CONNECT:
Making Learning Personal

Reports from the Field by the League of Innovators

Information Technologies to Advance Teaching and Learning

Ryan Baker and Janet S. Twyman

The **Center on Innovations in Learning** developed its "Conversations with Innovators" event as a forum for its **League of Innovators** to engage in intimate discussions with author/experts on selected topics. The 2016 event was held at Temple University on June 22nd and 23rd. In each of three sessions, pairs of experts—each of whom had written a chapter for the Center's recent publication, ***Handbook on Personalized Learning for States, Districts, and Schools***—made brief 5–7 minute presentations on the designated topic, after which the floor was then opened for participants' questions and discussion. The lively oral discussion was enhanced by participants' postings on Padlet, an online virtual bulletin board. In three issues of ***Connect,*** the conversation from each session continues, with author/experts responding to the overflow of questions and comments. In Session 1, Joe Layng and Sam Redding discussed **Personal Competencies as Propellants of All Learning**; and in Session 2, Melinda S. Sota and Karen Mahon addressed the topic of **Personalizing Instruction: Student Voice and Choice.**

In this issue of *Connect*, Ryan Baker and Janet S. Twyman respond to questions related to the topic of Session 3, **Information Technologies to Advance Teaching and Learning**. Dr. Baker, who wrote the *Handbook* chapter "**Using Learning Analytics in Personalized Learning**," is an associate professor of cognitive studies at Teachers College, Columbia University. He is currently serving as the founding president of the **International Educational Data Mining Society** and as associate editor of the ***Journal of Educational Data Mining***. One of his many publications, "**Stupid Tutoring Systems, Intelligent Humans**," was just published in the June 2016 issue of the ***International Journal of Artificial Intelligence in Education***. His research combines educational data mining and quantitative field observation methods in order to better understand how students respond to educational software, and how these responses impact their learning.

Dr. Twyman, who wrote the chapter "**Personalizing Learning Through Precision Measurement**" for the ***Handbook***, is an associate professor of pediatrics at University of Massachusetts Medical School and the director of innovation and technology for the **Center on Innovations in Learning**. For over a decade, she has worked at the forefront of merging evidence-based educational methods with new and emerging technologies, including selecting technologies that incorporate adaptive instructional systems to support personalized learning. As a vice president at

This field report is the ninth in a series produced by the Center on Innovations in Learning's League of Innovators. The series describes, discusses, and analyzes policies and practices that enable personalization in education. Issues of the series will present either issue briefs or, like this one, field reports on lessons learned by practitioners recounting the successes and obstacles to success encountered in implementing personalized learning.

Neither the issue briefs nor the field reports attempt to present in-depth reviews of the research; for those resources readers are encouraged to access the Center on Innovations in Learning's resource database. Topics should be of particular interest to state education agencies and district and school personnel.

Headsprout, she led the design, development, and dissemination of the company's Internet-based reading programs and oversaw their implementation in over 1,500 public and private schools.

Below are the questions asked by attendees, followed by the authors' responses.

1. How is the topic of your session (Information Technologies to Advance Teaching and Learning) connected to evidence-based practices?

J.T.: Evidence-based practice involves making pedagogical decisions informed by relevant empirical research evidence. "Information technologies"—specifically digital tools that can track, measure, and analyze learninga—are valuable assets in making data-based decisions and using evidence to guide practice. Such tools and technologies make actual evidence-based practice easier to do, and perhaps more reliable and more informative both for immediate decisions and for designing and delivering instruction over time.

R.B.: I completely agree. And beyond this, modern information technologies have supported major advances in research on the science of learning and teaching, which have made significant additions to our base of evidence that can be used to support practice.

2. Twyman (2016) states that "true personalized learning varies the time, place, path, pace, practice, and trace of learning for each and every student." Can you help to clarify the difference between trace and path?

J.T.: My variation on "varying time, place, and pace" adds three other dimensions, path, practice, and trace. These additional dimensions add to the bigger picture of how we view student learning. "Path" refers to the route a student takes to move towards his or her learning objectives, whether it be a specific course sequence towards graduation or the series of steps a student took to meet a specific competency. A personalized learning system should offer many pathways, including opportunities that are created by the student. "Practice" is what the educators do to facilitate learning—in other words, the actual teaching (or, on a larger scale, the implementation of policies at the school, district, or state level). In education we see a plethora of strategies and interventions, but little specific, reliable guidance on what practices to use, when, with whom, under what conditions. Whenever possible, educators should always avail themselves with evidence-based strategies and tactics, varying their practice based on the needs, interests, performance, and goals of each of their students. The effects of practice lead us to "Trace," or what remains after teaching and learning occurs. How do we know when a student has learned something, or more importantly for practice, if a student is learning? Trace is the objective, notable change that comes from teaching and learning. It can be a change in things like knowledge or ability, and can be seen in many ways such as performing new skills, increased accuracy to questions, greater confidence, greater fluency, or in data displays. Good personalized learning embraces the multitude of ways the trace of learning can be seen.

3. When discussing personalization, one of the first "pushbacks" we hear is that unmotivated or unprepared students will make the choice not to learn, or will take a significant amount of time to reach expectations. How can the affect EDM/LA [Educational Data Mining/Learning Analytics] assist with identifying and preventing this type of choice?

J.T.: I think Fred Keller, in his famous 1968 article, "Good-Bye, Teacher...," succinctly debunked this argument; he wrote, "I learned one very important thing: the student is always right. He is not asleep, not unmotivated, not sick, and he can learn a great deal if we provide the right contingencies of reinforcement. But if we don't provide them, and provide them soon, he too may be inspired to say, 'Good-bye!' to formal education." [*Journal of Applied Behavior Analysis*, *1*(1), p. 88.]

R.B.: I think that the issue of students taking a significant time to reach expectations and the issue of students choosing not to learn are partially separable. Starting with the first issue, one of the first core themes in EDM/LA was the automated detection of disengaged behaviors such as gaming the system, off-task behavior, and carelessness. See, for example, the 2004 article by R. S. Baker, A. T. Corbett, and K. R. Koedinger, "Detecting student misuse of intelligent tutoring systems" in the *Proceedings of the 7th International Conference on Intelligent Tutoring Systems* (pp. 531–540, Maceió, Alagoas, Brazil). Although automated interventions based on these kinds of detectors have not yet made it into commercial products used at scale, this type of technology creates the long-term potential to support disengaged students. Perhaps more importantly, these types of detectors have been used to understand why students choose not to learn; it's connected to a lack of grit and academic persistence (see Angela Duckworth's work) and to boredom. Which is ultimately good news, because various interventions have been found to increase academic persistence (see David S. Yeager's work) and boredom is relatively feasible to reduce as well—not easy, but definitely possible. Boredom varies a lot between different online platforms. Additionally, specific platforms such as ASSISTments and Reasoning Mind have used evidence on disengaged behavior and boredom to redesign and enhance their platforms—see for example the 2015 article "Incorporating effective e-learning principles to improve student engagement in middle-school mathematics," by K. Mulqueeny, V. Kostyuk, R. S. Baker, and J. Ocumpaugh [*International Journal of STEM Education*, 2(15)].

Dr. Baker discusses how disengagement is one of the reasons why students flounder and take a long time to reach expectations.

To answer the other half of this question: Disengagement is one of the reasons why students flounder and take a long time to reach expectations. But it's only part of the story. If a student is far behind where he or she should be, a poorly-designed platform won't give that student the needed support. Getting 140 opportunities to learn Algebra equation-solving won't work very well, if the student repeatedly obtains wrong answers due to failures with arithmetic. In these cases, teachers need to be aware of what's going on, and intervene. Alternatively, an effective platform might still take a *looooong* time if the student isn't prepared. That's a more systemic failing—not in the platform but in the student's prior educational experiences—that may be out of the control of the platform, the current teacher, and even the current school district if the student transferred in from a failing school.

4. Are (measures of) learning outcomes at the unit level, or can they be extended over several months, or even years?

J.T.: Good (as in meaningful, useful) measures of learning occur at multiple levels, both slicing and dicing the behavior (the knowledge or ability), and extending across space and time. When we look at slices of learning, we take into consideration what that "learning" is made up of, what are the prerequisite skills, what does a task analysis tell us about what needs to be taught, what are the components of this composite repertoire we are trying to establish, and what do these indirect measures of learning tell us about what is going on? These various levels of data and measures inform teaching and learning, what the next steps may be, where any holes are that may need to be filled in. They help our instruction be more efficacious, and more personalized. When we extend measures across time and space, we are looking at maintenance of learning over time and in different contexts, settings, and situations. All of these aspects are useful when building, engaging in, and evaluating learning environments.

R.B.: Emphatically agreed. In fact, researchers in the learning sciences are increasingly dubious of short-term measures of learning. The pedagogical strategies which support immediate high performance—massed practice, say, or memorization of templates—often lead to much poorer learning outcomes over the long term than more conceptual and spread-out approaches.

Kinds of Data

5. In the Baker chapter, what does "gaming the system" mean? Is there a teacher/student analog? How does knowing this about students help teachers (not just the applications) personalize how they work with students going forward?

R.B.: Gaming the system is typically defined as attempting to succeed in an educational environment through exploiting properties of the system or platform rather than through learning the material. On this topic, I refer you back to the 2004 article I mentioned in response to a question above, but also see the 2006 article by R. S. J. d. Baker, A. T. Corbett, K. R. Koedinger, S. Evenson, I. Roll, A. Z. Wagner, et al., "Adapting to when students game an intelligent tutoring system," in the *Proceedings of the 8th International Conference on Intelligent Tutoring Systems* (pp. 392–401; Jhongli, Taiwan.) Typical forms of gaming the system in online learning include systematically guessing (1, 2, 3, 4, 5…) or asking for a hint over and over until the system gives up and provides the answer. Another recently documented form is creating multiple user accounts, with one account to harvest answers and another account to enter them correctly. There is a "live classroom" analogue in "executive help-seeking"—documented by Sharon Nelson-Le Gall and also by Amy Arbreton—where a student asks the teacher, teacher's aide, or other students for help and answers over and over. Whether this kind of behavior occurs online or offline, it's an opportunity for teachers to discuss the purpose of their learning with students, and explain how much gaming the system is a problem for their long-term success.

6. What kinds of student data can we obtain? Is it simple corrects/errors and the like, or can we measure things like "deeper learning"? For example: Mindset vs. Grit: How are these measured?

J.T.: I often use the iceberg analogy when addressing this question. The amount and type of data that we can collect and use is enormous; however, what we often get is a small bit, easily seen just above the surface. These are often measures such as number or percent correct, log-on frequency or duration, number of sessions, time in program, etc. Just below the surface, and incorporated usefully in some programs, are measures such as error patterns, learning sequence effects, fluency, response latency, and other dimensions that can be used to infer information about learning. While these are useful and informative metrics, other sorts of "deeper" data often are not collected or analyzed (notable exceptions are things like

Dr. Baker's work in using data to look at student off-task behavior, learner affect, or cognitive-affective states). These things are harder because they often are not directly measured, and must be inferred by a combination of other measures. As more of this type of research and work is done, the variables for "deeper learning" labels such as "grit" or "mindset" will become better understood. For more on measuring personal competencies and personalized learning Dr. Crean-Davis' chapter "**Proceed With Caution: Measuring That 'Something Other' in Students**" in the CIL ***Handbook*** is highly recommended.

Dr. Twyman discusses measuring things like "deeper learning" and Mind vs. Grit.

R.B.: Great points, Dr. Twyman. I would also add that Valerie J. Shute at Florida State University has made great strides to automatically measure grit and persistence in games, from learners' behaviors.

7. The Twyman chapter talks about how data can aid students and teachers in making instructional choices. Why is it that sometimes the act of using data with students is couched as impersonal in the education sector? How do we educate the educators that use of frequent, pinpointed data actually humanizes (rather than dehumanizes) the educational experience?

J.T.: It confounds and surprises me as well that data (in society, not just in education) seems to have a bad reputation: that of being cold and unfeeling. Peter DeWitt addresses some of the factors in his blog, **Are Educators Data Driven to Death**. As I wrote in my ***Handbook*** chapter, "**Personalizing Learning Through Precision Measurement,**" "Measurement has been accused of being a woefully inadequate means of getting at what really matters in education. However,...measurement is essential to any earnest teaching (or learning) effort. Without it we cannot truly or well personalize instruction for any student, let alone for all students. When done well and for the right reasons, measurement is one of the most caring and beneficial acts teachers can do" (p. 149).

8. In understanding data, what is the difference between correlational and causal relationships?

R.B.: All causal relationships are correlational, but not all correlational relationships are causal. A causal relationship, A → B, establishes that A actually causes B. For example, smoking cigarettes is now thought to *cause* lung cancer. Causality is only reliably demonstrated through true experiments, where you manipulate one variable and study the impact on another variable (other methods, such as quasi-experiments with propensity score matching, provide evidence about causality, but with lower confidence). By contrast, a correlational relationship between A and B may indicate that A causes B, that B causes A, or that some other factor C causes both A and B. Correlational evidence is still useful—as a predictor, to generate hypotheses for true experiments, or when a true experiment is simply impossible.

Data Competency for Educators and Students: Making Sense of Data AND Acting Upon It

9. How do we develop expertise (of educators) in evaluating evidence of what is effective for learning? Will teachers need to now become data scientists? How do we cultivate the field so

learning analytics are not only embraced, but also used wisely? What are the practices that allow teachers to use data well? (Redding: Good teacher use of data is to change practice). Similarly, how do we develop *student* data competency (i.e., teach students to look at their own learning data, to make sense of it, and to act upon it)?

J.T.: As noted in the header, data competency for educators and students is in making sense of useful data and acting intelligently upon it. Many of the problems characterized by a "lack" of data competency in educators lies in what sorts of data are collected, how and when they are collected, how they are accessed (i.e., not easily), when they are accessed (often too late), and how they are used (e.g., punitively). These conditions need to change and educator data competency can and will increase. Also, I believe it's imperative that, at an early age, students develop competencies around not only understanding their own data, but also in how to collect and act upon it. This gives them greater insight into their own learning and change that they can draw upon outside of school and throughout life. By asking questions and learning to chart information, even very young children can understand how to find and communicate information using data visualizations. (For example, see the 2003 article "**Collecting, Representing, and Interpreting Data: A Lesson for Kindergarteners and First and Second Graders,**" by L. Dacey and R. Eston in *Math Solutions Online Newsletter*, Issue No. 9). People (students, teachers, administrators, parents) begin to understand data when it conveys relevant information in easy-to-digest ways. When helping people begin to analyze and interpret data, it is often helpful to think of the data telling a story—what story are these data trying to tell. What questions are they answering; what are they telling us that help guide our actions, now or in the future? The website **Flowing Data** is one source of examples of how visual displays of data (data visualization) can be used to convey meaning or tell a story.

R.B.: I concur with Dr. Twyman. And let me add that competence with data does not imply needing to become a full data scientist, any more than one is required to be a professional mathematician to teach or use algebra. We need greater competence with data and data reasoning in our society, but we also need to create tools that facilitate data usage. A little bit of data literacy will go a long way in using learning analytics effectively and appropriately.

10. When a school leader follows the action principle (from Ryan Baker's chapter) to ask for raw data and validation studies (from vendors or creators of educational programs), what prerequisite knowledge will they need to understand the responses provided? Karen Mahon suggests asking for data from vendors that will verify ongoing benefit of the product (instead of the current practice of ROI [return on investment] defined by usage patterns).

R.B.: That's a hard question to answer. I don't think it's realistic that every school leader will become a data scientist (see answer above). But it is realistic that every large district and definitely every region bring on board one person trained in evaluating this kind of evidence —a data scientist or learning analytics specialist, at the master's degree level. Even taking (and really mastering) two MOOCs on this subject (one on learning analytics and one on statistics for program evaluation) would probably be enough to make thoughtful and critical decisions about the evidence provided by vendors. While society builds expertise in this more generally, schools and districts will have to talk to each other and share information and work together.

11. What is the role or connection of the teacher with those developing systems that collect data (i.e., vendors)? How do we get teacher's ideas and needs into the process of defining what is

needed from the data?

J.T.: I do see building the measurement and analytics into educational software as a two-way—or really a three-way—street (and four-way if we can get the students in there too!). By this I mean the designers and creators of the educational software, hands-on educators and administrators, and learning scientists/instructional designers (yes, and students too) should all influence the development of the level and type of data collected, how it is displayed and interpreted, and communicated. Both the **U.S. Department of Education's Office of Educational Technology** and organizations like **Digital Promise** have made efforts to connect educators and software developers, resulting in publications like the **National Education Technology Plan** or the **Ed Tech Developer's Guide**.

R.B.: Agreed. Vendors who *don't* work with teachers, or at least very recent former teachers, are making a huge mistake. This is one of the issues, by the way, that school leaders can also ask about, and should be well-prepared to assess whether they are receiving a decent answer.

Student Data Mining

12. Need guidelines for use of data mining in education: California has distinct laws relating to student data mining, that may not be an ideal. Wisconsin and Dallas are developing sophisticated approaches to student data mining.

R.B.: I'm not sure what the question is here? But ultimately, the key is to develop legislation and policy that enable research and development of high-quality solutions—and validation of their effectiveness—while protecting privacy. Complete bans on retention of personally identifiable information (PII)—or requirements to delete all data at the end of every semester—seem like they protect students, but they considerably reduce the potential to improve the quality of the systems students use to learn. Instead, PII mappings should be held in trust by school districts or by other entities, under careful control, but used when there is a good justification for benefit and careful controls of privacy.

KEY STATEMENTS

13. Any final takeaways?

J.T.: It's important to make educational data meaningful and actionable.

R.B.: Data quality and data analytics are not always opposed to each other; they can go hand in hand.

www.centeril.org

The **League of Innovators**, a network of state education agency and Regional Comprehensive Center personnel with an interest in learning innovations, is organized and administered by the **Center on Innovations in Learning**.

The **Center on Innovations in Learning** (CIL) is a national content center established to work with regional comprehensive centers and state education agencies (SEA) to build SEAs' capacity to stimulate, select, implement, and scale up innovations in learning. In partnership with the Academic Development Institute (ADI), Lincoln, Illinois, the Center on Innovations in Learning is located at Temple University in Philadelphia, Pennsylvania. The Center is funded by the U.S. Department of Education, Office of Elementary and Secondary Education (OESE), under the comprehensive centers program, Award # S283B120052-12A. The opinions expressed herein do not necessarily reflect the position of the supporting agencies, and no official endorsement should be inferred.