Reconceptualizing a College Science Learning Experience in the New Digital Era: A Review of Literature

Ji Shen, Shiyan Jiang and Ou Lydia Liu

Abstract Despite all the exciting new digital forms of living, our college science education remains relatively unchanged. Students sit quietly in large classrooms listening to lectures (or not), complete individual labs following cookbook instruction, and take exams only to solve problems of no practical importance. It is the time to reconceptualize a college science learning experience for all students. In this chapter, we review research on technology-enriched instruction and assessments for science education at the college level that target students' 21st century skills such as problem solving, critical thinking, and collaboration. We propose three interrelated core principles that can help design coherent science instruction, curriculum, and assessments at the college level that meet the needs of the new digital era: (1) Set the development of lifelong learning skills for students as a top priority; (2) incorporate multi-layered instructional supports using technologies; and (3) design new assessments for individual students that demonstrate and facilitate their growth of the lifelong learning capacity.

Keywords College science education · Learning objectives · Instructional support · Automated scoring

Introduction

One of the most important societal advancements of the 21st century is the rise of information and communication technology (ICT), which has fundamentally transformed our daily lives. We find dining places via smart phones, chat with

J. Shen (✉) · S. Jiang
University of Miami, Coral Gables, USA
e-mail: j.shen@miami.edu

S. Jiang
e-mail: s.jiang@umiami.edu

O. L. Liu
Educational Testing Service, Princeton, USA
e-mail: lliu@ets.org

strangers in a virtual world, and seek information through a large collection of social networks. Despite all the exciting new forms of living, our college science education remains relatively unchanged (Deslauriers, Schelew, & Wieman, 2011; Mazur, 2009). Students sit quietly in large classrooms listening to lectures, complete individual labs following cookbook instruction, and take exams only to solve problems of no practical importance. It is the time to reconceptualize a new college science experience for all students (Mervis, 2013). In this chapter, we propose a guiding framework that can help design coherent science instruction, curriculum, and assessments at the college level that meet the needs of the new digital era. The framework considers three interrelated core principles: (1) Set the development of lifelong learning skills (e.g., critical thinking, scientific reasoning, collaborative problem solving) for all students as a top priority; (2) incorporate multi-layered instructional supports using technologies; and (3) design new assessments for individual students that demonstrate and facilitate their growth of the lifelong learning capacity.

Reseting Learning Objectives

Learning objectives including content standards have been a common topic in any educational reform. This is more so in K-12 public education than in higher education. Many modern ideas on learning objectives can be traced back to Bloom's taxonomy (Bloom et al., 1956), which sets learning objectives for students in three domains: cognitive, affective, and psychomotor. For instance, within the cognitive domain, the learning objectives are placed along a hierarchy that includes (from the lower level to the higher level) knowledge, comprehension, application, analysis, synthesis, and evaluation. The higher-level objectives are often referred as higher-order thinking or higher-level skills.

More recently the 21st century skills have been proposed in various policy documents and reports (e.g., http://www.p21.org/). In order to synthesize the abundant and multifaceted works related to the 21st century skills, the committee of the recent NRC, (2012) report, *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*, identified three broad domains of competence: cognitive, intrapersonal, and interpersonal. These are summarized in Table 1. These skills relevant to science learning may take various forms such as problem solving (Hodges, 2012), scientific reasoning (e.g., Bao et al., 2009), and critical and collaborative argumentation (e.g., Osborne, 2010).

The essence of these new learning objectives, we believe, is to prepare students as adaptive, lifelong learners. Therefore, the first principle we propose to reform college science education is *setting the development of lifelong learning skills for all students as a top priority*. This first principle is particularly relevant for college science education in the 21st century because of the knowledge expansion dilemma. On the one hand, there is a large body of basic (textbook) scientific knowledge

Table 1 The three domains of the 21st century competencies proposed in NRC, (2012)

Domain	Clusters of competencies	Examples
Cognitive domain	Knowledge, Creativity, Cognitive processes and strategies	Critical thinking, information literacy, reasoning and argumentation, and innovation
Intrapersonal domain	Intellectual openness, Work ethic and conscientiousness, Positive core self-evaluation	Flexibility, initiative, appreciation for diversity, and metacognition (the ability to reflect on one's own learning and make adjustments accordingly)
Interpersonal domain	Teamwork and collaboration, Leadership	Communication, collaboration, responsibility, and conflict resolution

distilled through a long human history for students to learn. Without mastering this body of basic knowledge, students can hardly move on to their next level of education or work and eventually (for some of them) contribute to the frontier scientific research and development. On the other hand, new knowledge advances faster than ever. It appears that students are never able to catch up with the knowledge expansion if the focus is on assimilating existing knowledge. Therefore, if students develop lifelong learning skills in schooling, they can continue their own learning after graduation.

These 21st century skills or higher-level learning objectives are often enduring and do not change rapidly. They serve to prepare students for future learning (Bransford & Schwartz, 1999) in the ever-changing world. They should not be decorative additives appearing in course syllabi. Instead, they should be infused in every single activity of the courses students take. These objectives may differ over time because of "society's desire that all students attain levels of mastery—across multiple areas of skill and knowledge—that were previously unnecessary for individual success in education and the workplace" (NRC, 2012, p. 3). One particular new demand of the 21st century is the development of digital literacy (Lei, Shen, & Johnson, 2013). In the next sections, we highlight technological resources for college science education including new forms of assessment that take advantages of technology.

Maximizing Instructional Support Through Technology

Instructional Practices Promoting Lifelong Learning

Froyd (2008) listed eight promising instructional practices in undergraduate STEM education. Based on these, we propose the following four instructional practices that may promote students' lifelong learning skills:

1. *Designing activities to engage and motivate students in active learning*. The essence of this practice is to develop strategies to help students take more ownership and responsibility of their learning through making the classrooms more student-centered environments. These activities range from demonstrating interesting science phenomena, making science content relevant to students' personal lives, extending learning to outside class realms, and linking science to other interesting academic disciplines or even entertainment.
2. *Using scenario-based content organization*. Scenario-based approaches refer to the wide range of instructional practices that organize learning materials for a longer period of time around one or more scenarios. These practices are often labeled as problem-based, project-based, case-based, inquiry-based, or challenge-based learning.
3. *Organizing students in collaborative work*. This practice is combined from the two separate practices proposed by Froyd (2008), organizing students in small groups and organizing students in learning communities. Collaborative work can happen in many forms including within a course or across multiple courses, in or after class, and through face-to-face or virtual interaction.
4. *Conducting research*. This practice aims to involve undergraduate students, typically advanced ones, in science research either in an established lab or under the supervision of a faculty member.

These practices are closely related with each other and have overlaps (Fig. 1). For instance, scenario-based approaches and collaboration are often considered as important ingredients in active student learning environments. Nonetheless, active
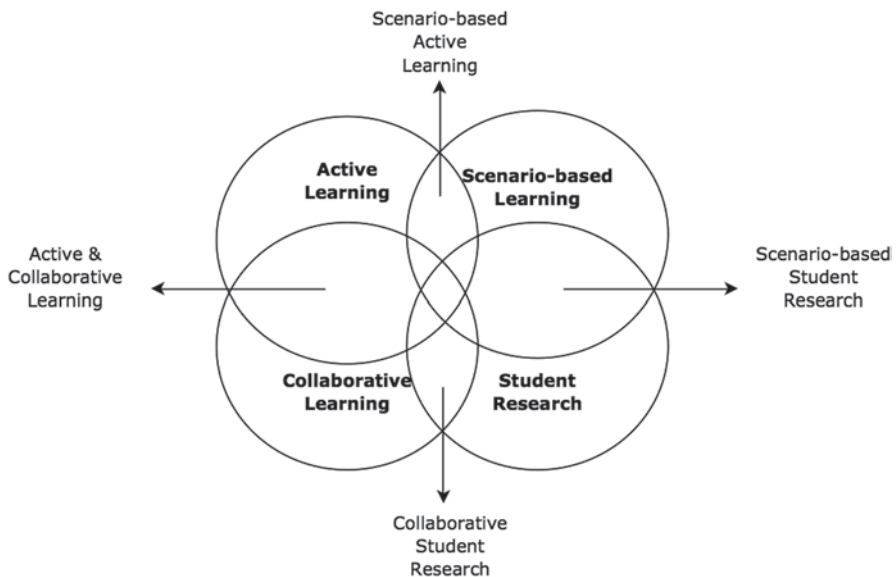

Fig. 1 The interrelated instructional practices that promote lifelong learning

learning can be individual-based and can occur in classes with more traditional ways of content organization. Interested readers can fill out the inner overlapping areas depicted in Fig. 1.

Technological Resources

Advanced technologies have made significant impacts on how students learn and how teachers teach (Lei et al., 2013; NSF Task Force on CyberLearning, 2008). In this section we highlight a few technology resources that can augment the afore-mentioned instructional practices to prepare college students to be lifelong learners in the 21st century.

Personal Response Systems Personal response systems, or clickers, have become a popular tool for large lectures in college science classrooms. The use of clickers is often accompanied with the instructional practice called Peer Instruction (Crouch & Mazur, 2001; Mazur, 1997), a pedagogy developed to engage all students in large classrooms in college science courses. Peer Instruction uses conceptually challenging questions to engage students in scientific reasoning and argumentation. In a Peer Instruction session, students are typically presented with a conceptual question in multiple-choice format. After they spend a minute or two to think about the problem, they use clickers (or other alternatives such as flashcards) to submit their individual answers. The instructor then provides corresponding feedback or follow-up questions based on the distribution of students' responses. For instance, if a large amount of students respond incorrectly, then the instructor can ask the students to discuss the problem with their neighbors (especially one with a different answer). The students then answer the question again before the instructor finally reveals and explains the answer.

Peer Instruction has shown success in improving college students' conceptual learning and problem solving, and retaining students in STEM majors (Kalman, Milner-Bolotin, & Antimirova, 2010; Mazur, 2009; Watkins & Mazur, 2013). Deslauriers et al. (2011) described a comparison study in which they measured the impact of deliberate practice (Ericsson, Krampe, & Tesch-Römer, 1993) in a large-enrollment introductory physics course. In the constructivism-based deliberate practice approach, students solve a series of challenging questions, make and test predictions, and critique their own and peers' arguments during class time that require them to practice physicist-like habit of mind and receive frequent feedback from peers and the instructor. Clickers were used to aid students' problem solving activities during class. Compared with a traditional lecture session taught by an experienced and highly rated instructor, in the 3-h intervention session taught by a trained but inexperienced instructor the students exhibited increased attendance, higher engagement, and much more conceptual learning.

Despite all the documented success, a major constraint of using clickers is that the instructor needs to develop a set of high quality and challenging questions for the students, similar to traditional approaches. This may drive students into thinking

deeply about the subject matter on the one hand, but may inhibit students from developing the essential skill of raising critical questions, an inherent trait of a lifelong learner, on the other hand.

Computer Visualizations and Simulations Computer visualizations and simulations (CVS) including computer-based modeling environments and virtual experiments have become popular instructional tools in science education at all levels (NRC, 2011; Scalise, Timms, Moorjani, Clark, & Holtermann, 2011; Shen, Lei, Chang, & Namdar, 2014). One well-known example is the PhET Interactive Simulations developed at the University of Colorado, Boulder (http://phet.colorado.edu/). PhET simulations include various science (and math) topics covering elementary, secondary, and university levels. They not only visualize abstract and complex scientific phenomena, but also provide opportunities for students to interact with the simulations and therefore, practice inquiry learning (e.g., Lancaster, Moore, Parson, & Perkins, 2013; Wieman, Adams, & Perkins, 2008). For example, Podolefsky, Perkins, and Adams (2010) examined how college students interacted with PhET simulations with minimal explicit guidance. They documented two cases on how students worked with a particular simulation – Wave Interference. Using this simulation, students may choose different objects to show, different measurement tools to use, and different variables to manipulate to make progress towards developing a scientific model of wave interference. Given the flexibility of the PhET simulation, the students followed different exploration paths, similar to how scientists investigate natural phenomena. Another good example is the ChemCollective (www.chemcollective.org) developed at Carnegie Mellon University. It is a collection of online activities including virtual labs, tutorials, and tests for general chemistry instruction. These virtual labs are designed to engage students in authentic chemistry problem-solving and complement algebraic computations for better conceptual understanding. Students' engagement in ChemCollective has been shown to help identify misconceptions, facilitate deeper conceptual understanding, and predict posttest performance (Yaron, Karabinos, Lange, Greeno, & Leinhardt, 2010). Taking a community approach, ChemCollective allows instructors from other institutions to contribute to the development of instructional materials.

With a workforce orientation, Stephens and Richey (2013) cautioned us that it is unlikely that computers and simulations will fully substitute for real world experiences. They observed that the new employees recently hired by the Boeing Company were generally good at using digital tools. However, many of them had rarely been put in situations where they had to create a product of value, and after training, were still weak in skills needed to manipulate materials effectively. Finkelstein et al. (2005) showed that well-designed computer simulations could be used productively in lieu of real laboratory equipment when they were used in proper contexts. The key factor that led to the success of their project was that the circuit simulation they used provided a variety of visual representations to make invisible physics concepts visible to students. de Jong, Linn, and Zacharia, (2013) reviewed the affordances and constraints of physical and virtual laboratories in science and engineering education, and recommended that:

…Combinations of virtual and physical laboratories offer advantages that neither one can fully achieve by itself…. Research on virtual and physical laboratories calls for nuanced decision-making…. Designers of instruction can improve outcomes by taking advantage of the affordances of each type of laboratory…. To design laboratories that take advantage of powerful guidance requires interdisciplinary teams involving domain experts, technologists, and learning scientists. (p. 308)

Computer-Supported Collaborative Learning The works on computer-supported collaborative learning (CSCL) rise with information, network, and Web technologies (for a conceptual review, see Goodyear, Jones, & Thompson, 2014; Stahl, Koschmann, & Suthers, 2006). This instructional approach focuses on developing computer-based learning environments that are built on a deep understanding of social structure, interaction, and dynamics with a relatively broader learning outcome in mind. Frisch, Jackson, and Murray (2013) described the WIKIed Biology course in which they infused Web 2.0 tools (del.i.cious, CiteULike, and Google docs and sites) to help college students collaborate with each other and learn biology knowledge. Using these tools, students worked together to find, create, and disseminate information and knowledge related to the course topics. Results showed that the students increased their understanding of certain biology topics as well as critical thinking skills. In order to understand how students collectively organize information in multiple modes and argue about social scientific issues accordingly, Namdar and Shen (2014) documented a study where they developed a science learning unit on nuclear energy for preservice science teachers. The learning unit incorporated a newly developed knowledge building and sharing platform (ikos. miami.edu) that offers three distinctive types of representational modes: pictorial, textual, and concept maps. The study indicated that the group of learners were able to generate a relatively dense knowledge network. Moreover, concept maps and wiki entries were more connected than the pictorial mode. The findings also suggested that students' knowledge organization and their argumentation practices informed each other in a complex way.

One challenge to incorporate CSCL in college settings is the grading part since in most college classes students are graded individually. How to balance individual accountability and productive collaboration in CSCL still needs more empirical research.

Educational Video and Computer Games Video and computer games have become a popular entertainment means for people of all ages. Gee (2007) asserted that in game playing, players are learning actively and critically to experience the world in a new way and developing resources for future learning. However, evidence for effectiveness of games for science learning is still contested and science learning with games rarely occurs in college settings (NRC, 2011). One major challenge to adopt gaming in college science education is to make game playing really educative and meaningful. A well-known example is Foldit (https://fold.it/portal), an online puzzle video game about protein folding. It takes a citizen science approach that allows users to contribute to actual scientific research related to protein structure and unfolding, which is critical in bioinformatics, molecular biology, and

medicine research. The highest scored solutions submitted by players are analyzed by researchers to evaluate their scientific values in solving real world problems. Notable accomplishments through FoldIt playing include deciphering the crystal structure of the Mason-Pfizer monkey virus retroviral protease (Khatib et al., 2011), and achieving the first crowd-sourced redesign of a protein (Eiben et al., 2012). A similar game is EteRNA (http://eterna.cmu.edu/web/) that enables players to solve puzzles related to the folding of RNA molecules. However, it is still an open question that how these games can be embedded in formal curricula.

OpenCourseWare With the goal of enhancing human learning worldwide through the Internet, OpenCourseWare (OCW) became a popular source for knowledge dissemination for many world's top universities during the first decade of the 21st century. For instance, a well-known physics series is offered by MIT professor Walter Lewin, including Newtonian Mechanics, Electricity & Magnetism, and Vibration and Waves (http://ocw.mit.edu/courses/). Recently, OCW has evolved into Massive Open Online Courses (MOOCs), web-based and large-scale free courses that have no restrictions on enrollment (Adamopoulos, 2013; Balfour, 2013). Overcoming geographic and financial restrictions, a massive number of learners can pursue their individual learning in MOOCs. Popular MOOCs platforms include edX, Coursera, and Udacity.

Hollands and Tirthali (2014) interviewed 83 individuals who were knowledgeable about MOOCs, including administrators, faculty members, researchers and other roles. The authors identified six major goals for MOOCs: (1) extending reach and access (the most stated goal), (2) building and maintaining brand, (3) improving economics, (4) improving educational outcomes, (5) innovation, and (6) research on teaching and learning. The authors suggested that institutions have achieved success to a certain degree regarding these goals except improving economics. Many interviewees agreed that MOOCs can improve educational outcomes. For instance, integrating MOOCs with on-campus courses has shown some signs of success – in this approach students can spent more class time in problem-solving instead of listening to lectures.

A major criticism of MOOCs is that the retention rate is quite low. Only 50–60% of the students enrolled in an MOOC return after the first course and only about 5% earned a credential after completing a course (Koller, Ng, Do, & Chen, 2013). Recent studies have explored students' engagement patterns and associated causes. Since videos are a central element in all MOOCs, Guo, Kim, and Rubin (2009) examined student engagement with videos. They obtained data from 6.9 million video watching sessions from four edX courses: Intro to CS and Programming (MIT, $n = 59{,}126$), Statistics for Public Health (Harvard, $n = 30{,}742$), Artificial Intelligence (Berkeley, $n = 22{,}690$), and Solid State Chemistry (MIT, $n = 15{,}281$). Students' engagement was assessed in terms of how long they watched the video and whether they attempted to answer post-video assessment problems. Video property was measured by the length, type, presentation style, quality, and speaking rate of instructors. The results showed that shorter videos, videos that combine instructors' "talking head" with slides, videos where instructors show their personal feeling,

and videos with Khan-style drawing (see https://www.khanacademy.org/) are more engaging than longer videos, videos only with slides, videos with high-fidelity studio recordings, and videos with still screencasts. Using the same courses, Kim et al. (2014) investigated within-video engagement behaviors. In order to understand the causes that lead to video interaction peaks that indicate points of interest or confusion within the video, the study combined peak profile analysis (log) with visual content analysis (image similarity metric). The results showed that the interaction peaks can be explained by five student activity patterns: starting from the beginning of a new material, returning to missed content, following a tutorial step, replaying a brief segment, and repeating a non-visual explanation.

Connection to Arts Efforts have been made to connect science education with art education because arts practices can promote inspiration and interests. Here we highlight a few examples that take advantage of technologies. A common approach is to develop a course or program that integrates arts and sciences. Jennifer Burg at Wake Forest University initiated a project that aimed to develop curricular materials that integrate mathematics, science, computer science, and digital sound production (http://csweb.cs.wfu.edu/~burg/CCLI/Templates/ home.php). The project brought college-level teachers and students from science and art disciplines to carry out, refine, and disseminate the curricular materials. Sciences and arts can support each other for students to learn science concepts. For example, Bopegedera (2005) conducted a light-related program in which students participated in both art workshops and science labs in order to help students to use scientific understanding of light to create artistic products. In the art workshops students could draw and paint products by hand or using graphing software (e.g., constructing light waves with yarn), while in the science labs students could learn concepts related to light (e.g., the relationships among frequency, wavelength, and the speed of light). Another approach to think about linking arts and sciences is to exploit the power of visualization. A good piece of software that can help practicing scientists to create and animate 3-D molecules is Molecular Flipbook (http://molecularflipbook.org). With the powerful visual aid of molecular graphs, scientists can communicate their findings to others aesthetically and informatively (Atwood & Barbour, 2003). Other creative ways to visualize and disseminate science ideas to the public have also been promoted. For example, Science Magazine hosted a competition named "Dance Your PhD" to encourage college students' using art to communicate scientific ideas and fuel creative thinking. The 2014 Dance Your PhD was awarded to a UGA plant biology student who danced out how forests regenerate after tornado (UGA Today, 2014). Despite these innovative approaches, however, research on connecting arts and science at the college level still needs much empirical work.

Summary In this section, we described a few notable examples of technologies that can be used to promote college students' lifelong learning competencies. We note that a number of important technologies have been left out in this review due to space limit. These may include physical sensors (e.g., Milner-Bolotin & Moll, 2008), virtual or mixed realities (e.g., Cheng & Tsai, 2013), mobile devices

Table 2 Technological resources that can be used to facilitate promising college science education instructional practices

	Designing activities to engage students in active learning	Using scenario-based content organization for complex problem solving and integrated learning	Organizing students in collaborative learning	Involve students in conducting scientific research for independent and critical thinking
Personal response system (i.e., Clickers)	Clicker questions engage students in critical thinking, reflection, and argumentation with peers	N/A	Students discuss and argue with their neighbors about their responses to clicker questions	N/A
Computer visualizations and simulations (CVS), including virtual labs	CVS can draw students' attention and increase students' active interaction	CVS can provide vivid depiction of specific scenarios	CVS renders collective artifacts for collaborative learning	CVS may be involved in research
Computer-supported collaborative learning (CSCL), including online discussion and social networks	CSCL may facilitate the development of a learning community and promote students' active learning	CSCL can be used in a scenario-based approach	CSCL by definition incorporates features to facilitate students' collaborative learning	CSCL can facilitate collaborative research
Educational video & computer games (VCG)	VCG requires users' active participation and learning	VCG is typically built on well designed storylines and scenarios	Multiplayer VCG including MUVE involves collaboration among players; game playing involves game-based learning community	Games can be developed to facilitate scientific research
OpenCourse-Ware (OCW) & MOOCs	OCW can engage motivated students, but still need improvement to engage all students	Exemplar scenario-based OCW is yet to be developed	OCW provides students with online collaboration opportunities	N/A

(e.g., Hwang, Yang, Tsai, & Yang, 2009), and artificial intelligence (e.g., Koedinger & Corbett, 2006), to name a few. We want to echo the position that it is not just the technology but how the technology is being used that matters (Mazur, 2009; Mishra & Koehler, 2006). Each individual instructor needs to consider the available resources and student needs to incorporate these technological resources. Table 2 summarizes the relevant features of these technological resources with respect to the instructional practices that promote life long learning.

Technology-Enriched Assessment for Learning

Without appropriate assessments, a pedagogical innovation will be incomplete (Pellegrino, 2013). Technological advancements have great potential to expand how science assessment can be designed and utilized. In this section, we describe a few assessment approaches that draw heavily on technology to nurture students' lifelong learning capacity.

Embedded Formative Assessment

Formative assessment has been increasingly used in science instruction. Black and Wiliam's seminal paper (1998) emphasized on the various ways that formative assessment can be practiced in classrooms and the ways evidence can be gathered to evaluate the effectiveness of the practices. A characteristic that distinguishes formative assessment from summative assessment is that formative assessment is *for* learning, not *of* learning (Black, 1993). Driven by this distinction, formative assessment offers opportunities for students to understand their misconceptions and improve understanding based on timely feedback.

Although formative assessment has great potential to complement instruction and enhance learning, a few prerequisites need to be satisfied for it to benefit students. First, sufficient professional development needs to be provided to teachers for them to fully understand formative assessment strategies and know when and under what circumstances each strategy should be practiced (Furtak et al., 2008). Second, formative assessment needs to meet quality standards for the assessment to elicit valid information from students. Last, mechanisms need to be developed for teachers to make use of the results from formative assessment. It is not uncommon that assessment results are left sitting on the shelf after a substantial amount of effort has been spent on collecting the results (Ruiz-Primo & Furtak, 2007). After a decade of research on formative assessment, Bennett (2011) provided a comprehensive review of formative assessment, and called for a more critical view of how formative assessment should be implemented and how its effectiveness should be assessed.

Formative assessment can take a variety of forms. For instance, in a college-level medical science course, Riffat, Quadri, Waseem, Mahmud, and Iqbal (2010) practiced a variety of learning and formative assessment tools such as small group discussion, self-direct learning and quizzes. The authors reported improved critical thinking skills and course understanding through the integrated learning and assessment methods. Lancor (2013) described an approach that used student-generated analogies as a formative assessment tool to elicit students' ideas about energy in biology, chemistry, and physics. Computer technology provides an efficient way to embed formative assessment in lesson sequences (Liu, Ryoo, Sato, Svihla, & Linn, 2013). Kibble (2007) reported a program using online quizzes as formative assessment. The study found that the students who participated in the formative quizzes

received higher scores on summative assessments and self-reported that the quizzes were useful in providing quality feedback.

A key component of formative assessment is the mechanism of providing informative feedback for students to improve their learning. For instance, the aforementioned Peer Instruction method (Mazur, 1997; Crouch & Mazur, 2001) is a form of formative assessment. In this method, students receive instant feedback from the automated response distribution of the whole class, from their peers through discussion and argumentation, and from their instructor for clarification and explanation. One constraint of this approach is that students have to attend the class, which they should, to receive the feedback. In contrast, Doige (2012) described an informal, email-based formative assessment program employed to encourage freshmen to constantly revisit their first-year general chemistry materials in a low-stake environment. The students would receive a formative assessment question through email twice a week and, if participating in the program, respond to the question through email. The instructor then would provide timely and personalized feedback to the participating students. The study revealed certain patterns of student participation in this voluntary-based program, and showed that students who participated regularly in the program were more likely to be successful in the summative assessments. One drawback of this approach is that if a large number of students participate in such a program, the responses and feedback for individual students would be extremely time-consuming.

In general, formative assessment should be practiced more frequently in college science classrooms given its potential to provide helpful feedback and improve learning. Formative assessment strategies are particularly needed for large-scale courses including MOOCs as they may be able to help increase student engagement and retention.

Automated Scoring

Automated scoring of constructed-response items is one of the most prominent technologies developed for assessing students' deep understanding (Bennett & Sebrechts, 1996; Dzikovska, Nielsen, & Brew, 2012; Leacock & Chodorow, 2003; Mitchell, Russell, Broomhead, & Aldridge, 2002; Nielsen, Ward, & Martin, 2008; Sandene, Horkay, Bennett, Braswell, & Oranje, 2005). Science educators call for the use of constructed-response items in measuring deep understanding and eliciting reasoning (e.g., Lane, 2004; Shepard, 2000). However, the use of constructed-response items has been limited due to the cumbersome scoring and long turn-around time. Automated scoring, if accurate, can shorten the time between test administration and score report, reduce the number of human raters, and avoid bias typically introduced by human raters (Burstein, Marcu, & Knight, 2003; Liu, Brew, Blackmore, Gerard, Madhok, & Linn, In Press; Williamson, Xi, & Breyer, 2012).

A number of studies have employed automated scoring to score college students' responses to science assessments. Attali, Powers, Freedman, Harrison, and Obetz (2008) applied c-rater®, an automated scoring tool developed by the Educational

Testing Service for content scoring, to score college-level science items in biology and psychology. The responses to the items were typically 1–3 sentences long. The average kappa indicating the agreement between automated and human scores was.62 for biology and.83 for psychology items. Dzikovska et al. (2012) used the content scoring engine BEETTLE II to score college-level physics and the responses were 1–2 sentences long. The kappa value was around.69 for the items tested.

Nehm and colleagues have applied machine-learning techniques to automatically score college students' written responses related to evolutionary biology (Ha, Nehm, Urban-Lurain, & Merrill, 2011; Nehm, Ha, & Mayfield, 2011). Nehm et al. (2011) evaluated the scoring performance of the machine-learning software Summarization Integrated Development Environment (SIDE; http://www.cs.cmu.edu/~cprose/SIDE.html) program against that of human experts, using a corpus of 2,260 student explanations on evolutionary change written by 565 college students. The study found that overall the SIDE software performed very well (i.e., kappa >0.80) and excellent for the natural selection understanding in terms of Key Concept Diversity. Similarly, Ha et al. (2011) applied SIDE to score biology major and nonmajor students' written responses (number of responses >1000) related to evolutionary change in introductory biology courses from two institutions. The results indicated that the automated scoring software did perform well in most cases, accurately evaluating students' understanding of evolutionary change. The authors also identified several common types of students' responses that led to poor performance of computer scoring. These include responses using many key terms but missing important aspects, responses using key terms that are scattered throughout a response, responses using uncommonly used or complex expression, and responses containing spelling and spacing errors.

Going forward, automated scoring has great potential to facilitate immediate feedback to students' written responses to open ended items. In a formative assessment setting, if students can receive instant feedback on their answers to a question and be pointed to relevant instructional steps, learning can be facilitated in a much direct and engaging way. Linn et al. (2014) provided empirical evidence that machine-generated automated feedback is as effective as the feedback provided by an expert teacher in terms of prompting students to revisit instruction and revise answer.

Automated scoring and feedback can be particular helpful for large classrooms including MOOCs in which students are unlikely to receive adequate feedback from the instructor given the mass number of students enrolled in these classes. Automated scoring and feedback offer the possibility for these students to receive meaningful and timely feedback, therefore, increasing their engagement and performance.

Learning Analytics

Since science learning involves complex processes such as inquiry, modeling, argumentation, and collaboration, new forms of assessments need to address the dynamic nature of these processes in order to better capture and facilitate student

learning (Gobert, Sao Pedro, Raziuddin, & Baker, 2013). Learning analytics is an emerging method in educational application that focuses on "developing tools and techniques for capturing, storing, and finding patterns in large amounts of electronic data; representing them in generative and useful ways; and integrating them into intelligent tools that personalize and optimize learning environments" (Martin & Sherin, 2013, p. 12).

There has not been much work conducted in applying learning analytics in college science education. Baker, Hershkovitz, Rossi, Goldstein, and Gowda (2013) presented a supervised method for analyzing student's moment-by-moment learning over time. In the study, participating students used an intelligent tutoring system for college level genetics called Genetics Cognitive Tutor. The researchers then applied a program to create graphs of student moment-by-moment learning. The graph is based on the probability a student knows a concept or skill at a particular time point (the BKT model, Corbett & Anderson, 1995) and learned the concept or skill at a particular step (e.g, a specific step during a problem-solving process; Baker, Goldstein, and Heffernan 2011). The study found that these graphs are correlated with different learning outcomes.

Learning analytics has also been applied in understanding students' engagement patterns in MOOCs. For instance, Kizilcec, Piech and Schneide (2013) proposed a mechanism to identify students' engagement trajectories in MOOCs based on patterns of learners' interaction with video lectures and assessments. Using k-means clustering analysis, they classified learners in three computer science MOOCs into four major patterns: auditing, completing, disengaging, and sampling. Based on learners' self-report, "completing" learners had a significantly better learning experience than the other three groups. They also compared clusters based on learner characteristics and behaviors. They found out two major factors motivated a learner's enrollment: (a) the course is challenging and (b) the learner is interested in the content of the course.

Apparently, more empirical studies need to be conducted in extracting information from learning analytics to facilitate college students' science learning. One possible direction is to utilize these fine-grained data to build more informative digital profiles of learners. In this way, students as well as instructors can better reflect on their learning experience and therefore, take appropriate actions to improve learning as needed.

Conclusion

In this chapter we propose that college science education needs to prioritize the goal of developing students' lifelong learning skills. We reviewed a set of promising pedagogies and new forms of assessments that exploit innovative technologies in college science instruction that can facilitate this goal. We applaud that some of the technology-infused approaches, rare in their kinds, make connections between science and arts instruction at the college level. We stress that it is not about technology per se,

but how it is integrated in instruction in different contexts that matters. To make this happen, we need to engineer creative ways to support faculty in using these innovative methods and technologies. A good example of a University-level imitative is the Science Teaching and Learning Fellows through the Carl Wieman Science Education Initiative at University of British Columbia (http://www.cwsei.ubc.ca). We believe that a large-scale implementation of these new forms of technologies, either through a bottom-up or top-down approach, has potential to bring about transformative changes to reach the goal of college science education in the 21st century.

References

Adamopoulos, P. (2013, December). *What makes a great MOOC? An interdisciplinary analysis of student retention in online courses*. Proceedings of the 34th International Conference on Information Systems (ICIS), Milan, Italy.

Attali, Y., Powers, D., Freedman, M., Harrison, M., & Obetz, S. (2008). *Automated scoring of short-answer open-ended GRE Subject Test items (ETS GRE Board Research Report No. 04-02)*. Princeton: Educational Testing Service.

Atwood, J. L., & Barbour, L. J. (2003). Molecular graphics: From science to art. *Crystal Growth & Design, 3*(1), 3–8.

Balfour, S. P. (2013). Assessing writing in MOOCs: Automated essay scoring and calibrated peer review. *Research & Practice in Assessment, 8*(1), 40–48.

Baker, R. S. J. d., Goldstein, A. B., & Heffernan, N. T. (2011). Detecting learning moment-by-moment. *International Journal of Artificial Intelligence in Education, 21*(1–2), 5–25.

Baker, R. S., Hershkovitz, A., Rossi, L. M., Goldstein, A. B., & Gowda, S. M. (2013). Predicting robust learning with the visual form of the moment-by-moment learning curve. *Journal of the Learning Sciences, 22*(4), 639–666.

Bao, L., Cai, T., Koenig, K., Fang, K., Han, J., Wang, J., et al. (2009). Learning and scientific reasoning. *Science, 323*(5914), 586–587.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice, 18*(1), 5–25. doi:10.1080/0969594X.2010.513678.

Bennett, R. E., & Sebrechts, M. M. (1996). The accuracy of expert-system diagnoses of mathematical problem solutions. *Applied Measurement in Education, 9*(2), 133–150. doi:10.1207/s15324818ame0902_3.

Black, P. (1993). Formative and summative assessment by teachers. *Studies in Science Education, 21*(1), 49–97. doi:10.1080/03057269308560014.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7–74. doi:10.1080/0969595980050102.

Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives: The classification of educational goals, Handbook 1: Cognitive domain*. New York: David McKay.

Bopegedera, A. M. R. P. (2005). The art and science of light: An interdisciplinary teaching and learning experience. Journal of Chemical Education, 82(1), 55.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of research in education* (Vol. 24, pp. 61–199). Washington, D. C.: American Educational Research Association.

Burstein, J., Marcu, D., & Knight, K. (2003). Finding the WRITE stuff: Automatic identification of discourse structure in student essays. *IEEE Intelligent Systems: Special Issue on Advances in Natural Language Processing, 18*(1), 32–39. doi:0.1109/MIS.2003.1179191.

Cheng, K., & Tsai, C. (2013). Affordances of augmented reality in science learning: Suggestions for future research. *Journal of Science Education and Technology, 22,* 449–462. doi:10.1007/s10956-012-9405-9.

Corbett, A. T., & Anderson, J. R. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction, 4,* 253–278.

Crouch, C., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics, 69*(9), 970–977.

de Jong, T., Linn, M. C., & Zacharia, Z. (2013). Physical and virtual laboratories in science and engineering education. *Science, 340*(6130), 305–308.

Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science, 332*(6031), 862–864.

Doige, C. A. (2012). E-mail-based formative assessment: A chronicle of research-inspired practice. *Journal of College Science Teaching, 41*(6), 32–39.

Dzikovska, M. O., Nielsen, R. D., & Brew, C. (2012). *Towards effective tutorial feedback for explanation questions: A dataset and baselines.* In Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (pp. 200–201). Montreal, Canada.

Eiben, C., Siegel, J., Bale, J., Cooper, S., Khatib, F., Shen, B., & Baker, D. (2012). Increased diels-Alderase activity through backbone remodeling guided by Foldit players. *Nature Biotechnology, 30*(2), 190–192. doi:10.1038/nbt.2109.

Ericsson, K. A., Krampe, R. Th., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363–406.

Finkelstein, N. D., Adams, W. K., Keller, C. J., Kohl, P. B., Perkins, K. K., Podolefsky, N. S., et al. (2005). When learning about the real world is better done virtually: A study of substituting computer simulations for laboratory equipment. *Physical Review Special Topics-Physics Education Research, 1,* 010103–010110.

Frisch, J. K., Jackson, P. C., & Murray, M. C. (2013). WikiED: Using Web2.0 tools to teach content and critical thinking. *Journal of College Science Teaching, 43*(1), 71–80.

Froyd, J. (2008, June). *White paper on promising practices in undergraduate STEM education.* Paper presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, D. C.

Furtak, E. M., Ruiz-Primo, M. A., Shemwell, J. T., Ayala, C. C., Brandon, P., Shavelson, R. J., et al. (2008). On the fidelity of implementing embedded formative assessments and its relation to student learning. *Applied Measurement in Education, 21*(4), 360–389. doi:10.1080/08957340802347852.

Gee, J. P. (2007). *What video games have to teach us about learning and literacy* (2nd edn.). New York: Palgrave Macmillan.

Gobert, J. D., Sao Pedro, M., Raziuddin, J., & Baker, R. (2013). From log files to assessment metrics: Measuring students' science inquiry skills using educational data mining. *Journal of the Learning Sciences, 22*(4), 521–563.

Goodyear, P., Jones, C., & Thompson, K. (2014). Computer-supported collaborative learning: Instructional approaches, group processes and educational designs. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (4th ed., pp. 439–451). New York: Springer.

Guo, P. J., Kim, J., & Rubin, R. (2014, March). *How video production affects student engagement: An empirical study of MOOC videos.* In Proceedings of the first ACM conference on Learning@ scale conference (pp. 41–50). New York: ACM Press.

Ha, M., Nehm, R. H., Urban-Lurain, M., & Merrill, J. E. (2011). Applying computerized-scoring models of written biological explanations across courses and colleges: Prospects and limitations. *CBE-Life Sciences Education, 10,* 379–393.

Hodges, K. V. (2012). Solving complex problems. *Science, 338*(6111), 1164–1165.

Hollands, F. M., & Tirthali, D. (2014, April). *MOOCs: Expectations and reality.* Paper presented at the 2014 American Educational Research Association annual conference, Philadelphia, PA.

Hwang, G. J., Yang, T. C., Tsai, C. C., & Yang, S. J. (2009). A context-aware ubiquitous learning environment for conducting complex science experiments. *Computers & Education, 53*(2), 402–413.

Kalman, C. S., Milner-Bolotin, M., & Antimirova, T. (2010). Comparison of the effectiveness of collaborative groups and Peer instruction in a large introductory physics course for science majors. *Canadian Journal of Physics, 88*(5), 325–332.

Khatib, F., DiMaio, F., Foldit Contenders Group, Foldit Void Crushers Group, Cooper, S., Kazmierczyk, M., Gilski M., Krzywda S., Zabranska H., Pichova I., Thompson J., Popović Z., Jaskolski M., Baker D. (2011). Crystal structure of a monomeric retroviral protease solved by protein folding game players. *Nature Structural & Molecular Biology, 18,* 1175–1177. doi:10.1038/nsmb.2119.

Kibble, J. D. (2007). Use of unsupervised online quizzes as formative assessment in medical physiology course: Effects of incentives on student participation and performance. *Advances in Physiology Education, 31,* 253–260.

Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., & Miller, R. C. (2014, March). *Understanding in-video dropouts and interaction peaks inonline lecture videos*. Proceedings of the first ACM conference on Learning@ scale conference (pp. 31–40). ACM.

Kizilcec, R. F., Piech, C., & Schneider, E. (2013, April). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In D. Suthers, K. Verbert, E. Duval, & X. Ochoa (Eds.), *Proceedings of the Third International Conference on learning analytics and knowledge*. New York: ACM.

Koedinger, K., & Corbett, A. (2006). Cognitive tutors: Technology bringing learning science to the classroom. In K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 61–78). Cambridge: Cambridge University Press.

Koller, D., Ng, A., Do, C., & Chen, Z. (2013). Retention and intention in massive open online courses. *EDUCAUSE Review,* May/June, 62–63.

Lancaster, K. V., Moore, E. B., Parson, R., & Perkins, K. (2013). Insights from using PhET's design principles for interactive chemistry simulations. In J. Suits & M. Sanger (Eds.), *Pedagogic roles of animations and simulations in chemistry courses* (pp. 97–126), ACS Symposium Series, 2013. American Chemical Society.

Lancor, R. A. (2013). The many metaphors of energy: Using analogies as a formative assessment tool. *Journal of College Science Teaching, 42*(3), 38–45.

Leacock, C., & Chodorow, M. (2003). C-rater: Automated scoring of short-answer questions. *Computers and the Humanities, 37*(4), 389–405. doi:10.1023/a:1025779619903.

Lane, S. (2004). Validity of high-stakes assessment: Are students engaged in complex thinking? *Educational Measurement: Issues and Practice, 23*(3), 6–14. doi:10.1111/j.1745-3992.2004. tb00160.x.

Lei, J., Shen, J., & Johnson, L. (2013). Digital technologies and assessment in 21st century schooling. In M. P. Mueller, D. J. Tippins, & A. J. Stewart (Eds.), *Assessing schools for generation R (Responsibility): A guide to legislation and school policy in science education* (pp. 185–200). New York: Springer.

Linn, M. C., Gerard, L., Kihyun, R., McElhaney, K., Liu, O. L., & Rafferty, A. N. (2014). Computer-guided inquiry to improve science learning. *Science, 344*(6180), 155–156. doi:10.1126/science.1245980.

Liu, O. L., Ryoo, K., Sato, E., Svihla, V., & Linn, M. C. (2013, April). *Designing assessment to measure cumulative learning of energy topics*. Paper presented at the 2013 Annual Conference of the American Educational Research Association, San Francisco.

Liu, O. L., Brew, C., Blackmore, J., Gerard, L., & Madhok, J. (In Press, early view). Automated scoring for inquiry science assessment: Prospects and obstacles. *Educational Measurement: Issues and Practice*. doi:10.1111/emip.12028.

Martin, T., & Sherin, B. (2013). Learning analytics and computational techniques for detecting and evaluating patterns in learning: An introduction to the special issue. *Journal of the Learning Sciences, 22*(4), 511–520.

Mazur, E. (1997). *Peer instruction: A user's manual*. Upper Saddle River: Prentice Hall.

Mazur, E. (2009). Farewell, lecture? *Science, 323*(5910), 50–51.

Mervis, J. (2013). Transformation is possible if a university really cares. *Science, 340*(6130), 292–296.

Milner-Bolotin, M., & Moll, R. F. (2008). Physics exam problems reconsidered: Using Logger Pro technology to evaluate student understanding of physics. *The Physics Teacher, 46*(8), 494–500.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record, 108*(6), 1017–1054.

Mitchell, T., Russell, T., Broomhead, P., & Aldridge, N. (2002). Towards robust computerized marking of free-text responses. In *Proceedings of the Sixth International Computer Assisted Assessment Conference* (pp. 233–249). Loughborough: Loughborough University.

Namdar, B., & Shen, J. (2014, June). *Knowledge organization with multiple external representations for socioscientific argumentation: A case on nuclear energy*. Proceedings of the 11th International Conference of the Learning Sciences (ICLS), Boulder, CO.

National Research Council. (2011). Learning science through computer games and simulations. Committee on science learning: Computer games, simulations, and education. In M. A. Honey & M. L. Hilton (Eds.), *Board on science education, division of behavioral and social sciences and education*. Washington, D. C.: The National Academies Press.

National Research Council. (2012). Education for life and work: Developing transferable knowledge and skills in the 21st century. Committee on Defining Deeper Learning and 21st Century Skills, J.W. Pellegrino and M.L. Hilton, Editors. Board on Testing and Assessment and Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, D. C.: The National Academies Press.

Nehm, R. H., Ha, M., & Mayfield, E. (2011). Transforming biology assessment with machine learning: Automated scoring of written evolutionary explanations. *Journal of Science Education and Technology, 21*(1), 183–196.

Nielsen, R. D., Ward, W., & Martin. J. H. (2008). *Classification errors in a domain-independent assessment system*. Proceedings of the Third Workshop on Innovative Use of NLP for Building Educational Applications (pp. 10–18), Columbus, OH.

NSF Task Force on CyberLearning. (2008). *Fostering learning in the networked world: The cyberlearning opportunity and challenge*. Washington, D. C.: NSF.

Osborne, J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. *Science, 328*(5977), 463–466.

Pellegrino, J. W. (2013). Proficiency in science: Assessment challenges and opportunities. *Science, 340*(6130), 320–323.

Podolefsky, N. S., Perkins, K. K., & Adams, W. K. (2010). Factors promoting engaged exploration with computer simulations. *Physical Review Special Topics-Physics Education Research, 6,* 020117–020127.

Riffat, S., Quadri, K. H. M., Waseem, A., Mahmud, S. N., & Iqbal, M. (2010). Experience with a theme-based integrated renal module for a second-year MBBS class. *Advances in Physiology Education, 34*(1), 15–19.

Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching, 44*(1), 57–84. doi:10.1002/tea.20163.

Sandene, B., Horkay, N., Bennett, R., Braswell, J., & Oranje, A. (2005). *Online assessment in mathematics and writing: Reports from the NAEP Technology-Based Assessment Project, research and development series (NCES 2005-457)*. Washington, D. C.: U.S. Government Printing Office.

Scalise, K., Timms, M., Moorjani, A., Clark, L., & Holtermann, K. (2011). Student learning in science simulations. Design futures that promote learning gains. *Journal of Research in Science Teaching, 48*(9), 1050–1078.

Shen, J., Lei, J., Chang, H., & Namdar, B. (2014). Technology-enhanced, modeling-based instruction (TMBI) in science education. In J. M. Spector, M. D. Merrill & J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communication and technology* (4th ed., pp. 529–540). New York: Springer.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4–14. doi:10.3102/0013189X029007004.

Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 409–426). Cambridge: Cambridge University Press.

Stephens, R., & Richey, M. (2013). A business view on U.S. education. *Science, 340*(6130), 313–314.

UGA Today (2014). UGA plant biology student wins international 'Dance Your Ph.D.' contest. http://news.uga.edu/releases/article/plant-biology-student-wins-international-dance-your-phd-contest-1114/. Accessed 23 Nov 2014.

Watkins, J., & Mazur, E. (2013). Retaining students in science, technology, engineering, and mathematics (STEM) majors. *Journal of College Science Teaching, 42,* 36–41.

Wieman, C., Adams, W. K., & Perkins, K. K. (2008). PhET: Simulations that enhance learning. *Science, 322*(5902), 682–683.

Williamson, D., Xi, X., & Breyer, J. (2012). A framework for evaluation and use of automated scoring. *Educational Measurement: Issues and Practice, 31*(1), 2–13. doi:10.1111/j.1745-3992.2011.00223.x.

Yaron, D., Karabinos, M., Lange, D., Greeno, J. G., & Leinhardt, G. (2010). The ChemCollective-virtual labs for introductory chemistry courses. *Science, 328*(5978), 584–585.

Ji Shen is Associate Professor in the Department of Teaching and Learning at the University of Miami (UM). Dr. Shen's scholarly work focuses on interdisciplinary science learning and assessment, technology-enhanced learning environments in STEM education, modeling-based teaching and learning, and preservice science teacher education. His work has been funded by the National Science Foundation. Dr. Shen holds a bachelor's degree in Geophysics from Peking University and doctoral degree in Physics from Washington University in St. Louis.

Shiyan Jiang is a third-year doctoral student in the program of STEM (Science, Technology, Engineering and Mathematics) education in the Department of Teaching and Learning at the University of Miami. Her research interests include visualizations of learning process, technology-enhanced learning, emerging technologies for teaching and learning, Computer Supported Collaborative Learning (CSCL), and instructional design. Ms. Jiang holds a bachelor's degree in educational technology from East China Normal University.

Ou Lydia Liu is Managing Senior Research Scientist at the Educational Testing Service (ETS). Her main research areas include assessment of student learning outcomes in higher education and innovative science assessment in K-12. She is currently leading a NSF funded project and Co-PI for three other NSF funded projects on science assessment. Dr. Liu served on the advisory panel for the OECD's AHELO (Assessment of Higher Education Learning Outcomes) project. Dr. Liu has published widely in applied measurement, higher education, and science education journals. She was the recipient of the NCME Jason Millman Promising Measurement Scholar Award in 2011. Dr. Liu holds a bachelor's degree in Science & English from the University of Science and Technology of China and a doctoral degree in Quantitative Methods and Evaluation from the University of California, Berkeley.