Technology in Support of Diverse Assessment

Gary L. Ackerman

Windsor (Vermont) School

Abstract

This paper presents the case of a small rural school that implemented an assessment system designed around three different measures, each designed to answer a different question about the secondary students enrolled in the school. Technological solutions were used to captured both qualitative and quantitative evidence of students' learning; those solutions are described and other findings are discussed.

Keywords: assessment, K-12, portfolios, educational technology

Technology in Support of Diverse Assessment

Compared to previous generations, today's educators gather more and more types of information regarding students' learning and they use it to make more decisions (Rajagopalan & Gordon, 2016). Data gathered during instruction are used formative purposes and data gathered in summative assessments are used for evaluation (Harlen & James, 1997). Both types of assessments are used to improve curriculum, to evaluate programs, to justify accreditation decisions, and for other institutional purposes (Walvoord, 2010). Despite the long tradition of relying on the results of tests to measure students' learning, there is increasing recognition that diverse assessment data are necessary to completely describe students' learning, understand the value of instruction and curriculum, and identify aspects of schooling that need to be improved (Andrews & Wulfeck, 2014; Dede, 2010).

As data and its purposes have become more sophisticated, information and computer technology (ICT) has been applied to many assessment problems, and a range of ICT tools have been deployed to facilitate assessment. Nearing the third decade of the 21st century, educational practitioners use ICT to collect easily quantifiable data such as scores on tests that are administered given to large populations (Bennett, 2002) and content questions that can graded by computers in online and blended classes (Black, Beck, Dawson, Jinks, & DiPietro, 2007). ICT has also been used to document and assess learning in classrooms where teachers seek to promote and document complex learning (Webb & Gibson, 2015). Further, ICT tools are used to manage vast collections of assessment data as well as to manipulate, analyze, and report those data.

In this paper, the author presents the case of a public school in which ICT was adopted and adapted to facilitate several assessment practices. There are two purposes of this paper. First, to describe the assessment system and the roles of technology in the system. Second, to articulate and support conclusions regarding the design of effective technology systems in secondary schools.

The Setting and the Data Collection

The project described in this paper occurred in a comprehensive public school in a rural area of the northeast United States; it enrolls approximately 300 students in grades 7-12. The guidance counselor at the school described it as "similar to other schools in the state. We have slightly more students in poverty and receiving special education services than others, but we are closer to the middle than the extreme," and the school "sends its share of students to Ivy League schools, but we also have many students who go to work right out of school."

In the summer 2015, several groups of educators in the school were working on different initiatives; those groups included a committee working to increase the use of project-based learning (Krajcik & Shin, 2014) in all courses; a dyad preparing reports for an upcoming accreditation review; and a group of four working to upgrade the information technology hardware, software, and web services for instruction. As the principal was preparing for the beginning of the school year, he convened a meeting of individuals representing those groups for the purpose of "designing assessments that bring together the curriculum, technology, and accreditation work so it all fits together." In introducing the project to the group, the principal noted, "we have our students' SBAC, PSAT, SAT, and other test scores, but our students do so much more. By capturing a more complete view of our students, we can improve our curriculum and reporting to the accreditors." The members of this group agreed to assume a leadership role in defining a new assessment system. For the purposes of this paper, this group is called the Consolidated Committee. The author was involved as an internal consultant to the Consolidated Committee. He was primarily involved in designing and configuring the technology solutions, and revising them as they were deployed. He was present at most meetings of the Consolidated Committee to provide input specific to the technology infrastructure and to understand what actions the committee recommended so that the technology accurately reflected the intent.

This case is described with data from three sources. First, the agendas used to plan and the minutes used to record the meetings of the Consolidated Committee were reviewed. The primary purpose of this data was to accurately describe the work including design decisions and the rationale that supported each decision. Second, the author conducted a focus group interview (Fey & Fontanta, 1991) of the Consolidated Committee. Finally, the author conducted a second focus group interview with a group of teachers who had contributed to and made use of the assessment system. Both focus group interviews were conducted about 18 months after the project began and at a time when the systems were fully functional and active development had stopped (although the Consolidate Committee reconvened soon after the interviews were completed to continue development of the system).

The transcripts of the focus group interviews were coded using the constant comparative method (Glaser, 1965). The author along with a second researcher read and coded the transcripts of the interviews; themes that were identified by both researchers are elucidated in the findings.

The Assessment System

The Consolidated Committee adopted a planning method grounded in educational design research (McKenny & Reeves, 2012); this method is intended to facilitate the use of theory to create interventions that improve educational practice, and it comprises three phases (see Figure 1). McKenny and Reeves (2014) described educational design research as "the iterative development of solutions… to practical and complex educational problems" which leads to "new knowledge that can inform the work of others" (p. 133). Key aspects of educational design research include an iterative approach to understanding the local instantiation of a problem and continuous planning actions that reconcile theory with the local circumstances. Further, educational design researchers seek to

Figure 1. Educational design research adapted from McKenny & Reeves (2012)

both evaluate the degree to which the interventions produced the desired results and they seek to articulate and support generalizations that can inform similar efforts in different settings.

The principal described the vision for the assessment system that emerged during conversations with the faculty over the course of several months prior to the formation of the Consolidated Committee. These became the analysis/ exploration phase of the process. The principal observed,

> We seem to agree that a rich and varied curriculum leads to good test scores, but we know those scores do not show the true extent of our students' work. Plus, our local learning expectations [upon which] we spend so much time on don't show up in any tests. Everything students do is important to us, so we need to capture the data in whatever way [the students] exhibit their learning.

In this vision, the principal and his faculty sought to broaden their definition of assessment data to more closely align with high-quality and valid assessments (Pellegrino, 2014) than is possible relying on standard test scores, which "had been the focus of assessment efforts previously." The Consolidated Committee recognized the need for diverse data that documents creativity, problem-solving, communication, and collaboration to be included in the formal assessment system, as those skills have taken on new importance in organizations and cultures in which digital electronic information technology dominates (Black & William, 1998; Cumming & Maxwell, 1999; Cumming, Maxwell, & Wyatt-Smith, 2016; van Laar, van Deursen, van Dijk, & de Haan, 2017).

Dimensions of Learning

Three sources of data became the focus of the assessment system (see figure 2), and IT systems were designed to collect and use all three types of data. While the Consolidated Committee used the terms "Course Grades," "Standard Tests," and "Performances" to identify the types of information they were collecting, they suggested the assessment system was designed to answer three questions: a) Does the student have the habits of effective learners and workers? b) How does each student compare to students in other populations? c) Can the student create polished work that is valued outside the context of school? and the three sources of evidence in the assessment were interpreted as answering those questions.

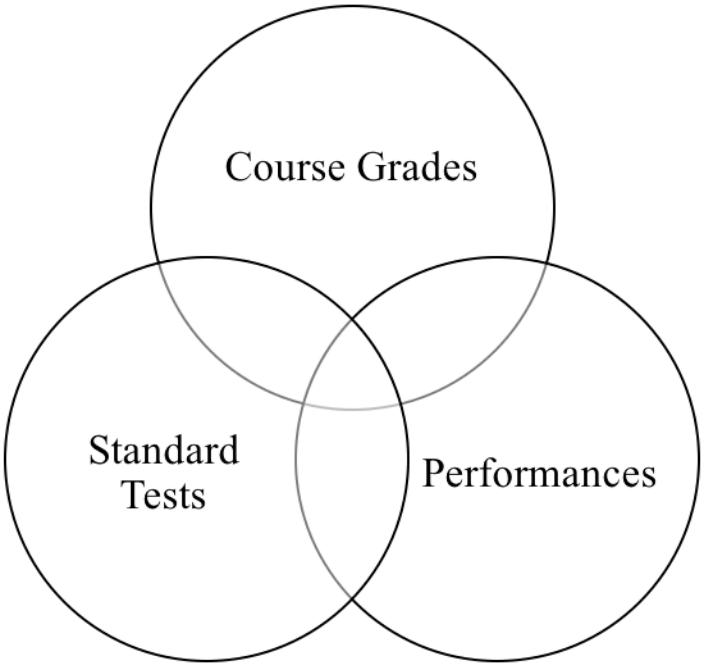

Figure 2. The three sources of evidence in the assessment system

The students in the school participate in a range of standards-based and standardized tests. These include tests associated with college admissions, as well as the tests associated with the Common Core State Standards (CCSS) (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Because he led a public school, the principal was obligated to ensure these test scores were shared with the public, so it was decided the school would continue its participation in the tests. Students would be encouraged to fully engage in the tests and the scores would be presented to students, parents, and the community as "one small part of

what we know students can do, and the only chance to see how they compare to students in the wider population." Following this rationale, students' scores on these tests became part of the assessment system.

While many schools appear to be abandoning traditional course grades in favor of reporting on students' progress towards standards (Stump & Silvernail, 2014), the Consolidated Committee reasoned learning in that dimension was already being documented in the Smarter Balanced Assessment Consortium (SBAC) tests administered to comply with CCSS, thus a standards-based course grade was redundant. Traditional course grades, on the other hand, are positively associated with several aspects of each student's skills and knowledge; including content knowledge, ability to complete tasks in a timely manner, compliance, and attitude and participation which appear to be necessary for many workplace settings (Manpower Report, 2015). The committee recorded agreement with the teacher who suggested, "these are all things that matter when students leave the school, no matter where they go, shouldn't we somehow record their ability to succeed in similar ways in school?" They further reasoned the subjective assessment of students by different professionals in different courses was a predictor of students' success, thus they decided to continue collecting traditional course grades.

The final dimension of learning that needed to be captured in the assessment system was students' participation in project-based learning (Krajcik & Shin, 2014) and other authentic learning environments (Herrington, Reeves, & Oliver, 2006) in which the products they created resembled the work done by professionals working in their field of study. The Consolidated Committee recommended an electronic portfolio platform with which students could collect and cull artifacts (mostly images and video) of the projects they created along with their reflections on the importance of that work be included to complete the assessment system

Technologies for Assessment

Once the three dimensions of learning had been defined, the Consolidated Committee turned its attention to designing ICT systems that would allow teachers and others to easily record data documenting learning in each dimension and that would allow school leaders, teachers, and others to manipulate and present the data that documented each dimension of learning. Specifically, the committee decided to adopt three systems.

- In the summer when the Consolidated Committee began designing the new assessment system, the student information system (SIS) was upgraded as part of a long-established technology plan. The Consolidated Committee also recognized the school had been using the SIS for several years and that "teachers, parents, and

the administration along with the technology staff were familiar with and satisfied with the method of recording course grades," so that system would remain in place and be used to document course grades.

- A spreadsheet file would be created to track students' scores on standards-based and standardized tests that are administered to high school populations throughout the state, region, and country; this was known as the "Dashboard Spreadsheet" following the convention of displaying quantitative assessment data on dashboards.
- A new electronic portfolio system to capture artifacts of students' performances and products on authentic projects was to be installed and configured.

The first design of the Dashboard Spreadsheet compiled students' SBAC, SAT, PSAT scores. Upon reviewing that draft, the Consolidated Committee concluded there was too little data, so it was immediately redesigned; this is consistent with the design/ construction phase of educational design research (McKenny & Reeves, 2012). Changes included adding grade point averages and categorical data such as foreign language courses passed, participation in athletics, arts, and other activities. Included in this first design/ construction iteration was defining procedures for exporting relevant data from the SIS and importing it into the Dashboard Spreadsheet. Further, formulas were added to the spreadsheet to automate the calculation of descriptive statistics.

Once that second design was completed, the Consolidate Committee reviewed it and found it more complete and satisfactory given the anticipated uses of the data to recommend student choices for courses, respond to leaders' requests for information, and otherwise improve the school curriculum. The Consolidate Committee also sought input from the professionals who managed the SIS to confirm the work of exporting data from the SIS did not significantly increase that person's duties. One additional design iteration was undertaken so that correlations between data sets contained in the Dashboard Spreadsheet could be displayed graphically and numerically.

At the end of the initial design/ construction session and the two following iterations, the Consolidated Committee reported, "the dashboard is ready for use, and it should be populated with real data." A final design/ construction iteration followed during which two individuals from the Consolidated Committee, along with the author met to improve the appearance of the data, lock cells on the spreadsheet to prevent accidental editing of data, anonymize data as appropriate to maintain privacy, and otherwise finalize the file.

The second new ICT tool designed for the assessment system was an installation of Mahara (n.d.), an open source platform for creating, curating, and presenting electronic portfolios. In adopting this platform, the

Consolidated Committee sought to resolve the problem of attempting to use the SIS to capture authentic student work. They observed, "not only are the comments we have been using difficult to manage, they don't seem well-understood by parents, and it is impossible to record students' reflections." This system also underwent several design/ construction iterations.

The first design established "'proof-of-concept' as no one on the committee has any experience using Mahara, so we want to see it and determine of it is an improvement over [the previous platform]." Two additional design/ construction iterations followed. During the first, Mahara was reconfigured to be integrated with the learning management system that had been adopted by the school. The LMS had been adopted under an earlier initiative to deploy a one-to-one computing initiative, but it had not been widely used. Once two LMS and Mahara were integrated, the second design iteration added rubrics to the LMS and configured the systems so Mahara portfolios could be submitted to the LMS. The Consolidated Committee recommended, "all assessment and reporting of projects and authentic learning be migrated to the LMS and Mahara. This will allow for full documentation of learning and reporting of qualitative data exclusively through this system."

With the three technology-based assessment tools, the system was completed. The SIS was used to display traditional course grades, and also as a source of data for the Dashboard Spreadsheet. After several drafts, the dashboard displayed both individual student data and data summarizing entire cohorts of students. Finally, the electronic portfolio platform was integrated with the LMS and configured to capture ad display artifacts of works created by students and associated qualitative data.

Designs of Technology-Rich Assessment

The final phase of educational design research (McKenny & Reeves, 2012) is evaluation/ reflection during which the researcher-practitioner both determines the degree to which the interventions solved the problem as it was locally instantiated and articulates generalizations that can be used to solve similar problems in different settings. In this case, there appears to be support for three observations regarding the design of technology-rich systems to support assessment in schools; the system is evaluated in the Discussion.

Efficiency.

One advantage of using ICT to manage the vast assessment data is the efficiency with which sophisticated analyses can be deployed. When initially designing the Dashboard Spreadsheet, the Consolidated Committee sought to minimize the amount of data that would need to be input. During design/ construction, they discovered the data

was received in the school or existed in the SIS in formats that could be imported with minimal effort. Batch uploading and simply copying and pasting allowed "us to enter data far more quickly than expected." The Consolidated Committee reported, "in the time we apologized for asking [the SIS manager] to get us the data we wanted, he had it into the dashboard." Further, the formulas used to generate the descriptive statistics could be written into a single spreadsheet file which was saved as a template, so creating the dashboard for new students and for new cohorts was "the matter of unlocking the file and making a few clicks."

Because Mahara was a new tool for all faculty and because the faculty expressed some reluctance to adopt it because previous platforms had proven ineffective and were abandoned, the Consolidate Committee deployed the system in steps. For example, as faculty gained experience with Mahara, they used paper copies of rubrics to assess students' work. As they become more familiar with the platform and realized digital rubrics in the LMS could be used to reduce the time needed to provide and review feedback, the faculty asked for the design/ construction iteration during which rubrics were added to the integrated LMS and Mahara system.

Ease of Use.

For several years, the school had been attempting to increase the use of online classrooms as well as consistently develop electronic portfolios; those efforts had largely been unsuccessful because "the systems all required a separate log on." With the single step of allowing single sign-on so that the LMS and Mahara used the same credentials as students and teachers used to access email and other web-based services, this assessment system was found to be easier to use. Teachers reported, "it seemed a simple barrier, but having even one or two students fighting to remember how to log on made us avoid them. Now that the systems are integrated, my department uses both [LMS and portfolios] more than we did."

Ease of use was also identified as an obstacle to using the dashboard to the extent it was intended. Specifically, teachers found it difficult to customize the graphs so they displayed the correlations between scores and other data in which they were interested. While most agreed the steps were not difficult, they were unfamiliar, so many avoided the need to learn by finding a colleague who was willing to create the graphs he or she needed.

Understanding Learning.

The final observation that was consistently reported by both the Consolidated Committee and the faculty who used the assessment system was that it had increased their understanding of the degree to which students were learning. Their understanding appeared to be for both individuals and for cohorts.

The addition of qualitative data regarding students' projects and authentic work in the electronic portfolios "gave [teachers] insight into how students were applying what we studied in class." Teachers also reported the technology was one factor that contributed to their ability to gain this insight. "Because the system was always available and easy to use, it was used by more teachers than the other portfolios and they added more work so we had more to look at," observed on teacher, and that was followed almost immediately be another who added, "having the rubrics there also made it more clear to us what we thought we should be seeing, so I could focus on what the [student] did and what we thought they should be doing." The qualitative data that was consistently gathers and easily assessed led to deeper understanding of individual students.

The Dashboard was found to be particularly useful for understanding students as a cohort. The fact that data was compiled for all of the students in a grade in a single place made data available in a manner that it was not previously. "Especially for the students in the middle quartiles, the dashboard seemed to given of information we did not have previously," reported the teachers. "We could see differences between what they had taken for courses, for example, and see that those who had taken certain classes appeared to have higher tests scores. This is useful for making recommendations," added the principal.

Discussion

There two purposes of this paper. First, to describe the design of the assessment system. In this case, the technology-based assessment system appeared to provide educators with information that they did not have prior to the deployment of the system. The additional data appeared to be both the result of the system being easier to use than other systems and the system being designed to gather data that was previously lost or not collected. Because the system was designed through iterative processes with changes being recommended by users of the system, it appears to have more closely reflected their needs and their capacity to use it.

Second, to articulate generalizations regarding the design of ICT systems in school. These data appear to support the conclusion that the unified theory of acceptance and use of technology (UTAUT) (Venkatesh, Morris, Davis, & Davis, 2003) accurately describes the factors that predict the use of ICT in educational populations. Specifically, UTAUT posits effort expectancy is positively associated with the intention to use technology. In this case; case of use and efficiency, both factors associated with effort expectancy, were reasons teachers cited for using the assessment system. Also, performance expectancy, which is grounded in job fit is posited to positively affect

technology use. In this case, teachers claimed the system helped them understand their students learning, thus they were abler to perform essential job function.

While this project does appear to have resulted in an assessment system judged to be an improvement over those used previously, and this improvement is recognized as an aspect of user-based research (Bereiter, 2002), it is unknown the degree to which this improvement was recognized by all stakeholders. Because the assessment system was an initiative begun by the principal who had supervisory authority over all of the faculty, and the faculty is fewer than 30, it was difficult to identify potentially dissenting voices. Further, it is unknown the degree to which this specific effort and the iterative nature of it was responsible for the observations. Given the authority of the principal, any new assessment system could have produced these results.

Conclusion

Despite the limitations for the research, there does appear to be support in these data that an assessment system that used ICT configured to provide both quantitative and qualitative data regarding students' performance in three dimensions of learning contributed to teachers understanding of their students. Both the data they already gathered were more easily manipulated and displayed and they were able together data the was previously lost because it was not consistently collected.

References

Andrews, D. H., & Wulfeck II, W. H. (2014). Performance assessment: Something old, something new. In *Handbook of research on educational communications and technology* (pp. 303-310). Springer New York.

Bennett, R. E. (2002). Using electronic assessment to measure student performance: Online testing. *State Education Standard*, *3*(3), 23-29.

Black, E. W., Beck, D., Dawson, K., Jinks, S., & DiPietro, M. (2007). Considering implementation and use in the adoption of an LMS in online and blended learning environments. *TechTrends*, *51*(2), 35-53.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, *5*(1), 7-74.

Bereiter, C. (2002). *Education and mind in the knowledge age*. Mahwah, N.J: L. Erlbaum Associates.

Cumming, J., & Maxwell, G. (1999). Contextualising authentic assessment. *Assessment in education: Principles, policy & practice*, *6*(2), 177-194.

Cumming, J., Maxwell, S., & Wyatt-Smith, C. (2016). School leadership in assessment in an environment of external accountability: Developing an assessment for learning culture. In *Leadership in Diverse Learning Contexts* (pp. 221-237). Springer International Publishing.

Dede, C. (2010). Comparing frameworks for 21st century skills. In J. A. Bellanca (ed.) *21st century skills: Rethinking how students learn* (pp. 51-76). Bloomington, IN: Solution Tree Press.

Frey, J. H., & Fontana, A. (1991). The group interview in social research. *The Social Science Journal*, *28*(2), 175-187.

Glaser, B. G. (1965). The constant comparative method of qualitative analysis. *Social Problems*, *12*(4), 436–445.

Harlen, W., & James, M. (1997). Assessment and learning: differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice*, *4*(3), 365-379.

Herrington, J., Reeves, T. C., & Oliver, R. (2006). Authentic tasks online: A synergy among learner, task, and technology. *Distance Education*, *27*(2), 233–247.

Krajcik, J., & Shin, N. (2014) Project-based learning. In R. K. Sawyer, (Ed.). *The Cambridge handbook of the learning sciences* (Second edition) (pp. 298-318). New York, NY: Cambridge University Press.

Mahara [Computer Software]. (n.d.). Retrieved from https://mahara.org/, August 28, 2015.

Manpower Group (2015). Talent shortage survey: 10th annual talent shortage survey. Retrieved on March 12, 2016 from The Manpower Group Web site: http://www.manpowergroup.com/wps/wcm/connect/408f7067-ba9c-4c98-b0ecdca74403a802/2015_Talent_Shortage_Survey-lo_res.pdf?MOD=AJPERES&ContentCache=NONE

McKenney, S., & Reeves, T. (2012). *Conducting educational design research*. New York: Routledge.

McKenney, S., & Reeves, T. (2014). Educational design research. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 131–140). New York, NY: Springer New York.

National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Common Core State Standards.* Washington DC: Author.

Pelligrino, J. (2014). A learning science perspective on the design and use of assessment in education. In R. K. Sawyer, (Ed.). *The Cambridge handbook of the learning sciences* (Second edition) (pp. 233-252). New York, NY: Cambridge University Press.

Rajagopalan, K., & Gordon, E. W. (2016). *The testing and learning revolution: the future of assessment in education.* Springer, Chicago.

Stump, E., & Silvernail, D. (2014). Implementation of a proficiency-based diploma system: Early experiences in Maine. Gorham, Maine: Maine Education Policy Research Institute in the School of Education and Human Development, University of Southern Maine. Retrieved from http://digitalcommons.usm.maine.edu/cepare_proficiency/2

van Laar, E., van Deursen, A. J. A. M., van Dijk, J. A. G. M., & de Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, *72*, 577–588.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, *27*(3), 425–478.

Walvoord, B. E. (2010). *Assessment clear and simple: A practical guide for institutions, departments, and general education*. John Wiley & Sons.

Webb, M., & Gibson, D. (2015). Technology enhanced assessment in complex collaborative settings. *Education and Information Technologies*, *20*(4), 675-695..