***Journal of Educational Psychology,*** Vol. 81, No. 2, pp. 240-246. doi:10.1037/0022-0663.81.2.240

Means, B. (2006), "Prospects for transforming schools with technology-supported assessment" in R. K. Sawyer (ed.), ***The Cambridge Handbook of the Learning Sciences,*** Cambridge University Press, New York, pp. 505-519.

Minstrell, J. A. (1989), "Teaching science for understanding", In L. B. Resnick and E. Klopfer (Eds.), Toward the Thinking Curriculum: Current Cognitive Research. ***1989 ASCD Yearbook,*** Association for Supervision and Curriculum Development, Alexandria, VA.

Murphy, R., L. Gallagher, A. Krumm, J. Mislevy, and A. Hafter (2014), ***Research on the use of Khan Academy in schools,*** SRI International, Menlo Park, CA.

National Research Council (2011), ***Learning science through computer games and simulations,*** M. A. Honey and M. L. Hilton (eds.), The National Academies Press, Washington, DC.

OECD (2009), ***Giving Knowledge for Free. Open Educational Resources,*** OECD Publishing.

OECD (2010a), ***"Educational Information Systems for Innovation and Improvement: Workshop Summary Report",*** http://www.oecd.org/edu/ceri/46413418.pdf.

© Copyright 2015 SRI International.

OECD (2010b), *The Nature of Learning: Using Research to Inspire Practice,* OECD Publishing.

Pane, J. D., B. A Griffin, D. F. McCaffrey, and R. Karam (2013), Effectiveness of Carnegie Tutor Algebra I at scale. *RAND Working Paper* WR-984-DEIES, RAND Education, Santa Monica, CA.

Pane, J., D. McCaffrey, M. Slaughter, J. Steele, and G. Ikemoto (2010), “An experiment to evaluate the efficacy of Cognitive Tutor Geometry”, *Journal of Research on Educational Effectiveness,* Vol. 3, No. 3, pp. 254-281.

Pape, S. J., K. E. Irving, D. T. Owens, C. K. Boscardin, V. A. Sanalan, A. L. Abrahamson, S. Kaya, H. S. Shin, and D. Silver (2008), *Classroom connectivity in algebra I classrooms: Results of a randomized control trial,* American Education Research Association, New York.

Patton, C. M., D. Tatar, and Y. Dimitriadis (2008), “Trace theory, coordination games, and Group Scribbles” in *International handbook of information technology in primary and secondary education,* Springer US, pp. 921-933.

Pellegrino, J.W. and M.L. Hilton (2012), *Education for life and work: developing transferable knowledge and skills in the 21st century,* National Research Council, Washington, DC.

Penuel, W., Fishman, B., Cheng, B. H., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher,* 40(7), 331-337.

Prieto, L. P., S. Villagrá-Sobrino, I. M. Jorrín-Abellán, A. Martínez-Monés, and Y. Dimitriadis (2011), “Recurrent routines: Analyzing and supporting orchestration in technology-enhanced primary classrooms”, *Computers & Education,* Vol. 57, No. 1, pp. 1214-1227.

Prince, M. (2004), “Does active learning work? A review of the research”, *Journal of Engineering Education,* Vol. 93, pp. 223-232.

Roschelle, J. and S. Hegedus (Eds.) (2013), *The SimCalc vision and contributions: Democratizing access to important mathematics,* Springer, Berlin.

Roschelle, J., J. Knudsen, and S. Hegedus, S. (2009), “From new technological infrastructures to curricular activity systems: Advanced designs for teaching and learning” in M. J. Jacobson and P. Reimann (eds.), *Designs for learning environments of the future: International perspectives from the learning sciences,* Springer, New York.

Roschelle, J., N. Shechtman, D. Tatar, S. Hegedus, B. Hopkins, S. Empson, et al. (2010), “Integration of technology, curriculum, and professional development for advancing middle school mathematics: Three large-scale studies”, *American Educational Research Journal,* Vol. 47, No. 4, pp. 833-878.

Roschelle, J., D. Tatar, S. R. Chaudbury, Y. Dimitriadis, C. Patton, and C. DiGiano (2007), “Ink, improvisation, and interactive engagement: Learning with tablets”, *Computer,* Vol. 40, No. 9, pp. 42-48.

Roschelle, J., B. Penuel, and L. Abrahamson (2004), “The networked classroom”, *Educational Leadership,* Vol. 61, No. 5, pp. 50-54.

Roschelle, J. (1995), “Learning in interactive environments: Prior knowledge and new experience” in J. H. Falk and L. D. Dierking (eds.), Public institutions for personal learning: Establishing a research agenda, *American Association of Museums,* pp. 37-51.

© Copyright 2015 SRI International.

Sandoval, W.A. & Bell, P., (2004). Design-based research methods for studying learning in context: Introduction. *Educational Psychologist*, 39(3), 199-201.

Sawyer, R. K. (Ed.) (2006), *The Cambridge handbook of the learning sciences,* Cambridge University Press, New York.

Scardamalia, M. and C. Bereiter (2006), “Knowledge Building: Theory, Pedagogy and Technology” in R. K. Sawyer (ed.), The Cambridge handbook of the learning sciences, Cambridge University Press, New York.

Scardamalia, M. and C. Bereiter (1994), “Computer support for knowledge-building communities”, The *Journal of the Learning Sciences,* Vol. 3, No. 3, pp. 265-283.

Schmidt, W. H., H. C. Wang, and C. C. McKnight (2005), “Curriculum coherence: an examination of U.S. mathematics and science content standards from an international perspective”, *Journal of Curriculum Studies,* Vol. 37, No. 5, pp. 525-559.

Slavin, R. E. (1990), Cooperative learning: theory, research, and practice, Prentice Hall, Englewood Cliffs, NJ.

Smith, J. P., A. A. diSessa, and J. Roschelle (1993), “Misconceptions reconceived: A constructivist analysis of knowledge in transition”, *Journal of the Learning Sciences,*Vol. 3, No. 2, pp. 115-163.

Strother, S., J. Van Campen, and A. Grunow (2013), *Community college pathways: 2011-2012 descriptive report,* Carnegie Foundation for the Advancement of Teaching, Stanford, CA.

Tatar, D. (2007). The design tensions framework. *Human-Computer Interaction,* 22, (4), 413-451.

Teasley, S. D., and J. Roschelle (1993), “Constructing a joint problem space: The computer as a tool for sharing knowledge” in S. P. Lajoie and S. J. Derry (eds.), *Computers as Cognitive Tools,* Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 229-258.

Trilling, B., and C. Fadel (2009), *21st century skills: Learning for life in our times,* Jossey-Bass.

U.S. Department of Education (2013), *Expanding evidence approaches for learning in a digital world*, Washington, D.C.

U.S. Department of Education (2010), T*ransforming American education: Learning powered by technology. National Education Technology Plan 2010,* Washington, D.C.

U.S. Department of Education, Institute of Education Sciences (2012), *What Works Clearinghouse: Procedures and standards handbook,* Version 2.1, Author, Washington, D.C.

U. S. Government Accounting Office (2010), *Improved dissemination and timely product release would enhance the usefulness of the What Works Clearinghouse,* GAO-10-644, Washington, D.C.

Webb, N. M., and A. S. Palincsar (1996), “Group processes in the classroom” in D. C. Berliner and R. C. Calfee (eds.), *Handbook of educational psychology,* Macmillan, New York, pp. 841-873.

Zucker, A., R. Tinker, C. Staudt, A. Mansfield, and S. Metcalf (2008), “Learning science in grades 3-8 using probeware and computers: Findings from the TEEMSS II Project, *Journal of Science Education and Technology,* Vol. 17, No. 1, pp. 42-48.

© Copyright 2015 SRI International.

SRI Education

SRI Education, a division of SRI International, is tackling the most complex issues in education to identify trends, understand outcomes, and guide policy and practice. We work with federal and state agencies, school districts, foundations, nonprofit organizations, and businesses to provide research-based solutions to challenges posed by rapid social, technological and economic change. SRI International is a nonprofit research institute whose innovations have created new industries, extraordinary marketplace value, and lasting benefits to society.

Silicon Valley
(SRI International Headquarters
333 Ravenswood Avenue
Menlo Park, CA 94025
+1.650.859.2000
education@sri.com

Washington, D.C.
1100 Wilson Boulevard, Suite 2800
Arlington, VA 22209
+1.703.524.2053

www.sri.com/education

SRI International is a registered trademark and SRI Education is a trademark of SRI International. All other trademarks are the property of their respective owners. Copyright 2015 SRI International. All rights reserved. 1/15

STAY CONNECTED