STEMs: A Proposal for Calibrated Classroom Assessments that Increase Student Motivation and Provide Authentic Evaluation of Student Learning

By:
Hector Murrieta
&
Gordon Amerson

Publication Date:
3/14/11

Abstract

The purpose of this study was to validate the development and proposal of what the authors call STEMs (Standards Tests to Evaluate Mastery) and have defined them as calibrated classroom assessments that increase student motivation and provide authentic evaluation of student learning. Theoretical and empirical research on classroom assessment and grading, and high school Algebra failure rates were studied to provide the justification for the development and proposal of STEMs. The research indicates that there is a need for classroom assessments that increase student motivation, reduce grade inflation and deflation, and provide teachers with an authentic evaluation of student learning of the content. The research shows that student motivation to complete and expend energy on tasks is largely defined by the assessment environment created in the classroom. Students, educators, and parents would benefit from the development of calibrated classroom assessments that increase student motivation and success. (Contains 1 table)

STEMs (Standards Tests to Evaluate Mastery) are calibrated assessments that are designed to motivate students to take academic risks and make academic decisions about their learning while providing teachers with an authentic measure of the student's level of understanding and achievement of academic standards. They are based on the concept of conation. "Conation is defined as the mental or behavioral tendency to act or strive to act. It involves both motivation…being disposed to act or wanting to act, and volition…taking steps or causing oneself to act. Both motivation and volition are involved in student effort in the classroom." (Brookhart, 1997)

The justification for the efficacy of STEMs will follow from an analysis of their structure and how that structure foments conation, and from empirical and anecdotal evidence of current practices. First, we will explain the physical structure of STEMs and provide a justification for that structure. Next, we will look at current research that substantiates their fundamental premise. Lastly, we will provide anecdotal evidence from years of observation of current practices and results from many classrooms and schools to demonstrate the need for STEMs.

STEMs are 35-problem assessments where the questions are categorized and designed to fit into five levels; the five levels correspond to Bloom's Taxonomy. There are ten level 1, ten level 2, ten level 3, and five level 4 and 5 questions. Level 1 questions correspond to Bloom's "knowledge" level of learning, where students demonstrate their understanding of the standard at the most basic level. Level 2 questions correspond to the "comprehension" level of learning. Level 3 questions correspond to the "application" level of learning. Lastly, level 4 and 5 questions correspond to the "analysis", and "synthesis" and "evaluation" levels of learning, respectively. Each STEM has the standard written at the top, and below the standard the different levels are explained so that students are reminded of what is being assessed and the level of learning they choose to attempt. Each set of problems is preceded by a heading indicating their level.

The two elements of conation drive and serve to explain the structure of STEMs, thus the development of conation in the students begins with the structure itself. The motivational piece of conation is developed by the presentation of the questions. In STEMs, students are able to readily identify the level of questioning they choose to attempt and feel confident with. They will not have to guess at the level of questioning and sift through an entire test to find the problems they can and will answer. Gipps (1994a) found that student motivation is affected by the way instructional activities are presented in the classroom. Motivation follows from a feeling of success, and as students are free to choose the level they will be successful with on a STEM, the motivation will follow. As Brookhart (1997) pointed out, "judgments about what they can do will influence what they will do or how much effort they will expend."

Current classroom assessment practices will spread a varied number of levels throughout the test, forcing the students to attempt every problem to complete the test. For example, a traditional 35 question test may have "knowledge" level questions at the beginning, middle, and end, but some students may never get to the ones in the middle and end because they lose too much time on the higher order questions they encounter and may likely not answer correctly anyway. The end result is an incomplete exam and a student feeling like they don't know anything. Schunk (1994) noted that students make decisions about their self-efficacy. The opposite is also true; why have a more advanced student attempt lower level problems if they can show mastery with the higher level questions? Current practices focus on quantity, and students are graded on how *many* they answer correctly, irrespective of the level of questions they answered. STEMs step away from assessing on quantity and focus on the quality of the questions. As a result, motivation shifts from a focus on answering "as many as I can" to "the ones I know I can."

The other conative element of STEMs, volition, follows from the fact that they do not deceive. On a STEM, a student will never have to guess at their grade or become disappointed with a grade they did not expect. There are fewer things more stressful and tense, to a student than anticipating a low grade on an exam; and fewer things more inhibiting than repeated low exam scores. "Information that a student can use to make himself or herself more competent is intrinsically motivating" (Ryan, Connell, & Deci, 1985). STEMs empower students to choose the grade they will receive on the exam depending on their level of understanding. When students receive the STEM, they evaluate the different levels of questions, and make a determination as to which level they can successfully respond to. For example, on one exam they may feel comfortable answering level 3 questions and on another level 4; they attempt only the level they choose. So, a 35 question test becomes a 10 question one with a student confident of success at that level and a teacher knowing exactly at what level of understanding the student lies. The result is students motivated, as opposed to resistant, to taking exams; the fear of failure has been removed. Nolen and Haladyna (1990) found that students' perceptions that their teachers valued independent thinking and mastery of material were positively related to task orientation and to beliefs in the value of strategies. With STEMs, there is no failure, just differing levels of understanding and performance.

The current need for calibrated exams like STEMs arises from many years of low grades and poor performance in core subjects, particularly math, and specifically Algebra. There are several articles and studies that attest to Algebra failure rates around the country between 20 and 90 percent (see table one). Although it can be argued that the low grades and poor performance result from bad "teaching", ultimately assessment is the element of teaching that teachers use to judge performance and assign marks (grades). "The grades teachers… assign to students' work and performance have long been identified by those in the measurement community as prime examples of unreliable measurement."

(Guskey, 2006). So, it is possible to have good instruction (teaching) but low performance on badly written and designed assessments. STEMs do not interfere with a teacher's style of instruction (how she teaches) but rather provide a genuine measure of a student's understanding of the content.

Grade inflation and deflation are also reduced with STEMs. Conventional grading practices involve a percentage of correct questions from the total number of questions, commonly known as point-system grading (Docan, 2006). So, on a conventional 35 question test, a student would have to score 31.5 to achieve 90% correct, 28 to achieve 80%, 24.5 to achieve 70%, 21 to achieve 60%, which normally correspond to A, B, C, and D grades respectively. However, conventional exams contain two inherent elements that compromise the accuracy with which they measure student understanding of the material. First, conventional teacher created exams tend to have a large number of lower level order of questions spread throughout the exam and a "current propensity to focus on low-level skills" (Shepard, 2000). So, while an A paper means that the student answered 90% of the questions correctly, that doesn't mean that the 90% are at a higher cognitive level. This type of situation causes grade inflation: students with high grades but a minimal level of understanding. Second, the large number of problems on the exams places a greater emphasis on speed than showing mastery. A student taking his/her time may only reach question 25 and answer all of them correctly, but the score will remain a C, not because the student has not demonstrated understanding, but because he/she did not complete all 35. This may cause grade deflation.

An understanding of achievement goal theory (AGT) will help in appreciating how and why the inherent structure of STEMs eliminates the aforementioned problems. AGT proposes that the goal structure of an environment might affect students' motivation, cognitive engagement, and achievement within that setting (Ames & Archer,1988). Wolters (2004) suggests that there are essentially two goal structures that students perceive in classrooms, mastery and performance. Mastery goals primarily focus on engaging in achievement behavior through developing competence (Kaplan, Middleton, Urdan, and Midgley, 2002). With performance goals, success is demonstrated through extrinsic rewards, demonstrating ability, and doing better than other students (Midgely et al., 1998). Research examining secondary school students indicated that a mastery goal orientation was associated positively with students' self reported effort and persistence at academic tasks (Miller, Greene, Montalvo, Ravindran, & Nichols, 1996). Additionally, Condry and associates (1977, 1978) demonstrated that learners' curiosity, interest, and mastery of a subject remain more prevalent when rewards are not involved.

With STEMs, because students are able to identify and choose the level of questions desired, and the questions are calibrated to Bloom's Taxonomy, the teacher and student are assured that the student will receive a grade that reflects the level of understanding. So, students completing five level 4 and 5 problems correctly undeniably deserve an A as those problems are all at an evaluation, analysis, and/or synthesis level; this eliminates grade inflation. Also, as students are required to complete only ten problems per level on levels 1 through 3, and only five for levels 4 and 5, STEMs place due emphasis on student demonstration of mastery and not on the rate at which they answer. STEMs eliminate the doubts and discrepancies regarding student achievement levels.

An intrinsic feeling of success and autonomy are large motivators in any endeavor and at any age. According to Wolters (2004), the types of tasks assigned, the grading procedures, and the degree of autonomy students are provided affect the achievement goals students adopt, and thus embody the classroom goal structure. Hence, compulsion cannot foster intrinsic feelings of success. When people are validated for what they know and do, the intrinsic feeling of success follows; success breeds motivation. Years of anecdotal and empirical data have demonstrated that motivation is one of the largest factors in the success of a student, and STEMs provide a way for teachers to instill that motivation by validating the different levels of understanding that they find in every class.

References

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Student's learning strategies and motivational processes. *Journal of Educational Psychology, 80,* 260–267.

Brookhart, S (1997). A Theoretical Framework for the Role of Classroom Assessment in Motivating Student Effort and Achievement. *Applied Measurement in Education*, 10(2), 161-180.

Condry, J. (1977). Enemies of exploration: Self-invited versus other-initiated learning. *Journal of Personality and Social Psychology, 35,* 459-477.

Docan, T (2006). Positive and Negative Incentives in the Classroom: An Analysis of Grading Systems and Student Motivation. *Journal of Scholarship of Teaching and Learning,* 6(2), 21-40.

Gipps, C. V. (1994a, April). Quality assurance in teachers' assessment. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED 372 086)

Guskey, T (2006). "It Wasn't Fair!" Educators' Recollections of Their Experiences as Students with Grading. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Kaplan, A., Middleton, M., Urdan, T. and Midgley, C. (2002). Achievement goals and goal structures. In. C. Midgley (Ed.), Goals, goal structures, and patterns of adaptive learning (pp.21-53). Mahwah, NJ: Erlbaum.

Midgley, C. et al. (1998). The development and validation of scales assessing students' achievement goal orientations. *Contemporary Educational Psychology, 23,* 113-131.

Miller, R., Greene, B., Montalvo, G., Ravindran, B., & Nichols, J. (1996). Engagement in academic work: The role of learning goals, future consequences, pleasing others, and perceived ability. *Contemporary Educational Psychology, 21,* 388–422.

Nolen, S. B., & Haladyna, T. M. (1990). Personal and environmental influences on students' beliefs about effective study strategies. *Contemporary Educational Psychology,* 15, 116-130.

Ryan, R. M., Connell, J. P., & Deci, E. L. (1985). A motivational analysis of self-determination and self-regulation in the classroom. In C. Ames & R. Ames (Eds.), Research on motivation in education: Vol. 2. The classroom milieu (pp. 13-51). Orlando, FL: Academic.

Schunk, D. H. (1994). Self-regulation of self-efficacy and attributions in academic settings. In. D. H. Schunk & B. J. Zimmerman (Eds.), Self-regulation of learning and performance: Issues and educational applications (pp. 75-99). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Shepard, L. (2000). The Role of Classroom Assessment in Teaching and Learning. *CSE Technical Report,* University of Colorado, Boulder

Wolters, C. (2004). Advancing Achievement Goal Theory: Using Goal Structures and Goal Orientations to Predict Students' Motivation, Cognition, and Achievement. *Journal of Educational Psychology*, 2004, 96(2), 236–250

Table 1

Failure rate	**Author**	**Source**
44%	Duke Helfand	A Formula for Failure in L.A. Schools, L.A. Times, 2006
8.9% increase in failure	Debra Viadero	Algebra-for-All Policy Found to Raise Rates Of Failure in Chicago, www.edweek.org, Vol. 28, Issue 24, Page 11
84%	Emily Richmond	Students show math gains despite high failure rates, http://www.lasvegassun.com/news/2009/feb/04/students-show-math-gains-despite-high-failure-rate/
20% & 33%	Meirowsky, Fellers, Wertenberger	Expanding High School Math Curriculum, Proceedings of the 3rd Annual GRASP Symposium, Wichita State University, 2007
20%-30%	Lori Higgins	Algebra I stumping high school freshmen, Detroit Free Press, 2008
60%	Sidener, Thomas, Turner	Algebra Alert, Superintendent's Urban Principal Initiative (SUPI), 2006
55%	Rick DuFour	Teachers Key to Reversing High Failure Rate in Math, http://www.allthingsplc.info/wordpress/?p=87, 2009