

The Need for Holistic Implementation of SMART Assessment

Dr. Ron Averill, Michigan State University

Ron Averill joined the faculty at Michigan State University in 1992. He currently serves as the Associate Chair of Undergraduate Studies in the Department of Mechanical Engineering. His research focus is on pedagogy, design optimization of large and complex systems, and design for sustainable agriculture.

Dr. Geoffrey Recktenwald, Michigan State University

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.

Sara Roccabianca, Michigan State University

Sara Roccabianca is an Assistant Professor in the Department of Mechanical Engineering at Michigan State University (MSU). She was born and raised in Verona, Italy and received her B.S. and M.S. in Civil Engineering from the University of Trento, Italy. She received her Ph.D. in Mechanical Engineering from the University of Trento in 2011. She then was a Postdoctoral Fellow at Yale University, in the Department of Biomedical Engineering, working on cardiovascular mechanics. Sara's research at MSU focuses on urinary bladder mechanics and growth and remodeling associated with bladder outlet obstruction (e.g., posterior urethral valves in newborn boys or prostate benign hyperplasia in men over the age of 60). Her goals are to (i) develop a micro-structurally motivated mechanical model to describe the non-linear elastic behavior of the urinary bladder wall, (ii) develop a stress-mediated model of urinary bladder adaptive response, and (iii) understand the fundamental mechanisms that correlate the mechanical environment and the biological process of remodeling in the presence of an outlet obstruction.

Dr. Ricardo Mejia-Alvarez, Michigan State University

Dr. Ricardo Mejia-Alvarez obtained his BS degree in Mechanical Engineering from The National University of Colombia in 2000 (Summa Cum Laude), and a MSc degree in Thermal Engineering in 2004 from the Universidad de Antioquia. The same year, he joined the University of Illinois at Urbana-Champaign as a Fulbright Scholar to pursue a MS and a PhD degrees in Theoretical and Applied Mechanics, which he completed in 2010. After concluding his PhD program, he joined the Physics Division at Los Alamos National Laboratory as a Postdoctoral Research Associate and later became a Research Scientist. At Los Alamos, Dr. Mejia-Alvarez conducted research in shock-driven instabilities for the experimental campaign on nuclear fusion of the DOE-National Nuclear Security Administration. In 2016, Dr. Mejia-Alvarez joined the Department of Mechanical Engineering at Michigan State University, where he is currently the director of the Laboratory for the Physics of Living Tissue Under Severe Interactions and the Laboratory for Hydrodynamic Stability and Turbulent Flow. Dr. Mejia-Alvarez was the recipient of the 2011 Francois Frenkiel Award for Fluid Dynamics from the American Physical Society, and the Outstanding Young Alumni Award from the Department of Mechanical Science and Engineering from the University of Illinois.

The Need for Holistic Implementation of SMART Assessment

Abstract

The SMART Assessment model has been developed and tested during the past four years at Michigan State University. This new approach has been shown to significantly increase students' problem solving proficiency while encouraging more effective study habits and a positive learning mindset.

Here, we describe the main components of SMART Assessment along with the natural relationships among these components. The components of SMART Assessment work synergistically, and adopting them in isolation is not recommended. For each component, we discuss the best practices and the importance of a holistic approach to achieve a successful implementation.

Introduction

The SMART (Supported Mastery Assessment using Repeated Testing) Assessment course model aims to reduce or eliminate ineffective study strategies that many students are now using to pass STEM courses [1]. These practices include: 1) copying of homework solutions from online resources; and 2) memorization of a small number of problem solutions that can be used to mimic understanding and maximize partial credit on exams. Enabled by technology and social networking, the rapid proliferation of these detrimental strategies is increasing, and their long term impacts are just now being fully realized. Based on our observations, the net effect is that the current level of learning is well below what is needed for an engineering graduate and much lower than most currently-used course assessment methods would indicate. This is a world-wide trend, and its potential consequences are perilous.

When implemented holistically, the SMART Assessment model has produced consistently positive results, irrespective of instructor or student cohort. Compared to a standard assessment model with graded homework and "correct approach"-based partial credit on exams, students in courses that used SMART Assessment scored between two and three letter grades (20-30 points out of 100) higher on common exams designed to assess mastery [1]. A more detailed analysis of these results shows no statistical difference in the performance of men compared to women or of underrepresented minorities compared to non-underrepresented ethnicities [2]. Implementation has now begun in additional courses [3] and at other universities, where early positive results and feedback indicate that the approach is transferable among universities and department cultures.

There have been a small number of unsuccessful implementations of SMART Assessment, each of them notably omitting important components of the system. In this paper, we discuss the key principles and components of SMART Assessment as well as their interdependencies. We

emphasize the features of successful implementations to serve as a guide to instructors and programs who may choose to implement this approach in the future.

Grading Rubric

The primary feature and key to SMART Assessment is the grading rubric. This sets the expectation of solving problems completely and correctly, which discourages the ineffective strategy of maximizing partial credit. The rubric we have used successfully is described in Table 1.

Table 1. Rubric used to grade each problem on exams.

Competency	Level	Score	Description
Meets Minimum Competency	I	100%	Correct answer fully supported by a complete, rational, and easy to follow solution process, including required diagrams and figures
	II	80%	Incorrect answer due to one or two minor errors but supported by a correct solution process (as in Level I)
Does Not Meet Minimum Competency	III	0%	Incorrect answer due to conceptual error(s)

In Level II scores, there are two necessary conditions for classifying an error as minor:

1. The mistake is a minor algebraic error, computational error, error in units or significant digits, or other human mistake such as misreading a value in the problem statement.
2. If the identified error had not been made, the final solution would have been correct.

When either of these conditions is not true, the error is assumed to be conceptual, and no credit is given. Level III work does not demonstrate minimum competency.

The rubric in Table 1 is in some ways the antithesis of the commonly used “correct approach” partial credit scheme, which is highly subjective and unintentionally encourages students to memorize example problems in an attempt to maximize partial credit. In contrast, the rubric in Table 1 minimizes the benefits of memorizing example problems, while strongly encouraging learning and practice strategies that foster deeper understanding of the underlying concepts.

This model is occasionally misrepresented as a “no partial credit” model, which is clearly untrue. This mislabeling gives the wrong impression and can have the effect of discouraging students. Making mistakes is an important part of the learning process, and the freedom to make mistakes while working toward mastery of new concepts and skills should always be acceptable, and perhaps encouraged. It is much more accurate to describe the current rubric as a “defined partial credit” model, wherein minor human mistakes that are common across many concepts are penalized only mildly, but mistakes in application of concepts and solution steps represent lack of mastery.

Even for minor errors, it is important to impose a penalty. If there is no penalty or if the penalty is too small, then students may not develop the proper appreciation for accuracy, which is a critical part of the engineering mindset. Accuracy is achieved more often when the work process is *consistent* and results are carefully *checked* [4]. We have found that students work hard to develop these strategies under the current rubric.

On the other hand, if the penalty for minor errors is too large, then students may in fact get discouraged. Students who demonstrate an ability to solve engineering problems with an occasional minor error are meeting our most important course objectives, and we want to encourage this level of achievement with a high score.

In other words, the scoring for each problem should give an appropriate weighting to correct process and accurate solutions. The 80% / 20% weighting we have chosen seems to be working very well thus far in this regard.

An additional benefit to the rubric is it sets a clear standard for students to achieve. There is no ambiguity in grading leading to students misunderstanding their scores and more importantly their level of mastery. The standard is clearly set and students themselves are trained to assess the difference between a conceptual mistake and a minor mistake. This is an important part of education. Too often students assume that they missed a problem because of something they “should have known” and they assume the mistakes are simple. This rubric forces the students to confront conceptual errors head-on rather than giving themselves a pass.

Exams Early and Often

If students are experiencing SMART Assessment for the first time, there will likely be a significant adjustment period, during which students realize that their previous strategies for “getting through” a course will not work under the SMART model. If this realization comes too late in the semester, there may not be enough time to make the necessary adjustments, or students may not get enough feedback to convince them that a change is needed.

For these reasons, we have found it is important to schedule at least one, and preferably two, examinations within the first three weeks of the course. These examinations may primarily cover the topics from prerequisite courses, which students presume they already know. This effectively separates the current teaching style or course format from the measurement of knowledge, while

providing direct feedback to students about the expected level of performance. Additionally, students have the opportunity to realize that their pre-requisite knowledge may not be as strong as they think it is.

Beyond the first few weeks of the course, frequent assessment continues to have many benefits. In terms of learning, testing has been shown to be as valuable as many other forms of studying [5,6]. In our first attempt at implementing SMART Assessment, we gave 13 exams plus a final exam. This required a 50% reduction in our usual lecture time, yet students scored 25-30 points (out of 100) higher on a common final exam compared to a traditional class model with more lecture and less testing [1]. We believe most of that benefit came from the mastery style grading rubric, but the amount of testing time (we might also call it intensive practice time) probably played some role as well. In more recent semesters, we have reduced the number of exams to 8-10 plus a final exam with no reduction in overall benefits, but we still consider this to be in the category of frequent testing. (We have now modified the course structure to regain all of the lecture time that was initially lost to frequent testing (as described in [1]).)

We offer two attempts at each exam. So, for example, if students are tested on five modules and have the opportunity of sitting for two attempts at five separate exams, that would result in ten exams during a semester. The two attempts contain different problems and questions, but the structure, topical coverage and level of difficulty are kept the same, as much as possible. The advantages of multiple attempts on each exam are discussed in the next section of this paper.

In the past four years, there have been a few implementations that have not used early and frequent testing, which have not been as successful. In these cases, students did not adjust properly and the high stakes associated with each exam created a very high level of stress. We believe that this counterproductive situation can be avoided using the principles described here.

Most students adapt relatively quickly and successfully to the expectations of SMART Assessment. Based on our limited experience thus far, we have observed that students who take subsequent courses under the SMART model slide right into the proper mindset of studying and practicing during the second course. Does this mean that early and frequent testing is not as crucial in subsequent courses? The answer to this question is not yet clear, but we prefer to err on the side of caution by continuing to test often. The testing process is a key part of learning [5,6] and frequent testing helps to reduce testing anxiety [7-9], so we see no reason to abandon these significant benefits.

Multiple Attempts at Exams

For some students, a mastery level examination feels like a high stakes situation, accompanied by stress and anxiety [7-9]. Allowing multiple attempts at exams may help to reduce some of this stress. And there are many other benefits.

Prior to a second attempt at an examination, students receive direct feedback on areas that need improvement. In this way, the first attempt at each examination can be considered a formative

assessment. Then, during the time between examination attempts students can seek additional assistance or receive corrective intervention aimed at improving understanding or skill in targeted areas. This process could be formalized, though this has not been done to date.

When assessing at a mastery level, one small issue could have an overly-weighted effect on the results. Just a bad day. Misunderstanding a problem. Distraction over a family member's health issue. These and many other issues are legitimate reasons why a student's performance may not reflect their true understanding or ability on a particular exam. With multiple attempts at exams, the effects of these types of issues are greatly reduced, though not eliminated.

Our exams tend to have four sections (described in [1]), and we take the best section score from the two exams to obtain the final aggregated exam score. Due to this scoring process, the class average score on any one of these exam attempts is relatively meaningless.

For a variety of reasons, not all students will give their best effort on the first exam. Some students will use the first attempt as a practice session and then buckle down and prepare harder for the second exam. Some will limit their study to only a subset of the covered topics and focus mostly on problems related to those topics during the exam. For these and many other reasons, the average score on the first exam is often much lower than that of the final (aggregated) exam score.

A similar situation exists for the second exam attempt. Students who performed well on some sections of the first exam will not need to attempt those sections on the second exam. Some students will try to maximize total points by spending more time on some parts of the exam to ensure accuracy while ignoring those parts for which their confidence is low.

The lesson here is to not try to interpret the data or assume anything about class performance based on individual versions of exams within a multiple exam system. We also suggest not sharing the class average or any other performance data with the class, since that data will likely be misinterpreted and may have a negative influence on the attitude of the students.

Compass

If the expectation is for students to solve problems completely and correctly, then some direct attention must be given to this topic and some resources should be provided that demonstrate a clear process to follow when solving problems. We have introduced the concept of a Compass for problem solving [1,4], which serves as a guide to students during their problem-solving practice.

A Compass is a guide, or a set of suggested steps, for solving a certain class of problems. A Compass can be developed for most, if not all, types of problems in science, engineering and math. This is one part of the SMART Assessment approach that will be unique for each course [3].

For example, here is a Compass for drawing a Free Body Diagram (FBD) of a beam, truss or frame structure:

1. Create a *new* drawing of the structure, representing each member as a line.
2. Represent internal connections as either pinned or welded.
3. Define a global coordinate system (GCS) that is convenient for the current problem.
4. Replace all boundary icon symbols with the reaction forces and moments that these boundary supports impose on the structure.
5. Draw all external loads.
6. Include all key dimensions, including units.
7. Label all points corresponding to boundaries, joints, load discontinuities and key sections.

A Compass suggests what direction to go next rather than which detailed steps to take. The details of each step will be problem-dependent, and students will learn to effectively apply the key concepts at each step through varied practice.

A Compass facilitates creativity by reducing the mental load associated with developing an overall solution process. With practice, the solution steps in a Compass become habitual, and this consistency frees the mind to focus on the unique aspects of a problem or to concentrate on performing accurate computations. A Compass also provides structure to a solution process that helps with the communication of that solution.

In the beginning, students may depend heavily on a Compass while they build healthy habits. However, observations of student behavior suggest that a Compass is like training wheels on a bicycle. They enable you to ride without falling when you first get started, but you shed them quickly when you gain confidence in your own abilities. In other words, when the solution steps become instinctual it is no longer necessary to refer to the Compass.

Anecdotally, many students really appreciate having a Compass to guide their practice, and some feel that it had a significant impact on their becoming a skilled problem solver. In addition, when instructors in various sections of a course use a common Compass, consistency among the sections is increased.

A key feature of the Compass is consistency. Giving students a Compass will only confuse them if the instructor follows a different method in class, modifies the Compass mid-semester, or switches between different methods or notations during lectures. Additionally, solutions to exams and other assignments should follow the Compass with rigor. Instructors can point out what is a fundamental concept vs. a convention, and that other texts may use different conventions. However, expecting students to follow a method and using another does not give them confidence in the method.

Level of Difficulty on Exams

In conventional grading methods involving poorly defined partial credit, there is a tendency for exam problems to be lengthy and complicated. Often there is no expectation that students will solve such problems completely or correctly, so the grading is based on interpreting the attempts of students to write something meaningful regarding the approach to solving a problem of this type.

The opposite approach is used in the SMART Assessment model, where the expectation is that most students (those earning a grade of C or better) should be capable of earning at least 80% of the available points on the exam by solving most of the problems completely and correctly, according to the rubric in Table 1. The other 20% (or so) of the credit can be reserved for what we call “challenge problems,” which require a more complicated solution process involving multiple steps and multiple concepts. But even for the challenge problems, a complete and correct solution is required to receive credit.

When the grading rubric is known ahead of time, the job of developing exam problems and grading them becomes easier.

The key is to set the level of difficulty of the exam problems so that the established grading rubric will be an accurate measure of whether a student is achieving the course learning objectives.

These three components – course learning objectives, grading rubrics and exam questions – together define “the bar” that students must reach to pass the course.

Naturally, there is a tendency toward a higher level of difficulty in exam problems. This “creep” must be managed carefully so as to maintain fairness and time limits on exams. As a simple rule of thumb in core courses, if a problem starts to seem “interesting” to the instructor, it is probably getting too difficult for students who are learning this material for the first time. The really interesting problems might be better used in classroom exercises and homework problems.

An additional feature of SMART exams is the timing. In SMART assessment, students need sufficient time to carefully follow a process, review their work and correct issues. To that end, exams should be designed to take no more than 70-80% of the class period. This is especially important because the normal remedy for a lengthy exam is to be extra lenient in grading or to curve the grades. These practices are antithetical to SMART Assessment.

Finally, problems should be written in such a way as to help students build intuition. Intuition comes from practicing problem solving and reflecting on what ‘reasonable’ answers look like. If an exam has un-realistic answers (e.g. a Factor of Safety of 0.0004 or 100000), then the exam will not help students build intuition.

Exam Grading Process

Frequent exams imply a high volume of grading under traditional grading strategies. To make this activity more manageable and to add even greater value to the testing process, we implemented a different process for grading exams.

Below are the steps of the examination and grading process we now use for large sophomore and junior level mechanics courses. The number of steps may seem high, but the net effect of this approach is a *significant reduction in total time* spent to create and grade exams. The biggest benefits, though, may be in student learning.

1. **Make up the exam.** An exam problem should be 100% solvable by a student who has attained the target level of mastery of the topic(s) contained in that problem. Each problem tests a different set of topics and may require a different level of learning. In some cases, it is possible to break up a problem into multiple steps and then allocate a portion of the total problem score to each part. This is a form of partial credit that fits well into the current approach, provided the number of parts is small and subsequent steps are not awarded points when the solution to previous parts is incorrect. In any case, correct answers are always the expectation, not the exception.
2. **Print the exam.** We use an online grading tool called Crowdmark [10], so printing exams involves a few small steps that are not discussed here. These steps take only a few minutes, and this small investment pays big dividends later.
3. **Administer the exam.** This is done in the usual way, except that students rarely leave an exam early. Because correct answers are expected, they use the available time to double and triple-check solutions for completeness and correctness, like an engineer should do.
4. **Digitally scan the completed exams and match them to students.** This is another part of the Crowdmark process. The end result is an organized array of exam pages in a convenient online grading environment that facilitates efficient grading. It works especially well if teaching assistants or teams of people are involved in the grading. For a medium-sized exam or larger, the time spent on the scanning and matching process is comparable to the total time spent flipping pages and shuffling papers in a paper-based grading method.
5. **Grading – Round 1.** In the initial grading round, the answers to each problem are checked for correctness, including units and significant digits. If the numerical value of an answer is incorrect, then a grade of zero is assigned for that problem with no review of the work done. If the numerical value is correct, then the solution is reviewed to ensure that the answer is fully supported by a complete, rational and clearly-communicated solution process. If so, then full credit is awarded for the problem, except that a deduction is made for errors in units or significant digits. If required solution steps are omitted, then

a grade of zero is assigned. This grading approach and the solution requirements are clearly communicated to students at the beginning of the semester. This first round of grading requires minimal time and effort.

6. **Return graded exams to students.** With the press of a button, a pdf of every student's graded exam is sent to them by email. This is a very big time savings compared to passing out exams in class, and ensures compliance with FERPA requirements [11].
7. **Students review graded exams and submit written appeals to receive partial credit for minor mistakes.** With detailed instructor-generated solutions in hand, students are expected to review each step of their work and identify the errors made. If the errors are conceptual, then this review will help to improve understanding. If the solution is incorrect because of a simple mistake, the student may submit a written appeal to the course learning management system (LMS) to request partial credit. Appealable mistakes are defined in the rubric. For each problem, an appeal consists of a short paragraph describing the type of mistake and where it was made, followed by a complete and correct rework of the problem. The problem rework must clearly demonstrate that, in the absence of the identified error, the final solution would have been correct. When this condition is not true, the error is assumed to be conceptual and no credit is given.
8. **Grading – Round 2.** Appeals are reviewed and partial credit is awarded, when appropriate. Appeals are granted at the discretion of the instructor, though a well-trained teaching assistant can manage almost all of these. There will be an occasional judgment call, but most of the appeals are easy to interpret based on the rules described in the rubric. When a student receives a grade of zero on an exam problem, it causes the stress level to increase. For this reason, it is important that appeals be received and processed within a short time window after the exam. A rapid grading process also helps students prioritize sections on the B exam attempt.

The detailed steps above are consistent with the SMART Assessment model, which partly enables this streamlined method. This process can be adapted for many different courses. However, if you need to change a few things to apply this strategy to your courses, note that what happens in one step often influences the best way to perform other steps. In other words, there is moderately strong coupling among the steps.

In addition to the advantages already mentioned, the combination of the rubric and this grading approach:

- sets clear expectations for performance, removing the partial credit tug of war that often exists between students and faculty;
- virtually eliminates the disturbing student strategy of trying to maximize partial credit instead of learning how to solve problems;

- requires *almost* no judgment calls related to partial credit, so scores are as consistent and fair as they possibly can be, assuming the exam itself is written at the appropriate level of difficulty;
- enables teaching assistants or teams of graders to perform most or all of the grading steps with few concerns about grading interpretations or consistency;
- encourages students to perform additional reflection on exam questions and solution procedures, which improves learning [5];
- helps students understand the differences between conceptual errors and minor mistakes, which increases the depth of their knowledge;
- is one of the few existing methods for returning exams to students quickly without violating FERPA (the Family Educational Rights and Privacy Act) [11];
- returns graded exams to students quickly and easily, without wasting any class time;
- reduces cheating, since the original submitted version of all exams is scanned and stored for future reference; and
- may be helpful for collecting data for assessment of ABET student outcomes, again due to the scanning and storing of exams.

For exams with questions or sections that do not involve problem solving, this process still works. For example, multiple choice or matching problems can be scored exactly as described above, with no appeals needed. Short answer or essay questions, for which there may be multiple ways to express a correct answer and a different scale of partial credit may be appropriate, can be scored in step 5 based on a clearly defined rubric.

The elements of the approach described here are not entirely new. Various forms of these ideas have been used by others, though perhaps not assembled in this way. The most important pedagogical elements of this process would work just as well without an online grading tool, though some of the efficiency benefits would certainly be lost.

Homework

In our first few semesters of using SMART Assessment, we did not collect or give any credit for homework. This decision was based on the widespread knowledge that a large percentage of students now copy homework solutions from online resources, making the traditional homework model a high cost, low value activity.

We did, however, assign homework problems for practice. We know that students spent a lot of time practicing because we received some complaints from other faculty that students were spending too much time on our class and not enough on theirs. Aside from being humorous, this story illustrates the power of the SMART model to motivate students toward a different type and level of study and preparation.

In a recent implementation of SMART Assessment, we decided to give 5% credit for submitted homework, with a very important distinction. The homework was graded based on “completeness” and not on “correctness.” This was done to encourage students to work on an

even wider class of problems, including ones that might not appear on exams. This provides an opportunity to rethink the role of homework so that it is even more useful.

To encourage students to work on the homework in a way that would provide the greatest possible benefits, we included the following note in the header of each assignment.

“The purpose of this homework assignment is for you to practice solving problems and exploring concepts without any concern for how your performance will affect your grade. This assignment will be graded for effort and completeness, not accuracy. The use of existing solutions from any source will undermine the goals of the assignment, and is therefore forbidden. You are encouraged to have open discussions with other students in the class, with mentors in the Learning Center, and with the instructors. But the final solutions that you submit must be your own work.”

We do not have any data related to the benefits of this approach to homework, but we believe the potential benefits are high and the cost is relatively low. And if the grading weight assigned to it is low (say 5%), then the risk of grade manipulation from cheating on homework is also low.

Mindset of the Instructor

The importance of this component is often under-appreciated. While the grading rubric is the primary structural component about which all other components are aligned, the instructor’s attitude toward SMART Assessment is often the determining factor in how students perceive the model, which in turn will have a major influence on their motivation and performance.

Change is difficult for most people, and response to change often depends on feedback. When the grades on early exams are lower than students expect, this negative feedback may cause students to doubt the fairness of the course model rather than the effectiveness of their study methods, which have seemingly been successful in prior courses. This is a logical conclusion by the students, one which is difficult to overcome with truth about how learning actually happens.

During this adjustment phase, it is more important to manage emotions than to combat logic with truth. *A positive and encouraging environment* created by the instructor’s attitude and comments are necessary for a successful transformation. Reinforcing the conviction that students are *capable* of achieving the required levels in the course if they make simple changes to their *approach to studying* will have many positive effects.

Students must believe that the instructor has complete faith in the approach and in the ability of the students to be successful under the SMART Assessment model. It is not easy for an instructor to have this faith during a first implementation, so it is recommended that instructors become intimately familiar with the philosophy, principles and previous results of SMART Assessment so that these ideas can be shared in a positive way with students who may challenge the approach.

Here are some comments from two instructors regarding their first implementation of SMART Assessment.

Comments from Instructor 1:

“The first time I heard about the SMART methodology is when I was approached by Dr. Averill and Dr. Recktenwald to participate in the study as the “control section”. I simply kept teaching using the traditional methodology, and I was only required to share the (common) final exam papers so that they could be graded with the SMART methodology. While I suspected some shortcomings in the traditional teaching methods, in my mind I thought that my students were still learning a great amount in my class. I based this feeling on the grades I was assigning at the end of the year. Granted, while grading the finals I caught myself quite often thinking ‘I thought these concepts should have been clearer to the students by now,’ but ultimately, I felt my students were walking away from the class with a good grasp of the material. When I saw how much more knowledge and confidence in the material the students attending the sections implementing SMART had compared to the students in my section, I was shocked. Soon after, I was eager to try the SMART method, and prove that I could also achieve those results with my students.

Of course, there is a learning curve for everyone involved when applying the SMART methodology, students and instructors alike. As a young female instructor, I was worried it would be hard to enforce things like limited partial credit and many tests throughout the term. For this, I think that the support and mentoring I received by Dr. Averill and Dr. Recktenwald was key for me to be able to implement these changes. Also, while there was a little resistance from the students in the beginning, quite quickly most of class was on board with the method. By the end of the semester, some students were thanking me, saying things like ‘I feel like I really get it now’ or ‘I feel like I never worked this hard, but also that I have never learned this much before.’ More students than ever before perceived my efforts as working ‘with them’ aiming to make them successful, rather than working ‘against them.’

It was also quite intense to get in the mind set of preparing many tests that had to be appropriate in difficulty, as well as adhere to the Compass, which was new to me. In the end, this turned out to be mostly a matter of practice. I think I was on board with the ‘rhythm’ of the class about 4 weeks in, and I am confident it will not be an issue the next time I teach this class. However, I believe that potentially having a repository of problems to use could greatly help the implementation.

Finally, it was a great feeling when at the end of the semester my students demonstrated a level of knowledge and performance equal to the sections taught by the other instructors who were using this method!”

Comments from Instructor 2:

“I thought that this method would be discouraging to the students, given that each exam iteration prior to appeals for partial credit would show very low grades. Some of them were in fact initially discouraged, but the initial grade anxiety gradually decreased. I eventually had students coming up with meaningful questions based on their revision of exam problems during the appeals process. I felt that they were now thinking about learning the basic concepts as their tool to maximize their grade.”

And here is an unsolicited email comment from a former student, reflecting the transition that often occurs during a first semester under the SMART Assessment model:

“I wanted to let you know that I believe in your teaching method of testing to learn. Despite not finishing with as high of marks as I had hoped for, I saw continuous improvement in my scores and learning strategy throughout the semester when I have often times experienced the opposite. I believe that many students would agree with me in hindsight that the pressure of frequent exams as opposed to required homework result in more intentional studying, and that the focus required to take an exam of that caliber was very beneficial.”

In the student comment above, the phrase “in hindsight” is notable.

Training Teaching Assistants and Graders

A necessary part of the process is training Teaching Assistants (TAs) and graders (if they are part of the instructional team) in the new process. This is important in several ways. First, their approach needs to be consistent with the Compass. They may have learned a different method, and will need to adjust to the approach used in the course in order to reduce confusion. Second, they are ‘in the trenches’ with the students; like faculty, they need to have faith in the overall course model and in the solution approach embodied in the Compass. Finally, some TAs have a tendency to be overly generous with points when grading. They may also have the mindset that a correct answer should get full points, even when critical steps are missing. In the SMART system, the right answer is a necessary but not a sufficient condition for receiving full points on a problem. Graders also need to assess the completeness, quality and consistency of the approach.

Training Students

Rarely do instructors train students on how to take exams and succeed in the class, but SMART Assessment is a new paradigm for most students. If students are to give the SMART model a chance, then it is important to explain the philosophy and the details of the implementation, express confidence in the students’ abilities to achieve the required levels of mastery, and suggest best practices for studying and test preparation. Once the method has been implemented successfully in one course, these steps continue to be important, but the bulk of this work will

now be done by their peers who have successfully finished the course and realized how much they learned!

Curving of Grades

This is NOT a component of SMART Assessment. It is mentioned here because curving of final grades has the potential to undermine ALL of the positive effects of the method.

Compared to traditional loosely-defined partial credit grading methods, the rubric in Table 1 places pressure on students to perform at a higher level. Solving a variety of problems completely and correctly is not easy. The only way to motivate most students to attain that level of competency is by directly linking the course grade to that expectation. Curving removes the pressure that causes the effects we seek.

If students know (or expect) that their final course grade will be curved upward in some way, then many of them will not expend the extra effort to achieve the desired level of competency. They will put their faith in the “magic of curving” to attain a reasonable grade, and the crucial link between performance and the grade will be lost.

Conclusions

When implemented holistically, the positive benefits of the SMART Assessment model have been demonstrated uniformly across multiple courses by multiple instructors. A small number of partial implementations have been less successful, largely due to a loss of completeness, connectedness and consistency among the components of the model. We thus conclude that a holistic approach is necessary, especially when students are experiencing the model for the first time.

References

1. Averill, R., Roccabianca, S. and Recktenwald, G. “A Multi-Instructor Study of Assessment Techniques in Engineering Mechanics Courses.” *Conference Proceedings of ASEE Annual Conference & Exposition*. June 16-19, 2019. Tampa (FL).
2. Recktenwald, G., Grimm, M., Averill, R. and Roccabianca, S. “Effects of SMART Assessment Model on Female and Underrepresented Minority Students.” *Conference Proceedings of ASEE Annual Conference & Exposition*. Submitted for review. June 21-24, 2020. Montreal (Canada).
3. Recktenwald, G. and Averill, R. “Implementation of SMART Assessment Model in Dynamics Courses.” *Conference Proceedings of ASEE Annual Conference & Exposition*. Submitted for review. June 21-24, 2020. Montreal (Canada).
4. Averill, R. “The Seven C’s of Solving Engineering Problems.” *Conference Proceedings of ASEE Annual Conference & Exposition*. June 16-19, 2019. Tampa (FL).

5. Brown, P.C., Roediger III, H.L. and McDaniel, M.A. *Make It Stick: The Science of Successful Learning*, Cambridge, MA: The Belknap Press of Harvard University Press, 2014.
6. Lang, J.M. *Small Teaching: Everyday Lessons from the Science of Teaching*, San Francisco, CA: Jossey-Bass, 2016.
7. 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>
8. 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.
9. Zeidner, M. (1998) *Test Anxiety: The State of the Art*. Plenum, New York, NY.
10. Crowdmark. (2020) Crowdmark, Inc. <https://crowdmark.com/>
11. <https://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html>