ASEE 2022 ANNUAL CONFERENCE
Excellence Through Diversity
MINNEAPOLIS, MINNESOTA, JUNE 26TH-29TH, 2022 ASEE

Paper ID #37858

The Differing Impact of a New Assessment Framework on Student Success – The Effect of Socioeconomic Factors

Michele J. Grimm (Wielenga Creative Engineering Endowed Professor)

Michele J. Grimm, Ph.D., is the Wielenga Creative Engineering Endowed Professor of Mechanical Engineering and Biomedical Engineering at Michigan State University. Dr. Grimm completed her B.S. in Biomedical Engineering and Engineering Mechanics at The Johns Hopkins University in 1990 and her Ph.D. in Bioengineering at the University of Pennsylvania in 1994. In addition to her interest in injury biomechanics research, Dr. Grimm has dedicated her career to enhancing educational opportunities in engineering and serving as a mentor to students and colleagues.

Ron Averill (Associate Chair of Undergraduate Program)

Associate Professor at Michigan State University, Department of Mechanical Engineering.

Geoffrey Recktenwald (Teaching Faculty)

Geoff Recktenwald is a member of the teaching faculty in the Department of Mechanical Engineering at Michigan State University. Geoff holds a PhD in Theoretical and Applied Mechanics from Cornell University and Bachelor degrees in Mechanical Engineering and Physics from Cedarville University. His research interests are focused on best practices for student learning and student success. He is currently developing and researching SMART assessment, a modified mastery learning pedagogy for problem based courses. He created and co-teaches a multi-year integrated system design (ISD) project for mechanical engineering students. He is a mentor to mechanical engineering graduate teaching fellows and actively champions the adoption and use of teaching technologies.

The Differing Impact of a New Assessment Framework on Student Success – The Effect of Socioeconomic Factors

Abstract

In 2016, Michigan State University developed a new model of classroom education and assessment in their Mechanics of Materials course. This model used a modified mastery approach that stresses formative assessment, guidance in the problem-solving process, and structured student reflection. We now refer to this new approach as SMART Assessment - short for Supported Mastery Assessment using Repeated Testing. The effects of this model have been very positive, and results on overall student success in Mechanics of Materials have been presented in full at prior ASEE conferences.

In this paper, we focus on the effects of this new assessment model on the performance of students who may be at greater risk due to their first-generation status or economic disadvantage, while accounting for other measures such as incoming GPA and performance in the prerequisite course, Statics. The evaluation was conducted across 3.5 academic years and involved 1275 students divided among 9 experimental sections and 6 control sections.

Statistical analysis indicated that there were no significant differences between the performance indices for students in the SMART sections based on their parents' history of university education or their eligibility to receive a Pell Grant. While students in the Traditional section tended to have higher grades in ME222, this cannot be compared directly to the grades in the SMART section due to the difference in grading framework. Previous work, however, has indicated that students who complete the SMART framed sections have a deeper understanding of the course material, as demonstrated by their improved performance on common final exam problems that were evaluated with a mastery-focused rubric.

Background

The SMART pedagogical method was developed at Michigan State University in 2016 [1]. The acronym SMART stands for Supported Mastery Assessment using Repeated Testing. The goal of the SMART method is to address concerning trends in student understanding and performance in STEM courses, especially those that focus on problem-solving. The method was developed in response to growing indications that students were passing classes by achieving a level of learning that is lower than what is expected for an engineering graduate. This lower level of learning is often not discerned by current assessment methods. Preliminary investigations determined that this trend was not due to deficiencies in quality of instruction or lack of student effort. Rather, students had figured out a way to subvert the dominant learning paradigm and pass classes without developing the requisite understanding of fundamental concepts or problem-solving skills [1].

This trend in lower levels of learning has coincided with a rise in student usage of online solution guides for homework and other out-of-class learning activities. This approach eliminates the need for meaningful practice and struggle, which are crucial to the learning process [2,3]. Then, in order to pass examinations, students have increasingly relied on memorization rather than understanding. Instructors have unwittingly facilitated this change by using generous partial credit rubrics that cannot distinguish between simple mistakes and conceptual errors. After passing several college classes by using solution guides and short-term memorization, students begin to use them as their dominant study methods. In fact, after a year or more of this practice, students may not be adequately prepared to pass advanced courses in any other way.

SMART pedagogy leverages known pedagogical best practices to create a learning environment where students are challenged and supported [1]. A key component of the SMART method is to motivate students to practice and struggle through problem-solving. This is achieved through an *irreducible* set of course components that modify the standard course structure [4]. Frequent, formative assessments are used to provide spaced repetition and feedback on skills development. An exam appeals process is used to encourage structured student reflection on process and progress. A mastery rubric is used to set clear expectations of successful problem-solving and to help students develop skills at assessing and troubleshooting their own work. A Compass, or guided problem-solving process, is used to help students develop systematic problem-solving processes [5]. These components work in tandem to guide, support, and motivate students to learn key concepts and establish problem-solving skills.

Previous work on this topic has shown that the SMART method is transformatively effective. Results from previous studies have shown that **students in the SMART method outperform Traditional section students by 1 full standard deviation [1]** when assessed based on correctly answering common final exam problems**.** These results were independent of instructor and were validated over a span of 3 years.

The SMART Assessment grading rubric, examination schedule, and grading strategy [1] have several potential benefits -- ones that apply to all students but that may have an even greater impact on at-risk students:

- Testing anxiety may be reduced by increasing the number of lower-stakes examinations and providing multiple attempts at each one [6-8].
- Multiple chances on each examination result in students receiving direct feedback on areas that need improvement prior to the second attempt at the examination. The first attempt at each examination can serve as formative assessment. The time between examination attempts can then be used to provide additional assistance or corrective intervention aimed at improving understanding or skill related to the identified shortcomings. This process could be formalized, though this has not been done to date.

Understanding whether this framework negatively or positively benefits at-risk students is imperative if it is to be more widely deployed. Students may be at greater risk academically if they are members of one or more of the following categories: being underrepresented based on gender or race; having a disability; being a first-generation university student; or coming from an economically disadvantaged background. An analysis published in 2020 found that men from underrepresented race/ethnic groups and women from non-underrepresented race/ethnic groups were not negatively affected by the introduction of the SMART framework [9]. In fact, both groups earned slightly higher grades than their male, non-underrepresented peers. However, a very small subset of BIPOC women did earn statistically lower grades than their peers, indicating a need for additional research and the possible development of interventions targeted at this group of students.

The current paper looks at 3.5 years of student data to determine how student sub-groups benefit from the SMART method. It is the authors' hypothesis that all subgroups benefit equally from the new method, and -- in this particular study -- the authors investigate the effects of first generation status or economically disadvantaged background. This study looks specifically at student performance in ME222. As such, the purpose of the study is to investigate whether student success in SMART sections of ME222 is similar for all subgroups.

Study Methodology

The SMART method was first implemented in ME222 in the fall of 2016. In this semester, two instructors (B & C) adopted the new method and one instructor (A) maintained a traditional approach to act as a control. A common final exam was used, but course grades were determined independently by each individual section instructor based on their established assessment methods. A similar format was used in the fall of 2017. In the fall of 2018, the control instructor (A) adopted the SMART assessment method (Table 1). Concurrent with the introduction of the SMART format, four sections of ME222 were offered using the traditional assessment system without a comparative, common final exam to the SMART sections. A total of 6 instructors taught the course over this period.

Summer offerings of ME222 have been excluded in this study because of their condensed, 7-week format rather than the typical 15-week semester, making comparisons less relevant. One additional section was excluded as the professor started with the SMART format but shifted to a traditional format of grading after only 2-3 weeks of the semester. It was determined, therefore, that this section could not be classified as either "SMART" or "Traditional".

Table 1 – SMART implementation in ME222 at Michigan State University. Instructor A taught with both methods, Instructors B and C used the SMART method, and instructors D and E taught only using traditional methods. Enrollments varied from 60-130 students per section over this period.

	Fall 2016	**Spring 2016**	**Fall 2017**	**Spring 2018**	**Fall 2018**	**Spring 2019**	**Fall 2019**
Traditional Sections	A	D & E*	A	D		F**	
SMART Sections	B & C		B	C	A & B	C	B & C
Total	3 sections	3 sections	2 sections	2 sections	2 sections	2 sections	2 sections

* Instructor E had two sections this semester

** Instructor F taught in a style that cannot be characterized as Traditional or SMART. This data is excluded from the study.

A total of 1275 students are included in this study. There are 9 sections taught by three instructors (A, B, & C) for the SMART method and 6 sections taught by three instructors (A, D, & E) using the traditional approach. Of this population, 201 (about 16%) are first generation students (Table 2), and 180 (about 14%) are Pell Grant Recipients (Table 3).

All data was obtained from student records. IRB approval was received from the university to allow the researchers to access this data. First generation status was defined as not having a parent who had earned at least a bachelor's degree [10], and this was self-reported by students on their application for admission. Students were classified as being from an economically disadvantaged background if they qualified for a Pell Grant [11].

Table 2 – Study population summary – Study population based on First Generation status.

	Traditional Sections	SMART Sections	Total
First Generation	84 (17%)	117 (15%)	201 (16%)
Parent with a BS Degree or Higher	403 (83%)	671 (85%)	1074 (84%)
Totals	487 (38%)	788 (62%)	1275 (100%)

Table 3 – Study population summary – Study population based on Pell Grant status.

	Traditional	SMART	Total
Pell Grant Recipient	70 (14%)	110 (14%)	180 (14%)
Not a Pell Grant Recipient	417 (86%)	678 (86%)	1095 (86%)
Totals	487 (38%)	788 (62%)	1275 (100%)

Anonymized data on each student also included the students' grade in ME222. If students received a grade for ME222 based on more than one attempt, their first attempt was used in this statistical analysis. Students' grades in the pre-requisite course CE221 (Statics) and their GPA at the beginning of the semester in which they enrolled in ME222 were also obtained. Both of these markers were used as indicators of student preparation prior to enrollment in ME222. All grades at Michigan State University are numeric in nature (e.g., 4.0, 3.5, 2.0, on a 4.0 scale) rather than alphabetic (e.g., A, B+, C). This allowed for simple ratios to be used to calculate indices of performance in ME222 in comparison to indicators of student preparation. The first index (prerequisite performance) divided the grade in ME222 by the grade in CE221. The second index (overall academic record) divided the grade in ME222 by the cumulative GPA through the previous semester of coursework.

Several statistical tests were used to explore the performance of the various study populations as measured through differences in ME222 grades, both based on the actual assigned grade and normalized by CE221 or GPA. A one-sample t-test was used to compare a sub population to the overall ME222 student results, while the Welch's t-test was used to understand the significance of differences between population subgroups. A significance level was set as $p < 0.05$ for all tests.

Study Results

Assessing the impact of the SMART method requires a profile of the students who are taking the courses. The students entering the Traditional and SMART sections had statistically different grades in CE221, with the students taking the Traditional sections having a higher grade than those registering for the SMART sections. In fact, when the CE221 grades for the students in the Traditional ME222 section were examined, the grades were found to be statistically higher than the overall population ($p = 0.012$). However, the two groups were not statistically different in terms of their cumulative GPA as of the semester preceding their enrollment in ME222, comparing both between the two groups and comparing each group with the overall population.

Students who were classified as First Generation did not demonstrate any difference in their preparation for the Mechanics of Materials course when compared with their peers who had at least one parent with a bachelor's degree or higher. This held true for students enrolled in both the SMART section and those enrolled in the Traditional sections, despite the fact that the grades in CE221 appeared to be lower for the First Generation students in the Traditional sections. Similarly, there were no statistically significant differences between the students who were Pell Grant eligible and those who were not when looking at the preparation for either the SMART sections or the Traditional sections.

Table 4 – Incoming student comparisons. Statistically significant differences between subgroups, denoted by **, were only found in the grade in the prerequisite Statics course when comparing students enrolled in the two section types. When comparing to the overall mean, students in the Traditional sections had a statistically higher grade from the prereq course, as indicated by the +.

Population		CE221 Grade Average	CE221 Grade Std. Dev	CUM GPA Average	CUM GPA Std. Dev
SMART	**788**	**3.187****	**0.809**	**3.421**	**0.342**
First Gen	117	3.147	0.828	3.367	0.350
Non-First Gen	671	3.194	0.806	3.430	0.340
Pell Grant	110	3.176	0.829	3.415	0.322
No Pell Grant	678	3.189	0.806	3.422	0.345
Traditional	**487**	**3.296**,+**	**0.754**	**3.447**	**0.337**
First Gen	84	3.160	0.786	3.385	0.314
Non-First Gen	403	3.324	0.745	3.460	0.340
Pell Grant	70	3.210	0.778	3.415	0.309
No Pell Grant	417	3.311	0.749	3.453	0.341
Grand Total	**1275**	**3.229+**	**0.790**	**3.431**	**0.340**

Metrics of student performance in ME222 are presented for the study groups in Table 5, both in terms of their assigned grade and in terms of the ratio of the ME222 grade with measures of student preparation.

Table 5 – Average student performance in ME222 and performance in ME222 normalized by grade in pre-requisite CE221 (preparation index 1) and incoming GPA (preparation index 2). Statistically significant differences, indicated by **, were found between the students in the SMART sections and those in the Traditional sections for each of the three performance indices, with students in the Traditional sections having higher performance (i.e., ME222 grade closer to 4.0, preparation indices closer to 1).

Population		ME222 Grade	Preparation Index 1: ME222/ CE221 Grade	Preparation Index 2: ME222/GPA
SMART	**788**	**2.335****	**0.726****	**0.663****
First Gen	117	2.175	0.667	0.613
Non-First Gen	671	2.363	0.736	0.671
Pell Grant	110	2.350	0.732	0.675
No Pell Grant	678	2.333	0.725	0.661
Traditional	**487**	**2.515****	**0.773****	**0.717****
First Gen	84	2.476	0.795	0.714
Non-First Gen	403	2.524	0.768	0.718
Pell Grant	70	2.457	0.776	0.709
No Pell Grant	417	2.525	0.772	0.718
Grand Total	**1275**	**2.404**	**0.744**	**0.684**

A Welch's t-test (or unpaired variance t-test) was used to assess the significance in differences seen in the performance of groups and subgroups. Welch's t-test allows for comparisons between groups with different variances and is robust against skewness. Results of the Welch's t-test are shown in Table 6. For both SMART and Traditional classrooms, no statistical significance was found between the group considered academically at risk and the rest of the students.

Table 6 – Welch's t-test results for ME222 grades for the SMART and Traditional methods.

Groups being compared	SMART		Traditional	
	t	p	t	p
ME222 Grades – Differences between First Gen and Non First Gen Students	-1.23	0.221	-0.288	0.774
ME222 Grades - Differences between Students Who Did and Did Not Receive a Pell Grant	0.117	0.907	-0.454	0.650

A Welch's t-test was also applied to the calculated preparation indices – the ME222 grades normalized by either the grade in CE221 (Preparation Index 1) or cumulative GPA through the prior semester (Preparation Index 2) (Table 7). There are no statistically significant differences, indicating that the SMART method benefits were equally shared by students in these groups.

Table 7 – Welch's t-test results for Preparation Indices – ME222 grades normalized by CE221 grade or previous cumulative GPA.

Groups being compared	SMART		Traditional	
	t	p	t	p
Preparation Index 1: ME222/CE222 – First Gen vs Non First Gen Students	-1.45	0.148	0.560	0.576
Preparation Index 1: ME222/CE222 – Pell Grant vs No Pell Grant Students	0.161	0.873	0.082	0.935
Preparation Index 2: ME222/GPA - First Gen vs Non First Gen Students	-1.38	0.170	-0.081	0.936
Preparation Index 2: ME222/GPA - Pell Grant vs No Pell Grant Students	0.352	0.725	-0.230	0.819

Conclusions

The SMART method has been shown to improve student understanding and problem-solving skills in ME222 [1]. This study confirms that these benefits are generally shared among all students, even those who are at greater academic risk due to their status of being a First Generation student or those who are from an economically disadvantaged background, as evidenced by their receipt of a Pell Grant.

Though not statistically significant, there was a greater difference between the subgroup means for each of the pieces of compared data (i.e., ME222 grade and both preparation indices) within the SMART sections than in the Traditional sections. First Generation and Pell Grant receiving students also had both slightly lower (but not statistically different) grades in CE222 as well as slightly lower cumulative GPA's than their peers who had a parent that had received a bachelor's degree or higher. Early intervention with these students – perhaps as soon as they declare their interest in studying engineering – may reduce this disparity and provide a stronger foundation before they enter this important mechanical engineering course.

It should be noted that the grades between the two formats of the course cannot be directly compared – as the grades in the Traditional sections are much more dependent on partial credit rather than mastery. As a result, it is not surprising that the students in the Traditional sections have both slightly higher grades in ME222 as well as preparation indices, as the numerators of those ratios tend to be higher in the Traditional sections and there is no difference between the grades in CE222 nor the previous cumulative GPA. The key indicator of learning – and mastery of the subject matter – can be examined by grading common final exam problems using the same mastery-focused rubric. This was previously examined [1], and it is based on these results that the faculty continue to research and expand the application of the SMART framework.

Another notable observation about the SMART method is that the focus on mastery tends to leave the A and B students relatively unaffected in terms of their course grade. The lower, C- students experience the biggest effect. Presumably, this is because students who are in the low C range typically rely much more on 'partial credit' points rather than conceptual understanding. In the SMART approach, these students either modify their study habits and succeed or score well below 50% in the course (and must repeat the course) because they rarely demonstrate conceptual understanding. The analysis in this paper included only the first attempt at ME222 for each student during the study period, and so student grades are expected to trend lower than they would if the final attempt at the course was included. The current policy of Michigan State University is to only allow students to repeat a course if they have received a grade of 1.5 or lower – students cannot repeat a course in which they have earned a 2.0 or better simply to improve their grade.

In conclusion, this study indicates that the SMART framework of assessment does not differentially disadvantage students who may be considered "at risk" in terms of academic success based on their family history of university graduation. This is especially important in the state of Michigan, where only 28.3 percent of residents aged 25 or older have earned at least a bachelor's degree, which is less than the national average of 31.3 percent [12]. In addition, students who received a Pell Grant – an indicator of economic disadvantage – were not negatively impacted by taking the SMART version of ME222. It is interesting to note that the means of all of the performance indicators were more closely matched (i.e., higher p-value for Welch's t-test) for the subgroups that examined economic background than the subgroups that looked at family history. While First Generation students may have a relative other than a parent who has received at least a bachelor's degree (e.g., grandparent, sibling), it may be appropriate to give closer attention to this group of students in order to better understand how parental success in university impacts the undergraduates' approach to their engineering studies in comparison to those whose parents did not earn a 4-year undergraduate degree.

References

1. Ron Averill, Sara Roccabianca, and Geoffrey Recktenwald. "A Multi-Instructor Study of Assessment Techniques in Engineering Mechanics Courses." *Conference Proceedings of ASEE Annual Conference & Exposition*. June 16-19, 2019. Tampa (FL), https://peer.asee.org/31973.
2. Peter C. Brown, Henry L. Roediger III and Mark A. McDaniel, *Make It Stick: The Science of Successful Learning*, Cambridge, MA: The Belknap Press of Harvard University Press, 2014.
3. James M. Lang, *Small Teaching: Everyday Lessons from the Science of Teaching*, San Francisco, CA: Jossey-Bass, 2016.
4. Ron Averill, Geoffrey Recktenwald, Sara Roccabianca, and Ricardo Mejia-Alvarez, . "The Need for Holistic Implementation of SMART Assessment." *Conference Proceedings of ASEE North Central Section Conference.* March 27-28, 2020. Morgantown (WV), https://216.185.13.174/35749.
5. Ron Averill, "The Seven C's of Solving Engineering Problems." *Conference Proceedings of ASEE Annual Conference & Exposition*. June 16-19, 2019. Tampa (FL), https://peer.asee.org/32470.

6. Bangert-Drowns, R. L., Kulik, J. A., and Kulik, C.-L. C. (1991) "Effects of Frequent Class-Room testing," *Journal of Educational Research* 85:89-99, https://doi.org/10.1080/00220671.1991.10702818.
7. Asghari, A., Kadir, R., Elias, H., and Baba, M. (2012) "Test anxiety and its related concepts: A brief review," *GESJ: Education Science and Psychology* 22, 3–8, ttps://doi.org/10.3390/bs12040098.
8. Zeidner, M. (1998) *Test Anxiety: The State of the Art*. Plenum, New York, NY.
9. Geoffrey Recktenwald, Michele J. Grimm, Ron Averill, and Sara Roccabianca, "Effects of a New Assessment Model on Female and Underrepresented Minority Students," *Conference Proceedings of 2020 ASEE Virtual Annual Conference*, June 22-26, 2021, https://peer.asee.org/34507.
10. "Are You a First Generation Student," National Association of Student Personnel Administrators, https://firstgen.naspa.org/why-first-gen/students/are-you-a-first-generation-student, 2020.
11. "Federal Pell Grants are usually awarded only to undergraduate students," Federal Student Aid Office, Department of Education, https://studentaid.gov/understand-aid/types/grants/pell.
12. Wilkinson, M., and French, R., "Why Michigan Needs Newcomers. Told in 5 Data Maps," *Bridge Magazine*, Nov 30, 2017, https://www.bridgemi.com/talent-education/why-michigan-needs-newcomers-told-5-data-maps.