INCREASING STUDENT ACHIEVEMENT THROUGH MEANINGFUL, AUTHENTIC ASSESSMENT

Nathan Balasubramanian
Applied Technology Teacher
Angevine Middle School
United States
nathan.bala@bvsd.org

ABSTRACT

This paper describes how the author uses an online communication and assessment tool, *SchoolFusion*, to
(a) monitor and manage middle-school students' work and provide them immediate feedback,
(b) collect real-time data on students' understanding of science and engineering concepts, and
(c) use the information gathered to guide subsequent instruction. Quantitative data analysis showed that the mean test scores increased significantly from the pre-test to the post-test across the entire class. Students' responses in online think-writes also revealed students' improved conceptual understanding of scientific and engineering principles.

INTRODUCTION

It was Friday evening, four weeks after school reopened. I noticed my colleague at Angevine Middle School walking out of school carrying a "case" overflowing with over eight reams of students' mathematics worksheets for grading over the weekend. Not long ago I too had carried reams of physics papers for grading over weekends, not always completing my planned marking. As teachers we are often overwhelmed with a backlog of papers to grade and continue with our daily instruction, regardless. Meanwhile, students miss the timely feedback that impacts their learning. Obtaining and using data on students' individual learning in the classroom on a daily basis is a challenge.

On top of all this, with increased calls for accountability in school systems and fear of federal sanctions based on the *No Child Left Behind* (NCLB) Act of 2001, several schools have had to shuffle priorities and focus on reading and mathematics, while other subjects have been relegated to lesser importance. In Colorado, students are "tested" in science only at the end of Grade 8, and soon will be in Grade 5. However, there will be no data on their yearly performance. Not surprisingly, the authors in the *Designs for Science Literacy* (AAAS, 2001) observed that "students learn too little in science" (p. 51) because there is not enough time for teaching and learning science.

Besides, going by reports from the *Program for International Student Assessment* (PISA) http://www.pisa.oecd.org and the third version of the trends in *Trends in International Mathematics and Science Study* http://nces.ed.gov/timss/ (TIMSS) studies, there is growing concern and debate about students' problem-solving abilities. Inspired by Russell's (2004) article in a special issue of *Curriculum Inquiry*, my comments in *Education Week* (Balasubramanian, 2004) highlighted related questions: How do we select the most appropriate materials to teach? How do we determine the most efficient and effective ways to teach this material? Are we teaching as well as we can? Are we teaching as many students as we can? Do the techniques employed in schools enhance students' self-worth and confidence? Do we have ways of using and monitoring ongoing formative assessment?

In this paper, I illustrate how teachers can motivate and empower *all* students through positive and timely feedback by using an online learning management system effectively.

LEARNING MANAGEMENT SYSTEM (LMS)

Besides weblogs and wiki-variations, commercial and open source Internet-based learning management systems – marketed as content management systems or course management systems – are burgeoning. These website-in-a-box technologies are often perceived and presented as the panacea for K-16 and lifelong education. Ellen Rose (2004), peppered with Cassandra-like prognostications, raises valid questions about the ideologies and assumptions underlying the emergent website-in-a-box technologies in her paper. She traces the origins of these purportedly "personalized, meaningful, empowering, and ultimately learner-centered educational environments" (p. 57) to the low-tech teaching machines of the 1950s, computer-assisted instruction (CAI) of the 1980s, and integrated learning systems of the 1990s that promised individualized instruction.

Ellen Rose argues that these emergent technologies are really Trojan horses that seek to replace human interactions in the classroom. They are neither teacher-centered not learner-centered but merely technology-centered systems, she alleges. Although computer managed instruction (CMI) might appear to be a teacher's tireless machine servant to record students' progress, "pre- and post-test scores, and the like" (p. 52), they are driven by the vested interests of CMI researchers, Ellen asserts. According to her, the emerging online website-in-a-box technologies promote a cult of efficiency, atomized content delivery, and static classroom packages while demanding that teachers relinquish their decision-making powers to these technologies.

In this paper, I illustrate how I use one of these commercial content management systems effectively in my classroom to monitor, manage, and use data to inform my classroom instruction – more like a CMI without the vested interests of a CMI researcher. Although I gathered information on students' aptitudes and attitudes in past years, I had not done any formal study on their learning. Since I had access to a new, free LMS that I use as an online communication and assessment tool, I felt confident that I could pursue a more systematic exploratory study using pre- and post-tests.

I was keen on finding out what, if any, students were learning in my classes. I wondered whether I would see an improvement in student scores after instruction. Specifically, would I see improvements that were independent of gender and ethnicity? The following sections illustrate how I designed my courses around four critical organizational questions that guide teaching and learning: learning, instruction, assessment, and alignment questions, following Anderson and Krathwohl (2001, p. 6).

THE LEARNING QUESTION

Students spend just 14% of their time in school each year (Bransford et al., 2000). *What is important for students to learn in the limited school and classroom time available?*

Following a training in Fall 2002, teachers at Angevine are expected to have learning objectives written on the board, following the SIOP Model (Echevarria, et al., 2000). During one of her classroom observations in October 2003, the Principal of Angevine remarked: "Try and focus on what the outcome of the learning is, rather than on the task." My learning objectives have since, gradually, become specific.

Students in the applied technology classes were learning various scientific and engineering principles through building activities. For instance, in Designing Beams, I had specific expectations on students' learning objectives and vocabulary, in contrast to Heavner et al., (2004) "Breaking Beams." My objectives were students will: (a) recognize various types of beam designs; (b) understand forces, especially forces of tension and compression; (c) explain, where in the beam, these forces are greatest and why; (d) rank building costs, depending on materials used, particularly for wood, concrete, reinforced concrete, and steel; (e) design a prototype of a beam using the design process; and (f) test, calculate, explain and evaluate the strength-to-mass ratio of their beams.

THE INSTRUCTION QUESTION

How does one plan and deliver instruction that will result in high levels of learning for large numbers of students?

Since Fall 2004, students at Angevine also take Cornell notes during classroom instruction. Teachers have encouraged and modeled quality note-taking in all classes, and this section describes how students tracked their progress in their technology classes.

Students' learning, their ability to *adapt and improve performance*, is influenced by both motivational and cognitive processes (Balasubramanian, Wilson, & Cios, 2005). I digress briefly to mention that in all my classes, students work in teams on numerous hands-on and minds-on activities by *doing* and *applying* concepts learned at school. Hands-on, in practice, translates to resource-intensive and more planning. Furthermore, early in Fall 2004, I learned how Monopoly-like money can be a significant motivator for learning in middle schools. My description of a creative activity (Balasubramanian, 2005a), provides a brief description of this serendipitous discovery. This "microeconomy" through monetary monitoring has evolved into a full-fledged classroom management system (Balasubramanian, 2005b) and is part of another paper describing ongoing assessment of students' learning.

Before students received any instruction on designing beams, they took a timed online pre-test. Such tests are easy to create using the flexible online K-12 learning management system developed by the www.SchoolFusion.com team. I could experiment with the tool for free because the developers offer their classroom course shells, free for life, for the “first three teachers in any school in United States.” My classes can be viewed at www.angevine.groupfusion.net

Students then received a half-hour instruction that addressed the six learning objectives outlined earlier. With money being a significant motivator, students actively participated throughout the class discussions. The tools students use in class are always free, but each team then bought their supplies. One yard of balsa wood cost $200 and a bottle of wood glue cost $50. After brainstorming and sketching their designs on graph paper in ten minutes, teams had another half-hour to build their design. They used clamps to secure their designs and let them dry over the weekend.

THE ASSESSMENT QUESTION

How does one select or design assessment instruments and procedures that provide accurate information about how well students are learning? More importantly, how does one use this information to inform instruction?

At Angevine, since Fall 2004, teachers are required to have students track their academic progress through Assignment Logs, containing a record of all their graded work. Students had well over ten graded assignments in their logs in just four weeks because they did their quizzes and think-writes on their class website.

In *Systems for State Science Assessment*, Wilson and Bertenthal (2005) summarize 18 assessment approaches to cope with impending NCLB mandates. The *written pre-test*, *drawing* and students’ *problem-solving* abilities in designing beams described so far illustrate three strategies that assess students’ understanding.

During their next class, students tested their beams (Movie # 12 in Balasubramanian, 2005c). The oral *presentations* afford opportunities to assess students’ communication skills. I use these presentations for *peer-assessments* too. After every presentation, the captains of each team confer with team members and write down scores based on a 50-point scoring rubric using five criteria: *design*, *creativity*, *explanation*, *cost efficiency*, and *test-endurance*.

Following their presentations, I wrote down the test results on a transparency and we discussed them as a class. The results of the ten teams are summarized in Fig. 1. The first five were from the first class and the last five are from the second class. Interestingly, in both classes, the beams that won (Team 3 and Team 10) did not withstand the most load. Consequently, students understood the importance of “strength-to-mass” ratio. In the second class, with the results being so close, Team 6 was disappointed at not winning because their beam withstood the greatest load of 229 g. Consequently, they had questions about Team 10’s weighing. The two teams verified each other’s weighing using a triple beam balance and, reluctantly, Team 6 declared Team 10 the winners.

Team No.	Mass of beam (in g.)	Load on beam (in lbs.)	Strength-to-mass ratio (in lbs/g)
1	23.0	358	15.6
2	16.0	185	11.6
3	14.0	254	18.1
4	22.8	41	1.8
5	14.5	44	3.0
6	19.0	229	12.1
7	10.0	29	2.9
8	17.6	174	9.9
9	20.5	213	10.4
10	12.7	161	12.7

Fig. 1. Strength-to-mass ratio of different teams

To follow up on their written assessment, students took an online post-test titled “Beams, Materials, and Forces.” Although the questions in the two tests were the same, the order of questions was different and the tests had different “titles.” Almost all the students thought they were taking different tests. This is another advantage of using *SchoolFusion*. As teachers, we can create these tests, besides mid-term or end-of-term tests, easily using the Online Quiz feature from a repository of questions we have created throughout the semester. Moreover, I found it was easier to cope with common classroom challenges associated with students’ tardiness, truancy, absence, and desire to improve their grades with make-up tests. The instructions, tests, and students’ responses are all online and easy to monitor.

I have embraced this pre- and post-test approach for all my classes since Fall 2005. In the following paragraph, I report results from my Applied Technology classes because I could perform meaningful statistically analyses. In the other classes, with less that 30 students, I have noticed similar trends on improvement in students’ performance but they were smaller sample sizes.

I printed the pre-test and post-test scores from *SchoolFusion* and analyzed them using the *Statistical Package for the Social Sciences* (SPSS). The results are summarized in Fig. 2 below.

Group	N	Pretest Mean (%)	Pretest SD (%)	Post-test Mean (%)	Post-test SD (%)	t-value	df	p- value
Entire Class	34	45.0	23.4	68.9	20.9	4.553	33	< .0001
Caucasian M.	18	47.7	24.3	69.1	21.5	3.621	17	.002
Minorities	16	42.0	22.7	68.8	21.0	2.928	15	.01

Fig. 2. Summary of two-tailed, paired sample t-tests for beams, materials, and forces

Even for such small sample sizes, the two-tailed, paired sample, t-tests show that the mean test *scores increased significantly* from the pre-test to the post-test, irrespective of gender and ethnicity, *across the entire class*. Even with disaggregated data, by ethnicity and minority students (Girls, Hispanics, African Americans, and American Indians), the *mean test scores increased significantly for the two groups*. Clearly, this study cannot be construed as a valid scientifically based research because there was no control group (Slavin, 2003). However, p = .01 means that there is 1% probability that the observed difference among minority students happened by chance.

To complete the Designing Beams activity, students reflected on their learning and completed a confidential *self-assessment* using another grading inventory and rubric. They also reported on team members' individual contributions during the design activity and assigned percentages. Using money, it was easy for them to express percentages, because they were asked "How would you divide $100 between the members of your team based on each individual's contributions?" I was pleasantly surprised by one team in particular. The individual and the other two in the team reported 45%, 45%, and 10%. Although I was moving between the five teams, I had not noticed that this one student was doing little work.

Although contentious, recently Jonassen (2005) argued that "the only legitimate goal of education is problem solving." Students repeatedly hear that the most important concept they will learn in my class throughout the semester is creative problem solving using a systems approach. Students receive a blank grading inventory and rubric at the end of each activity. A sample grading inventory and rubric, illustrating exemplary student responses is available at http://www.innathansworld.com/technology/SampleGradingRubricInventoryNGradingScale.htm The grading inventory reinforces the problem-solving process by requiring students to reflect on their learning while completing their self-assessment.

So far, I described briefly how students' learning might be assessed through multiple measures using written tests, drawing, problem-solving, presentations, peer-assessments, and self-assessments. The following section illustrates how students' work might be assessed though observations, questioning, research products, practical investigations, creative writing, and bundling activities, using examples from a Grade 8 science class.

THE ALIGNMENT QUESTION

How does one ensure that objectives, instruction, and assessment are consistent with one another?

The previous sections illustrated how an outcomes-oriented approach of identifying desired learning goals and then working backwards to develop meaningful learning opportunities and assessments could be used to promote meaningful learning. This backwards design approach (Wiggins and McTighe, 1998) is one way to align assessment with the curriculum.

With high-stakes testing, the slogan "what gets taught is what is tested" is common. Learners do not readily access numerous available online resources (like online discussions and website references). Reeves (2002) observed that often, the learners do not see a relationship between assessment and online resources because they are focused instead on other activities that might help them obtain the highest scores in traditional course assessments.

In my classes, students pay attention to classwork because it counts toward 50% of their grade. The quizzes are 30% and homework is 20%. The homework uses *questioning* and *creative writing* assessment approaches using standards-based online discussions. For example, the Grade 8 science students had to provide thoughtful online responses to the scenario illustrated in Fig. 3. This example illustrates how teachers can use information gathered through essential questions to plan, inform, and modify instruction using feedback from authentic assessments. *Authentic assessments* must be contextualized, be public, require collaboration with others, enable students to show off what they can do, and replicate the actual challenges that typically face a person in the field: conduct original research, analyze the research of others, argue critically, and synthesize divergent viewpoints (Wiggins, 1989).

Discussion Topic

"Given a situation/dilemma/issue, write or critique an investigative process/procedure." (Standard 1.c.2 in the Colorado Science Assessment Frameworks)

For our class discussion today, here is a situation.

Alahna brought an ice pack from Ms. Hanson to place it on her hurt elbow. After sometime, when Alahna put the ice pack down, Bishal noticed that the ziplock bag became smaller and thought the bag was lighter. He mentioned his observation to Alahna and she disagreed with him. The two got into an argument about the mass of Ms. Hanson's ice pack.

1. Do you agree with Alahna or Bishal?
2. Why do you agree with either one of them?
3. You must **justify** your answer with some convincing *experiment* and *reason* that you might use to demonstrate and test your idea.

Created By Nathan Balasubramanian
08/26/05 12:31AM

Fig. 3. Sample standards-based online discussion

Our school nurse typically made over 85 ice-packs everyday. Students use these ice-packs often and could engage with the dilemma presented. They had to demonstrate their understanding of the concepts of density, melting, evaporation, condensation, and closed systems through their thoughtful responses. Students' individual responses are available in their class website at www.angevine.groupfusion.net Here I illustrate how their thinking evolved over time using two sample student responses. Ryan and Leah's initial responses are illustrated in Fig. 4.

Ryan E. Posted 08/26/05 09:26AM

I agree with Alahna because I don't think anything that was in the bag fell out and it might have just looked smaller because the ice all melted.

Leah O. Posted 08/26/05 09:28AM

I agree with Bishal because if you leave an ice pack out of the freezer, then, most likely, the ice will melt into water, and ice is heavier that water, so the ice pack cannot become heavier.

Fig. 4. Sample students' initial responses

After reading their responses, I asked them to divide themselves into three groups based on their beliefs about the ice-pack becoming lighter, staying the same, or heavier in their second class. There were six, nine, and two students in the three groups, respectively. They had to plan their experiments, and record their discussions and observations on Cornell notes. I used these as the prompts for the following class. They discussed the prompts in pairs and subsequently posted their replies (Fig. 5).

Ryan E. Posted 09/06/05 10:10AM

Ryan's reply:

I thought that this was true because none of the water was leaving the bag, and it couldn't get lighter. (B5) I WAS RIGHT. I think Group A cheated or they had a faulty scale, and their measurement was different from ours.

Feng's reply:

I think that group A is cheating and I agree with Ryan's conclusion.

Leah O. Posted 09/06/05 10:10AM

Anna
I think that the groups got different answers because the groups could of weighed the bags wrong, there could of been more ice in 1 bag than the other

Fig. 5. Sample students' response in their third class.

Following this, we discussed the relevant concepts as a class and students then had to post their final responses using the necessary vocabulary (Fig. 6).

Leah O. Posted 09/08/05 09:53AM

[lEaH)
All of the measurements should've been the same, but group A's became lighter, group B's stayed the same, and group C's became heavier.
I think that group A's could've become lighter because some water in one of the bags could've evaporated faster than in the other bag.
Group B's could've stayed the same because their bag didn't evaporate like group A's did.
Group C's could've gotten heavier because the bag could've been buoyant or filled with condensation.

Ryan E. Posted 09/14/05 08:44AM

I think group A got a lighter measurement because the bag might have been open and some of the water evaporated or spilled out.

I think that group B (our group) measured the water first before any could get out.

I think that Group C had a heavier measurement because condencation caused moisture to gather on the outside of the bag.

Fig. 6. Sample students' response in their fourth class

By engaging constructively with the multiple perspectives, students could "confront stereotypes and simplifications about the subject matter" (Gardner, 1991, p. 244) and demonstrate significant learning and development of their analytical and critical thinking skills.

In conclusion: This paper described how students and teachers receive timely feedback on students' performance in pre-tests and think-writes before engaging in formal classroom learning. Although extremely powerful for formative evaluation and informing instruction, the online learning management system does not support graphics during test design. Besides, this exploratory study was not scientifically based because there was no control group. Nevertheless, it demonstrates how *all* students assumed responsibility for learning and I had evidence of individual students' improved conceptual understanding.

REFERENCES

American Association for the Advancement of Science (AAAS). (2001). *Designs for science literacy*. New York: Oxford University Press.

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.

Balasubramanian, N. (2004). Response from Nathan Balasubramanian on 12/26/2004 in response to U. S. Math and Science Skills: Improving or Not? Retrieved from *Education Week* website http://www.edweek.org/tb/2004/12/14/82.html

Balasubramanian, N., Wilson, B. G., & Cios, K. J. (2005). Innovative methods of teaching and learning science and engineering in middle schools. In F. Malpica, F. Welsch, A. Tremante, & J. Lawler (Eds.), *The 3rd International Conference on Education and Information Systems: Technologies and Applications: Vol. 1. Proceedings*, Orlando, Florida, USA (pp. 174–178). International Institute of Informatics and Systemics. Retrieved from http://www.gamestolearn.us/EISTA2005.htm

Balasubramanian, N. (2005a). Nathan Balasubramanian's response to Assignment 17. Retrieved from *Design, Learning, and Collaboration SWIKI* at the University of Colorado at Boulder http://webguide.cs.colorado.edu:3232/dlc-2005/339

Balasubramanian, N. (2005b). Recording students' monetary monitoring in Mr. Bala's classes. Retrieved from Useful Documents, *Microeconomy Data Collection Student Version.pdf* in the Applied Technology class at http://www.angevine.groupfusion.net

Balasubramanian, N. (2005c). # 12. Designing beams. Retrieved from *Movies from the Classroom* at http://www.innathansworld.com/technology/Macromedia%20Flash%20Movies.htm

Echevarria, J., Vogt, M., & Short, D. J. (2000). *Making content comprehensible for English language learners: The SIOP Model*. Boston: Allyn and Bacon.

Gardner, H. (1991). *The unschooled mind: How children think and how schools should teach*. New York: Basic Books.

Heavner, B., Yakachi, C., Zarske, M. S., & Carlson, D. (2004). TE activity: Breaking beams. *Teach Engineering: Resources for K-12*. Retrieved from http://www.teachengineering.com/view_activity.php?url=http://www.teachengineering.com/collection/cub_/activities/cub_mechanics/cub_mechanics_lesson07_activity1.xml

Jonassen, D. H. (2005, September 19). *Let us learn to solve problems*. Retrieved from University of Georgia, Instructional Technology Forum (ITFORUM) Web site: http://it.coe.uga.edu/itforum/paper83/paper83.html

Reeves, T. C. (2002). Keys to successful e-learning: Outcomes, assessment, and evaluation. *Educational Technology*, *42*(6), 23-29.

Rose, E. (2004). "Is there a class with this content?" WebCT and the limits of individualization. *Journal of Educational Thought*, *38*(1), 43-65.

Russell, H. H. (2004). Connections among factors in education. *Curriculum Inquiry*, *34*(3), 267-282.

Slavin, R. E. (2003). A reader's guide to scientifically based research. *Educational Leadership*, *60*(5), 12-16.

Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, *70*(9), 703-713.

Wiggins, G., & McTighe, J. (1998). *Understanding by design*. Alexandria, VA: Association for Supervision and Curriculum Development.

Wilson, M. R., & Bertenthal, M. W. (2005). *Systems for state science assessment*. Washington, DC: National Academy Press. Retrieved from http://www.nap.edu/books/0309096626/html/

Acknowledgment

I would like to thank the Principal of Angevine Middle School, Ms. Isobel Stevenson, for reviewing this manuscript and providing valuable feedback.