3-19-2019

Mastery-Based Testing in Undergraduate Mathematics

J.B. Collins
University of Mary Washington

Amanda Harsy
Lewis University

Jarod Hart
University of Kansas

Katie Anne Haymaker
Villanova University

Alyssa Marie Hoofnagle
Wittenberg University

See next page for additional authors

Recommended Citation

Collins, J., Harsy, A., Hart, J., Haymaker, K. A., Hoofnagle, A. M., Janssen, M., Stewart Kelly, J., Mohr, A. T., & O'Shaughnessy, J. (2019). Mastery-Based Testing in Undergraduate Mathematics. *PRIMUS, 29* (5), 441. https://doi.org/10.1080/10511970.2018.1488317

Mastery-Based Testing in Undergraduate Mathematics

Abstract

Mastery-based testing is an assessment scheme that challenges students to provide complete solutions to problems derived from clear course concepts. Students are allowed multiple attempts to demonstrate mastery, which helps create a classroom environment where students value persistence toward thorough understanding. In this paper, we describe in detail the benefits and implementation of mastery-based testing in college mathematics courses. We also summarize student response data that show positive reactions to this testing method.

Keywords

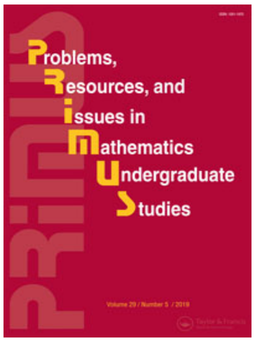
assessment, growth mindset, mastery

Disciplines

Higher Education | Science and Mathematics Education

Authors

J.B. Collins, Amanda Harsy, Jarod Hart, Katie Anne Haymaker, Alyssa Marie Hoofnagle, Mike Janssen, Jessica Stewart Kelly, Austin Tyler Mohr, and Jessica O'Shaughnessy



PRIMUS

Problems, Resources, and Issues in Mathematics Undergraduate Studies

ISSN: 1051-1970 (Print) 1935-4053 (Online) Journal homepage: https://www.tandfonline.com/loi/upri20

Mastery-Based Testing in Undergraduate Mathematics Courses

J. B. Collins, Amanda Harsy, Jarod Hart, Katie Anne Haymaker, Alyssa Marie (Armstrong) Hoofnagle, Mike Kuyper Janssen, Jessica Stewart Kelly, Austin Tyler Mohr, & Jessica OShaughnessy

To cite this article: J. B. Collins, Amanda Harsy, Jarod Hart, Katie Anne Haymaker, Alyssa Marie (Armstrong) Hoofnagle, Mike Kuyper Janssen, Jessica Stewart Kelly, Austin Tyler Mohr, & Jessica OShaughnessy (2019) Mastery-Based Testing in Undergraduate Mathematics Courses, PRIMUS, 29:5, 441-460, DOI: 10.1080/10511970.2018.1488317

To link to this article: https://doi.org/10.1080/10511970.2018.1488317

Accepted author version posted online: 15 Aug 2018.
Published online: 19 Mar 2019.

PRIMUS, 29(5): 441–460, 2019

ISSN: 1051-1970 print / 1935-4053 online
DOI: 10.1080/10511970.2018.1488317

Check for updates

Mastery-Based Testing in Undergraduate Mathematics Courses

J. B. Collins, Amanda Harsy, Jarod Hart, Katie Anne Haymaker, Alyssa Marie (Armstrong) Hoofnagle, Mike Kuyper Janssen, Jessica Stewart Kelly, Austin Tyler Mohr, and Jessica OShaughnessy

Abstract: Mastery-based testing is an assessment scheme that challenges students to provide complete solutions to problems derived from clear course concepts. Students are allowed multiple attempts to demonstrate mastery, which helps create a classroom environment where students value persistence toward thorough understanding. In this paper, we describe in detail the benefits and implementation of mastery-based testing in college mathematics courses. We also summarize student response data that show positive reactions to this testing method.

Keywords: Assessment, growth mindset, mastery

1. ASSESSMENT AS A GUIDE TOWARD MASTERY

Students of mathematics are asked to explore a rich and diverse landscape of ideas. Some features are more difficult to surmount than others; for a calculus student, differentiation of polynomials is a gently sloping hill, whereas the Fundamental Theorem of Calculus is more like a sheer cliff. Assessment is one of the ways in which we guide students to noteworthy landmarks and assist with the climb.

As teachers, we use assessment as an aid to determine student learning and achievement as well as communicate feedback to our students on their understanding of course concepts. Frequently, an A letter grade corresponds to exceptional performance and learning, whereas a C is

Address all correspondence to J. B. Collins, Department of Mathematics, University of Mary Washington, 1301 College Ave., Fredericksburg, VA 22401, USA. E-mail: jcollin2@umw.edu

average, and a B somewhere in between. What precisely is indicated when we average these letter grades over many assignments? Does an A signify exceptional competency on all learning goals or merely satisfactory competency on all learning goals? What does a C mean? Did the student attain some of the outcomes at a satisfactory level and others not at all? Did the student demonstrate partial understanding of all of the course concepts? When a grade is determined by a weighted average it could be either, or a mix of both. A final percentage grade of 70% could be obtained by earning a 7 out of 10 on all work, or 10 out of 10 on 70% of the work and zeros on the remaining 30% of the work.

Using assessments based on accruing points treats the landscape as though it were flat and tasks students with traversing 70% (or 80% or 90%) of the terrain. Such guidance often provides a perverse incentive to approach a new mountain, march around the gentlest parts of the base, and leave it behind without experiencing its important and beautiful heights. Indeed, traditional partial credit-based exams frequently encourage students to focus their attention on "earning points" as opposed to understanding why their work was assessed the way it was, and keeps them guessing about what constitutes the big ideas of the course. Moreover, maintaining consistency in awarding partial credit is difficult, even if the same person is awarding the points from exam to exam. Often teachers feel the need to give some partial credit if the student writes down anything relevant at all. In short, traditional, partial credit-based exams are often "unpleasant" for faculty [11]. Moreover, historically, the original function of grades as a private communication method gave way to using grades as an external communication tool [12], a transition that may result in an over-emphasis on summative assessment in the classroom.

A traditional exam structure is high-stakes: if a student does not demonstrate mastery of the content assessed on a given midterm, the student is usually stuck with that grade, regardless of whether or not the student eventually understands the relevant material. Traditional exams do not give students an opportunity to achieve success through effort and practice. Developmental psychologist Carol Dweck states that the most motivated students believe that they can develop their abilities through their effort towards learning concepts (the "growth mindset") rather than believing their achievement will be limited by innate intelligence (the "fixed mindset") [7]. According to Dweck, looking at the response of students to failure or mistakes, there is a clear dichotomy in reaction. Either students are "helpless" or "mastery-oriented." Those who have a more helpless/fixed mindset of learning quickly lose confidence in their ability and question their intelligence. On the other hand, students with the growth perspective take failure as an opportunity to learn from mistakes. They view their failures not as a reflection on their intelligence, but rather a challenge which will require more time and new techniques ([9], p. 9). As teachers, we would like our students to have a growth mindset towards mistakes, and thus should aim to have assessments which support this goal.

Assessment techniques have already been introduced to counteract entity (fixed) learning culture [4]. For example, Beatty discusses a standards-based grading scheme he used in an introductory physics course [1]. In the course, students were not awarded points per test, but points per standard or objective for the course. In another university physics course, Studman used a mastery learning method in which a certain set of objectives were given in small units to the students, and they had to achieve mastery on them to pass the course [13]. In Studman's course, students were allowed multiple attempts to show mastery, and testing was considered part of the learning process. Some teachers even use methods of grading in which no points are used. Instead, a grade is assigned based on how well the student meets the clearly laid out course objectives [2].

Since 2014, the authors have been exploring a method of assessment in our mathematics courses which we hope helps foster a growth mindset approach to learning [7]. We call this assessment technique *Mastery-Based Testing* (*MBT*). As we guide our students through the rich and vibrant landscape of ideas, MBT proceeds under the assumption that those final arduous feet climbing atop a peak are the most transformative. Indeed, the students' resistance to face a particular challenge is evidence that they have much to gain by undertaking it. Therefore, rather than directing students to explore a fixed percentage of the total area, we ask students to conquer a fixed number of carefully chosen peaks. When a student finally reaches a pinnacle such as genuinely understanding the definition of a logarithm, solving a related rates problem in full, or putting the finishing touches on a proof by induction, they experience growth on at least three fronts. First, students take careful stock of their content knowledge, patching up deficiencies in prerequisites and grappling with the subtleties of advanced ideas. Second, they synthesize this knowledge to produce a rigorous solution to a challenging mathematical problem. Third, they learn to persevere academically and develop the confidence that comes only from pushing past one's perceived limitations. This assessment method helps guide students to the top of these figurative mountains and gives them a chance to train if they are not ready to tackle the peak. Allowing students a chance and reason to revisit old ideas that they have not fully understood supports a growth mindset towards learning, since students who work hard and learn from their mistakes are able to persevere in reaching the top of their personal summit.

Although MBT is similar to other techniques intended to focus on complete understanding of material and counteract a fixed mindset toward learning, it is unique in that it accomplishes these goals by means of standard in-class assessments such as tests and quizzes only. Studman's implementation of mastery learning differs from MBT in that the units are small, and testing can occur at any time, not just during regular tests or quizzes. Hence, large class sizes may be prohibitive for implementing mastery-based assessment at the level described by Studman. Furthermore, in Studman's approach, mastery does not entail full conceptual understanding (as in MBT), but instead a student just needs to show general knowledge of the content, which is akin to getting a C [13]. In Beatty's standards-based grading, points-based scores are used, not mastery grading, but as with MBT, these scores relate to a particular skill or objective for the course [1]. Finally, the "no-points" method discussed in [2] is similar to MBT, but no concept of mastery is introduced. Williams also describes using specifications-based grading in an introduction to proofs course, which demonstrates how a related grading system can be implemented in the college mathematics setting [14]. Rigorously exploring the student impact of these methods of assessment is an ongoing area of research (see [10] for example).

This paper introduces details of implementing MBT in a college mathematics classroom, and also describes logistics and student responses to this assessment technique. In Section 2, we give the motivation and goals of MBT, as well as a sample implementation to highlight the common features that characterize this assessment method. Section 3 provides details on implementing MBT. Student feedback is summarized in Section 4, and Section 5 concludes the paper.

2. GOALS AND CHARACTERISTICS OF MBT

2.1. Goals of MBT

The purpose of MBT is to redesign assessment procedures to overcome some of the undesirable incentive structures that are inherent to typical partial credit examination formats. Our goals are to use MBT to emphasize a clear outline of important course concepts, motivate students to develop a deep conceptual understanding of course materials, to encourage a growth mindset for students, reduce test anxiety, and to do so in a way that is scalable for large classes without imposing an unrealistic workload for teachers. We aim to shift the incentive structure of exam grades to reward students that work to develop a deep understanding of course materials, and deter students from relying on memorization and other study habits that yield superficial understanding. Our testing structure provides students the opportunity to overcome knowledge deficits and initial misconceptions with minimal punitive effects on their final course grade, thus encouraging the attitude that they can improve their mathematical skill through persistence.

2.2. Characteristics of MBT

MBT is characterized by three essential features: clear course concepts, credit only for mastery, and multiple attempts to display mastery. Implementations of MBT vary greatly, but all possess these common attributes. We will explore these three characteristics as we build a basic sample implementation.

2.2.1. Clear Concepts

A good place to begin is to partition the course content into a dozen or more rich concepts. A single exam question should be associated with a clearly defined range of content pertaining to that concept.

For example, suppose you list "Basic Enumeration" as an assessment concept in a discrete mathematics course. You might indicate on a study guide that this item will require the student to demonstrate proficiency with the multiplication principle, permutations, and combinations. Having set clear bounds on the problem, you are free to create as rich a problem as you like. Perhaps you can devise a single problem that satisfactorily incorporates all these ideas. Alternatively, you may choose to divide the problem into two parts, the first of which involves only permutations and the second combining the multiplication principle and combinations, as in the following example:

1. How many words can be made from the letters A, B, C, and D (using each exactly once) in which the A does not come immediately before the B?
2. A ternary string is a sequence made from zeros, ones, and twos. How many six-digit ternary strings contain exactly three zeros?

In this example, it is important not to indicate which part requires which technique, since choosing the right tool for the task is part of the concept.

2.2.2. Credit Only For Mastery

There is no partial credit in a mastery-based exam; a question is either mastered or it is not. A refreshing aspect of the method is that the instructor may use their expert judgment to determine whether a student has understood the important aspects of the problem and award mastery accordingly.

For example, if the solution to a calculus problem is correct save for an inconsequential error in algebra, it may not be a good use of the student's time to study this *calculus* concept again. At the other extreme, a solution that demonstrates perfect technical proficiency with algebra, but poor understanding of the calculus concepts being tested will certainly benefit from studying precisely these concepts a second time. In any case, the standard set for a student to achieve "mastery" of a topic should reflect a high expectation of conceptual understanding of the concept. We can guide students to study the concepts that are most relevant to their understanding of calculus by awarding mastery to the former and withholding it from the latter. Withholding mastery creates a learning opportunity only if students are given additional chances to demonstrate their improved understanding, which brings us to the third key characteristic of mastery-based exams.

2.2.3. Multiple Attempts to Display Mastery

It is important to allow multiple attempts on each concept so that students can incorporate instructor feedback and make progress toward mastery. Moreover, previous failed attempts should not adversely affect a student's grade once the item is mastered. Penalizing failed attempts does little to incentivize timely mastery and reinforces the pernicious myth that one must fully apprehend new concepts immediately to be successful in mathematics.

There are many creative ways to allow for multiple attempts. A simple model for a course with, say, 16 concepts is to have four in-class exams, each of which features questions pertaining to all previous outcomes plus four new ones. That is, the first exam features questions pertaining to outcomes 1–4, the second features outcomes 1–8, the third features outcomes 1–12, and the fourth features outcomes 1–16. The final exam should not feature any new concepts, since students would have only one attempt to display mastery on such questions. Thus, if a student has mastered every concept before the final exam, that student would not need to take the final exam. Alternatively, a traditional final may also be giving if an instructor would like students to revisit ideas or the university requires all students to sit for a final exam.

After each attempt, the instructor returns the graded questions with feedback on how the student could display mastery on future attempts. The instructor must create alternate versions of each question that fairly address the same objective, but differ enough so that rote memorization of previous answers is of no help. Under this exam structure, a student could fail to master objective 1 four times without penalty and display mastery on the final exam, earning full credit for that objective. More commonly, a student will display mastery on the first or second attempt and simply ignore the alternate versions appearing on subsequent exams.

2.3. Benefits of MBT

Examination is a necessary part of math courses to assess student learning, and the way in which we structure exams greatly impacts study habits and outcomes for students. Our stated goals for MBT include restructuring the way students are compelled to learn in order to overcome some common issues in math courses. We would like to highlight some ways in which a mastery-based approach to exams surpasses ones based on points with respect to content knowledge, student mindset, and instructor time. These items, in part, explain the mechanism by which student behavior is influenced by MBT to achieve specific learning goals.

2.3.1. Content Knowledge

Our goal as math instructors is for our students to develop critical thinking skills and a deep understanding of mathematical concepts. Points-based exams encourage studying many topics superficially, since a student who memorizes the relevant formulas and performs some superficial algebraic manipulation may earn half credit or more without having truly engaged with the concepts in the question. On the other hand, a mastery-based approach encourages students to engage more meaningfully with the most important ideas in the course. The mastery-based approach requires students to take those final, difficult steps to develop a complete solution. Sorting out the fine details develops lasting problem-solving skills, even if it comes at the cost of addressing fewer concepts over the duration of the course. Moreover, this attention to detail comes much earlier in the course since students have the opportunity to display mastery on subsequent exams. Contrast this with a typical points-based exam, wherein students may be perversely incentivized to ignore old material until the final exam.

2.3.2. Student Mindset

Mathematical study is exceptionally effective in training students to persevere in solving complex problems. Points-based assessment can undermine this valuable experience by allowing students to halt their progress too early in the problem-solving process. Indeed, it is not uncommon for students to earn passing course grades without having completely solved a single exam problem. Under a mastery-based system, students must persevere in developing a rigorous solution in order to receive credit, and this may require multiple iterations of assessment, critique, reflection, and reassessment. This process of productive struggle is well-suited to the development of the growth mindset belief that failure, far from being evidence of an insurmountable shortcoming, is the very means by which one grows toward richer understanding.

Indeed Dweck shows that when students build towards thinking that their mathematical ability is not innate, they perform better [8]. By giving assessments like MBT which allow students to view their failures not as a reflection on their intelligence, but rather a challenge which will require them to spend more time developing their techniques, we are supporting a growth-mindset of learning as described by Dweck [9]. The reassessment attempts central to MBT provide opportunities for growth that have the potential to improve student outcomes. In a randomized study of sixth grade students, Butler and Modecai found evidence supporting the hypothesis that "intrinsic motivation would be maintained after receipt of nonthreatening, task-related evaluation and undermined after repeated nonreceipt of feedback or receipt of controlling normative grades" [3], which suggests that providing quick feedback along the lines of the MBT system can help to bolster student motivation.

2.3.3. Instructor Time and Pedagogy

The use of mastery-based exams adds value to the time invested by instructors related to office hours and grading. Exam performance is a function of the number of concepts that students understand deeply, which prompts thoughtful questions and productive discussion during office hours and review sessions. Grading can be less time-consuming on a *per question* basis since students tend to submit high-quality responses to a few questions and leave the others blank. Moreover, the time spent determining the appropriate point deduction for various sorts of errors is completely regained. Finally, the very act of providing careful feedback is made more meaningful by the strong likelihood that students will read it closely and focus their study accordingly.

Although MBT may have an impact on instructor pedagogy and planning due to having to establish course concepts before the course begins, this assessment method can be adapted to a wide variety of in-class pedagogical techniques.

3. IMPLEMENTING MBT

We provided some suggestions for implementation in Section 2 along with the fundamental features of MBT. In this section, we focus on the practical details for using MBT in your course. Keep in mind that there is no "right way" to implement MBT, and it can take a few semesters of using this assessment method to determine exactly the policies that fit best with your individual teaching style. Here, the authors address a few key areas to consider when converting a course to using MBT.

Table 1. Sample course concept lists. Calculus I features core concepts in bold

Calculus I	Real Analysis
Basic Limits	Logical Constructs
Delta-Epsilon Proofs	Proof Techniques
Continuity	Properties of Real Numbers
Limits at Infinity	Convergence of Sequences
Definition of Derivative	Monotone Convergence
Product/Quotient/Chain Rule	Cauchy Sequences
Implicit Differentiation	Limits of Functions
Related Rates	Limit Theorems
Mean Value Theorem	Continuity
Graphing Using the Derivative	Properties of Continuous Functions
Optimization Problems	Differentiability
Newton's Method	Properties of Differentiable Functions
Fundamental Theorem of Calculus	Mean Value Theorem
Approximate Area Under the Curve	Riemann Integration
Basic Integration	Properties of Riemann Integrals
***u*-Substitution**	Integration and Differentiation

3.1. Concepts

One of the first major decisions in creating a plan for implementing MBT is to determine the course concepts. Course concepts should encompass the "big ideas" and essential topics of the course. These concepts should be broad enough to allow the instructor to write multiple versions of rich questions. On the other hand, using too few topics may result in an unreasonably coarse grading scale or too few topics to fill all exams. Two samples of course concepts are given in Table 1.

One option is that an instructor may require that students master certain core concepts in order to pass the course. Although unlikely, a student who does not master all core concept questions will not pass the course, even if they have passed all other objectives. One author chooses to incorporate core concepts into her Calculus I course because she feels "[they] allow my students to focus their studies on key topics and be better prepared for follow-on courses." She further refines the Calculus I topics from Table 1 by identifying seven as core concepts (shown in bold).

Writing the exam questions themselves requires the instructor to establish a more focused notion of the specific learning objectives for each concept. DeLong and Winter provide an algorithm for generating lessons based on specific learning objectives (SLOs) [5]. Although the author's method is specifically aimed at in-class activities and lesson-planning, the ideas can also be adapted to develop exam questions in the MBT setting. Indeed, in a subsequent work Delong, Winter, and Yackel discuss assessments in conjunction with SLOs [6] (see Section 4.3.2).

3.2. Reassessment Opportunities

A key component of MBT is providing students multiple opportunities to test on each topic. Revisiting unmastered material allows students to address gaps in their understanding and promotes the development of a growth mindset toward learning. You may choose to limit reattempt opportunities only to subsequent exams as suggested in Section 2, or to allow additional opportunities for students to retest between the major exams. For the purpose of consistency, we call these additional retest opportunities quizzes, but they are not always quizzes in the traditional sense. These quizzes can be an excellent place for students to gain confidence in their knowledge and reduce anxiety before the next exam. Quizzes are also useful if the time allotted for exams is fixed. In such circumstances, later exams can become very long for students that have not mastered many topics. Using quizzes between exams can reduce the number of concepts students need to attempt and allow them a reasonable chance to master questions in the allotted time.

Quizzes can be given regularly throughout the semester on weeks where there is no exam. They can also be done in class or during office hours when the student is available. The quizzes can be organized to give students an opportunity to demonstrate mastery on just a single topic or multiple topics. One author utilizes "testing weeks," in which there is a particular week when students can use office hours to attempt a single topic at a time. They may try as many topics as they like, but only once for each topic during that week. Obviously there are multiple combinations and variations on each of these ideas, so each instructor can tailor retesting opportunities to each course, group of students, and teaching style.

3.3. Logistics of Writing Tests

Quizzes illuminate a key challenge to implementing MBT, which is the large number of exam questions to be generated and the large number of exams that must be written. Since each test includes learning objectives from previous exams, numerous versions of each of the questions are needed throughout the semester. In addition, the later exams will have an increasing number of questions on them, which makes the logistics of writing them increasingly complicated.

Generating exam questions can be made considerably easier by a thoughtful choice of questions. Often in MBT, questions are chosen so they can be easily modified while still testing the same concept. This can be a matter of changing numbers or modifying an equation from a polynomial to a trigonometric function. It can also be useful to generate a few classes of problems and modify the numbers within each class. In a related rates objective, for instance, one could identify the problem classes of sliding ladder, passing ships, and rising balloon. Changing the numbers within each problem allows one to generate a large number of versions with little effort.

Due to the many versions of each question that must be generated, it is helpful to store them in advance as a digital library. By generating all versions of each objective question (preferably before the semester starts) and organizing them by objective, writing the exams becomes essentially a matter of copy and paste. Therefore, once the initial work is done, writing exams and quizzes in an MBT setting is very quick and easy. One of the authors uses a \LaTeX database to randomly generate new exams, and an investigation of other implementation methods (such as the use of a learning management system) is an area of ongoing work.

3.4. Grading

Since MBT requires changing the way we approach in-class assessment, we provide suggestions and examples of grading schemes, how to deal with final grades, and methods for record-keeping during the semester. Although grade inflation may be a concern, in a 2-year study comparing MBT with traditional assessment, the grades earned in the courses were independent of the presence of mastery-based testing [10].

3.4.1. Exam Grading Schemes

It is essential to set an objective rubric for determining mastery. The important thing in this determination is to set the standards high enough to satisfy or exceed the desired expectations for a student who has passed the course. Keep in mind that this may vary between instructors and institutions. In general, the standards for mastery should be set high since it signifies that the student has demonstrated such deep understanding of the topic that they need not be tested further on it. It may be helpful to consider the question: *Will this student benefit from studying this topic again?*

Providing feedback for students on individual exam questions is important. It must be clear to the student whether they have or have not mastered a concept. Some choose to grade exam questions using a binary system: either a “mastered” or “not yet” performance on each objective. For others, a three-tiered system: often “mastered,” “progressing,” and “insufficient” levels is preferred. However, the addition of a “progressing” or similar level does not award students any credit in terms of their grades. They are used to provide encouragement and/or incentive for students to keep working towards mastery. This highlights an important point in terms of implementation in the classroom. It is crucial that students understand

Table 2. The chart above helps students estimate their current class average by dividing their total number of mastered concepts by the total number of concepts seen so far. It is filled out for a hypothetical student mastering three concepts on the first exam and four more concepts on the second

Date	Concepts Attempted	Total Number of Concepts Mastered	Total Number of Concepts Given	Estimated Average (%)
September 19, 2016	#1, 2, 3, 4	3	4	75
October 7, 2016	#5, 6, 7, 8, 9	7	9	78

that they only receive credit for “mastery” grades on an exam question, not for “progressing,” “insufficient” or other intermediate rankings.

3.4.2. Final Grade Calculation

The methods for calculating the final exam average generally fall into two categories: equal weighting and core concepts. The core concepts method regards certain concepts as more important than others. To obtain at least some minimum grade in a course, students are required to show mastery on all of the core topics. These core topics usually represent the necessary ideas for moving on to the next course in a sequence, or simply the material that is critical to the course content. Higher grades are determined using the topics that are not from the core. As an example, for a course that has 16 topics in which seven are core concepts, a student must master all seven core concepts to receive a 70%, and their grade above a 70% is determined by how many of the other nine topics they master.

Among those who do not use core concepts, some will remove 5% from 100% for each topic not mastered by the end of the course (resulting in a minimum average greater than zero). If one wishes to start from 0% and add credit, it will probably be necessary to weight the first several topics mastered by a student higher than the final topics so that the maximum score is 100%. This method has the added benefit of building a student’s confidence as they are beginning to master concepts. Mastery-based exams can easily be incorporated into a weighted average grading scheme by using a single large weight for the overall exam score (e.g., 70% exam, 20% homework, 10% participation).

3.4.3. Recording Grades

Recording exam grades in MBT is straightforward. One only needs to keep a “not mastered”/“mastered” record for each topic for each student.

Table 3. This chart helps students keep track of which concepts have been mastered and how many attempts have been made

Objective	1	2	3	...	14	15	16
Attempt #1							
Attempt #2							
Attempt #3							

It can be informative to also note on which exam the topic was mastered. Table 2 is a sample test grade sheet used for a first semester differential Calculus course taught using MBT with 16 topics. In this course, students were limited to three attempts of mastery, so students were given a chart to keep track of their progress (see Table 3). Having an efficient system of record keeping in mind prior to implementing MBT will help not only in calculating mid-semester grades, but also provide you with the ability to communicate effectively with students regarding their progress in the course.

MBT can be incorporated into an existing grading system by simply replacing the test score with the percentage of mastered topics. That is, it can fit into a typical course structure along with any quizzes, homework, or other class activities. Of course, one could also structure other class activities in a mastery-based format. Further implementation of mastery-based learning along these lines may be closer to other mastery-based and specs-based learning programs, for example in [1] or [11].

4. STUDENT REACTIONS

4.1. Survey Responses

We gave an end of semester survey (see Appendix A) to the students in each of the mastery-based assessment classes to see their overall impressions of MBT. These questions fell in three main categories: effectiveness, fairness, and study techniques. The data includes survey responses from 135 students at six different institutions. The data was compiled in aggregate and the quoted responses are not attached to a particular institution. Students were enrolled in mathematics classes that ranged from the Calculus sequence and finite mathematics to mid-level courses like statistics, linear algebra, discrete mathematics, and differential equations to math major courses like introduction to proofs, abstract algebra, and analysis. Universities represented in this data ranged from small (1500 to 2000 students) to mid-sized (5000 to 6500) private universities to regional public universities (9000 to 28,000 students).

Table 4. Student responses regarding effectiveness

Question	Strongly Agree (%)	Agree (%)	Disagree (%)	Strongly Disagree (%)	No Response (%)
Assessment tests knowledge	67.3	32.1	0	0.58	0
Helped learn material	49.1	46.8	2.3	1.2	0.6
Deepened my understanding	35.1	54.4	5.3	0.6	4.7
Prepared me for a variety of problems	22.8	61.4	9.9	1.2	4.7

In the effectiveness section, we asked students to decide how well they learned the material. Student responses are detailed in Table 4. Questions 1, 2, and 4 asked students if the assessment helped them understand the material more deeply. Students answered these questions overwhelmingly positive with between 81.8% and 100% of students agreeing or strongly agreeing. We also asked if students felt it prepared them for a variety of problems, and 84.2% of students agreed or strongly agreed.

In the section on fairness, we asked students if their results accurately reflected their knowledge, to which 81.3% assented. We also asked students whether they were anxious before exams, and 42.7% disagreed or strongly disagreed. (Note that students were asked whether they experienced *any* anxiety, not whether it was higher or lower than with other testing methods.) Detailed results can be found in Table 5.

Finally, we asked several questions on study methods. The majority of students in the classes said they spent 3–5 hours per week studying (54%) with no students reporting spending more than 11 hours outside of class. The two largest areas students look to for studying were review materials (56%) and their class notes (63%). We also asked whether students spent time memorizing solutions to past exams. This is a common concern when deciding to depart from traditional grading schemes, yet 73.1% of students did not feel they spent time memorizing past solutions.

We also compiled student feedback from surveys and end of the year course evaluations. The feedback was overwhelmingly positive. Students appreciated that MBT allowed them to practice and reevaluate material. One student in a Real Analysis course wrote:

> I like the way that [mastery-based exams] ensure you keep working at a topic until you know how to do it, rather than just deciding not to care because you won't see it again after the test. In Analysis a lot of the content builds off of each other as well, so I thought this was a good and fair way to make sure we were mastering old content while working on new content at the same time.

Table 5. Student responses regarding fairness

Question	Strongly Agree (%)	Agree (%)	Disagree (%)	Strongly Disagree (%)	No Response (%)
Accurately Reflected Knowledge	32.2	49.1	15.2	2.3	1.2
I felt anxious before exams.	22.8	31.0	31.0	11.7	3.5

Other students echoed this response saying that MBT allowed them the time to “actually grasp the concept of the topics” and “keep up on old concepts from earlier in the semester.” In general, most students said that this method of testing reduced stress when taking exams. One student remarked:

> I feel that mastery-based testing alleviated a lot of the pressure brought upon me as a college student from exams normally, and also re-enforced learning concepts I had difficulty mastering in a different way, helping to solidify knowledge of concepts that I did not understand as well.

Students also said that MBT helped them identify and focus on concepts. One student stated, “I like the mastery-based testing, it forced me to learn the concepts I did not understand initially.” Another said they enjoyed this assessment “because math revolves around practice.” All of these comments support a growth mindset of learning.

Students also mentioned that it was important for instructors implementing MBT to have clear grading systems and give their students enough opportunities to retest. One student said:

> I am an advocate for the mastery-based assessment system. However, this advocacy is contingent upon the instructor providing sufficient opportunities to succeed.

Some students complained that the increased size of the exams could feel “overwhelming” especially if they did not stay on top of mastering concepts. Note that this concern can be addressed by additional reassessment opportunities, as described in Section 3. Despite these reservations, the majority of student responses were in support of MBT.

4.2. Student Buy-In

It is important to make sure students completely understand the MBT grading system, as it is likely to be substantially different than other grading systems they have experienced. Giving a pitch or advertisement for MBT at the beginning of the semester is essential in order for students to both understand and buy-into the grading system which is probably unfamiliar. In the initial explanation of MBT to the class, it helps to emphasize that exams will be graded much harder on an "all-or-nothing" basis, but they will have several attempts for each topic. Encouraging and reminding students to take advantage of retesting opportunities will also help class morale and motivation. It is also beneficial to emphasize the pedagogical reasons behind this non-traditional assessment format.

Often students provided us with informal feedback about MBT during the semester. Usually the biggest adjustment has been the lack of partial credit. For example, if a student misses one part of a multi-part question, they may respond with frustration that the other correct answers do not "count." Reiterating the importance of mastery and emphasizing all future reassessment opportunities can address these concerns.

We have also observed that most students were able to perform at a higher level compared with traditional testing, and that some students who would have failed (due to being under-prepared coming into the course or not passing the first test) worked really hard and eventually passed the class.

Openly discussing the intended goals of MBT throughout the semester (as well as reiterating the means by which grades are determined) can help students accept the new method.

5. CONCLUSION

In this article, we have presented MBT as an alternative method of assessment in math courses. The fundamental attributes of MBT are clear course concepts, credit only for mastery, and multiple attempts with complete forgiveness. Under these parameters, course incentive structures are shifted dramatically. Students are encouraged to develop a deep understanding of the content goals, which can often be circumvented in traditional, partial-credit testing regimes.

Repeated testing of course materials with complete forgiveness for past performance motivates students to continue to revisit misunderstood concepts, thereby supporting the development of growth mindset in students. The direct alignment of course concepts with awarded grades puts the student and instructor on the same page in terms of course concepts and formulation of study habits. This in turn improves the quality and efficiency of instruction during office hours, and motivates students to more carefully consider feedback on their exams.

Implementations of MBT can be varied widely to fit different course content, teaching styles, and class sizes. We have used MBT in calculus, differential equations, introductory analysis courses, and others with largely positive experiences for students and teachers.

As instructors, we wish to impart deep and lasting mathematical skills to our students. MBT helps us guide students up to the “peaks” of a given mathematical topic by encouraging everyone to pursue thorough understanding. From there we hope they will have a better view of the broader mathematical landscape and of their own capacity to grow intellectually through persistence.

ACKNOWLEDGEMENTS

We wish to thank Dr. George McNulty at the University of South Carolina for introducing us to this style of assessment. We would also like to note that this collaboration began during the 2014 Project NExT professional development program of the Mathematical Association of America. We are grateful to the organizers of Project NExT for providing an environment where this sort of project can germinate, and to our home institutions for supporting our travel to the Project NExT workshop.

APPENDIX

ASSESSMENT SURVEY QUESTIONS

The following is the exit survey used to collect the student feedback presented in Section 4.

1. The assessments in this course test our understanding of key concepts.
2. Studying for the exams in this course helped me to learn the material.
3. The results of my in-class assessments accurately reflect my knowledge.
4. The in-class assessments deepened my understanding of the ideas in this course.
5. I was anxious before the exams in this course.
6. I relied mostly on memorizing solutions to earlier problems to prepare for in-class assessments.
7. I feel prepared to approach a wide range of problems in [Course Name].
    a. strongly agree
    b. agree
    c. disagree
    d. strongly disagree
    e. no response

8. How many hours per week did you spend on this course outside of class time?
   a. 0-2 hours
   b. 3-5 hours
   c. 6-8 hours
   d. 9-11 hours
   e. 12-14 hours
   f. more than 14 hours

REFERENCES

1. Beatty, I. D. 2013. Standards-based grading in introductory university physics. *Journal of the Scholarship of Teaching and Learning*. 13(2): 1–22.
2. Brilleslyper, M., M. Ghrist, T. Holcomb, B. Schaubroeck, B. Warner, and S. Williams. 2012. What's the point? The benefits of grading without points. *PRIMUS*. 22(5): 411–427.
3. Butler, R. and M. Nisan. 1986. Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal of Educational Psychology*. 78(3): 210–216.
4. Cherepinsky, V. 2011. Self-reflective grading: Getting students to learn from their mistakes. *PRIMUS*. 21(3): 294–301.
5. DeLong, M. and D. Winter. 2001. An objective approach to student-centered instruction. *PRIMUS*. 11(1): 27–52.
6. DeLong, M., D. Winter, and C. A. Yackel. 2005. Student learning objectives and mathematics teaching. *PRIMUS*. 15(3): 226–258.
7. Dweck, C. S. 2007. Boosting achievement with messages that motivate. *Education Canada*. 47(2): 6–10.
8. Dweck, C. S. 2007. Is math a gift? Beliefs that put females at risk. In S. J. Ceci & W. M. Williams (Eds.), *Why Aren't More Women in Science?: Top Researchers Debate the Evidence* (pp. 47-55). Washington, DC: American Psychological Association.
9. Dweck, C. S. 2013. *Self-Theories: Their Role in Motivation, Personality, and Development*. Philadelphia, PA: Psychology Press.
10. Harsy, A. and A. Hoofnagle (Armstrong). 2018. Comparing mastery-based testing with traditional testing in Calculus II. In revision.
11. Nilson, L. 2015. *Specifications Grading: Restoring Rigor, Motivating Students, and Saving Faculty Time*. Sterling, VA: Stylus Publishing, LLC.
12. Schneider, J. and E. Hutt. 2014 Making the grade: A history of the A–F marking scheme. *Journal of Curriculum Studies*. 46(2): 201–224.
13. Studman, C. J. 1984. A method of applying mastery learning to moderately large classes. *International Journal of Mathematical Education in Science and Technology*. 15(1): 95–100.
14. Williams, K. 2017. Specifications-based grading in an introduction to proofs course. *PRIMUS*. 28(2): 128–142.

BIOGRAPHICAL SKETCHES

Dr. Collins is an assistant professor at the University of Mary Washington and a member of the Gold ‘14 Project NExT cohort. He graduated from North Carolina State University with a Ph.D. in Mathematics. He spent four years at Colorado State University in a postdoctoral fellowship researching error estimation techniques for numerical methods. His research focus is on computational method for differential equations, as well as numerical algebraic geometry. He is also interested in active learning techniques in the classroom and undergraduate research.

Dr. Harsy is an assistant professor of mathematics at Lewis University. She is a Gold ‘14 Project NExT Fellow and in addition to being a part of the Mastery-based Testing Research group, is the faculty advisor for Lewis’s Math Club and is part of the leadership team for the Southwest Chicago Suburbs Math Teachers Circle Dr. Harsy has also been active in outreach to local K-12 schools with a focus on encouraging girls to pursue mathematics and computer science. In particular, she has run several summer math camps and has led several Girls Create with Technology sessions for middle schoolers. Dr. Harsy enjoys mentoring undergraduate research and has worked with students on projects involving math education, linear algebra, graph theory, and data science.

Dr. Hart is a postdoctoral researcher in the Higuchi Biosciences Center at the University of Kansas (KU) and is a Gold ‘14 Project NExT Fellow. He is actively involved with curriculum development, advising, and training for math and science students in underrepresented groups at KU. He has organized conferences and workshops, supported by the Initiative for Maximizing Student Development (IMSD) and Post-Baccalaureate Research Education Program (PREP) at KU, that provide students with training in math, quantitive reasoning, and programming. He has developed and taught interdisciplinary capstone-style courses that introduce undergraduate and graduate STEM students to using advanced mathematics in scientific research. For his courses, he has also formulated measurement and adaptive instruction techniques that respond to students at a cognitive developmental level. He was previously an Assistant Professor (Research) at Wayne State University, where he started using Mastery Based Testing (MBT). At the moment, he does not use MBT since he is teaching capstone-style courses, but he plans to use it again in the future. Dr. Hart enjoys doing research in harmonic and Fourier analysis, math education, as well as several subjects in applied math and math biology.

Dr. Haymaker has been an assistant professor at Villanova University since 2014. She received her PhD from the University of Nebraska-Lincoln. Dr. Haymakers mathematical research is in coding theory and applied combinatorial mathematics. She is a Project NExT Fellow (Gold 2014) and a member of the AMS, MAA, and AWM.

Dr. Hoofnagle is an assistant professor of mathematics at Wittenberg University and is a Gold '14 Project NExT fellow. Outside of her teaching duties, she is involved in curriculum development, statistics education, and outreach opportunities in the surrounding Springfield, OH community. She is also the advisor for the Wittenberg Math Club and the Wittenberg chapter of Pi Mu Epsilon. She is currently doing research involving mastery-based testing in introductory statistics, as well as leading undergraduate research projects in pure mathematics

Dr. Janssen is an assistant professor of mathematics at Dordt College, and a member of the Gold '14 Project NExT cohort. He has taught courses throughout the math major using mastery-based assessments. His research interests lie at the intersection of commutative algebra, algebraic geometry, and discrete mathematics, and he has enjoyed sharing those interests with undergraduate students.

Dr. Kelly is an assistant professor of mathematics at Christopher Newport University. She is a Gold '14 Project NExT Fellow and part of the Mastery-based Testing Research group. Dr. Kelly is the organizer of a regional high school mathematics contest, which involves over 100 high school students each year. Additionally, Dr. Kelly is active in research, working in orthogonal polynomials and special functions in addition to mentoring undergraduate research students.

Dr. Mohr is an assistant professor of mathematics at Nebraska Wesleyan University and a 2014 MAA Project NExT fellow. He is a co-PI on an NSF S-STEM grant supporting underrepresented students to earn STEM degrees and was the PI on an NSF EPSCoR grant for under graduate research experiences. He speaks frequently on topics in combinatorics, formative assessment, and active learning. Most recently, he co-organized a MathFest minicourse on mastery-based assessment.

Dr. OShaughnessy holds a PhD in mathematics from the National University of Ireland, Galway. She is a Gold '14 MAA Project NExT fellow and has a love of teaching mathematics. Her research interests in mathematics are in algebraic coding theory. Her interests in math pedagogy include formative assessment, in particular using mastery based testing in the classroom and using technology to support teaching and learning. She is currently an associate professor of mathematics at Shenandoah University.