A Novel Approach to Mastery-Based Assessment in Sophomore-Level Mechanics Courses

Prof. Keith D. Hjelmstad, Arizona State University

Keith D. Hjelmstad is President’s Professor of Civil Engineering in the School of Sustainable Engineering and the Built Environment at Arizona State University.

Amie Baisley, University of Florida

Amie Baisley is a lecturer in the Department of Engineering Education at the University of Florida. Her teaching and research interests are centered around the sophomore level courses that engineering students take and how changes in those courses can impact student learning and retention.
A Novel Approach to Mastery-Based Assessment in Sophomore-Level Mechanics Courses

Abstract

The Mechanics Project is a reimagination of the foundational mechanics courses that engineering students generally take in their sophomore year. The courses associated with the project (statics, dynamics, and deformable solids) were converted to a student-centered engaged learning environment with students spending most of their class time in a flipped recitation environment. The pedagogical transformation was complemented with the implementation of a novel assessment system based upon redundant demonstration of mastery of the course objectives. The assessment system requires two key things: (1) the definition of measurable course objectives that capture the fundamental concept strands—the DNA—of the course and (2) frequent assessment that incorporates the redundancy of demonstration required to confidently conclude mastery. The process of developing this system had a significant impact on the nature of the courses and informed the topical content and development of course materials. The main motivation for moving to a mastery-based grading system was to change the way students think about and experience assessment. The frequency of assessment reduces some of the stress of the testing environment and the redundancy promotes a spiral learning approach that helps students connect the components of the problem-solving process. A simple grading rubric and feedback system was created to provide timely, meaningful, and detailed feedback on their progress with each specific learning objective. This paper describes the learning objectives, the feedback the students receive following each assessment, and how mastery is assessed in each course.

Introduction

Grades have long served as the currency of quality of work in universities. The main purpose of the final grade in a course is to certify that a student has sufficient grasp on the concepts of the course. The certification is useful in deciding if a student is ready for more advanced study (i.e., satisfactory prerequisite knowledge) or worthy of a summary certification like a degree. Finer grade distinctions (beyond pass/fail) serve the purpose of distinguishing various levels of performance, from merely satisfactory to excellent. Such distinctions are then useful, for example, for deciding among two equally credentialed candidates for a job or a position in a graduate program.

Like any good currency, standardization is important. If one person is to decide something about a student based upon a grade another gave, that person would need to understand what that grade means. Hence, over time, grades have standardized to fixed scales like A=4.0, B=3.0, C=2.0, etc. While standard grade systems have been around for several hundred years, academia is still unable to consistently and meaningfully interpret what a grade represents about what a student has learned in a course [1]. Interpretations vary from institution to institution and from instructor to instructor within an institution. That said, grades remain a part of the fabric of formal education and strongly influence how we organize and execute teaching and learning.

Because grades are used for making decisions about what knowledge a student might possess and what opportunities he or she might qualify for, the stakes associated with earning those grades is high. And because the stakes are perceived as being high, grades have emerged as a dominating factor in the motivation of students. They also play a large role in forming a student’s
self-concept and professional identity. Grades function as an extrinsic motivator for students, and research has shown that extrinsic motivators can be ineffective, and sometimes counterproductive, especially for high cognitive tasks like learning [2].

A departure from standard grading

At a large R1 university in the southwest, we have redesigned the sophomore mechanics courses, through an effort called The Mechanics Project [3]. The goal of the project is to improve the learning outcomes of students in their sophomore year and to develop self-driven learners. The focus of the project is on the courses Statics, Dynamics, and Deformable Solids, which are required courses for students in civil, mechanical, and aerospace engineering curricula.

As part of this project, a mastery-based grading system (similar to standards-based grading or competency-based grading) was implemented in the three courses. There were two primary aims. First, we wanted to better evaluate each student’s learning progress so that we could provide more formative feedback and directed support for their learning. Second, we wanted to disrupt the negative influence of standard grading practices on student motivation.

Mastery-based grading is not new. In fact, Bloom outlined guidelines and instructions for mastery learning over 40 years ago [4]. Variants of mastery-based grading has been used in many different areas of engineering with different degrees of sophistication in implementation [5] – [8]. The implementation of the mastery-based grading system in The Mechanics Project involves all of the Bloom guidelines, but the specific implementation details are unique. Ultimately, the mastery-based grading system became the focal point of the pedagogical design of the three courses, impacting almost every aspect of the instructional strategy.

The primary driving factor behind the decision to introduce mastery-based grading is the opportunity it provides to change the role of assessment in the learning environment. We wanted our assessments to become a learning tool and the feedback to become the guide for fixing errors [9], [10]. Feedback without the opportunity to improve is not a strong strategy. Students will often ignore written feedback if the assessment is summative, electing to focus primarily on the score or grade. To generate a feedback cycle, we define mastery as the redundant demonstration of the ability to accomplish an objective. The redundancy imperative requires multiplicity of opportunity and that provides a route to improve in the feedback cycle. Hence, the mastery-based grading system requires frequent assessment.

An additional advantage of the mastery-based system, spawned by the frequency of assessment, is a reduction in test anxiety [11]. No single test makes up a large portion of the grade [10]. Students are encouraged to master all aspects of the course and not focus on only their strengths. The feedback after the assessment is formative and helps the student understand their specific mistakes rather than receiving a bulk score. Finally, the mastery system allows students to have time to develop their mastery in the subject over the course of the semester. Maturation of concepts over a semester is a natural growth path for all learners and the developmental process is often stunted in the standard midterm/final exam environment [4], [12].

The development of a mastery-based system in the mechanics courses created the need to have measurable course objectives unique to each course with plenty of assessment opportunities to measure a student’s mastery. Mastery is based on repeatedly showing successful completion of a task, so the courses had to be redesigned to allow that [10]. The course redesign started with modifying each of the three course schedules to meet twice a week for 75-minute recitation
periods and once a week for a 50-minute period. During the 50-minute period the entire course meets in a large lecture hall where they are given an assessment every other week. The assessment problems follow the same solution process as all homework and recitation problems, and the students get into a routine of taking an assessment at the end of every two-week module. Each assessment consists of one problem relevant to the current module topic of the course. The students are given the entire period to work through the problem and are asked to organize their solutions around the course mastery objectives. Assessing every two weeks results in seven assessment opportunities throughout the semester. In addition to these seven assessments, the final exam for each of the courses are designed as four additional assessment problems. Each assessment problem is equivalent to one mastery opportunity, which creates a total of 11 assessment opportunities during a given semester.

Mastery objectives

The mastery system requires creating redundant opportunities for a student to show they know certain aspects of a course. As a result, a set of general objectives in each course form the DNA of the course. These objectives represent the key strands to solving mechanics problems that are present in most, if not all, problems in the given branch of mechanics. Students should be developing these in a traditional learning environment. Nevertheless, the creation of an objective for that strand requires it to be explicitly written down. What gets measured gets improved. The objectives for each of the three courses are given in Table 1.

Table 1. Course Mastery Objectives. This table gives the short name for each of the mastery objectives used in the course. Objectives can be grouped into subobjectives with common themes.

<table>
<thead>
<tr>
<th>Statics</th>
<th>Dynamics</th>
<th>Deformable Solids</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. Solution strategy</td>
<td>A.2. Initial conditions</td>
<td>A.2. Boundary conditions</td>
</tr>
<tr>
<td>D. Free body diagrams</td>
<td>B. Describe position vector</td>
<td>C. Free body diagram</td>
</tr>
<tr>
<td>E. Force equilibrium</td>
<td>C. Compute velocity and accel.</td>
<td>E.1. Force equilibrium</td>
</tr>
<tr>
<td>F. Moment equilibrium</td>
<td>D. Free body diagrams</td>
<td>E.2. Moment equilibrium</td>
</tr>
<tr>
<td>G. Distributed effects</td>
<td>E.1. Balance linear momentum</td>
<td>F. Strain-displ. relationships</td>
</tr>
<tr>
<td>I. Internal forces</td>
<td>F.1. Vector algebra and calculus</td>
<td>G.2. Properties of areas</td>
</tr>
<tr>
<td>J. Units &amp; conversions</td>
<td>F.2. Integrate over spatial domain</td>
<td>H.1. Derive differential eqn.</td>
</tr>
<tr>
<td>K. Systems</td>
<td>G. Conservation of momentum</td>
<td></td>
</tr>
<tr>
<td>L. Notation</td>
<td>H.1. Classical soln. to diff. eq.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H.2. Natural freq. and modes</td>
<td>H.2. Implement BCs</td>
</tr>
<tr>
<td></td>
<td>J. Compute dynamic response</td>
<td></td>
</tr>
<tr>
<td></td>
<td>J.1. Stress formulas</td>
<td></td>
</tr>
<tr>
<td></td>
<td>K. Compute energy and work</td>
<td>K. Compute relevant response</td>
</tr>
<tr>
<td></td>
<td>L. Apply work/energy principles</td>
<td></td>
</tr>
</tbody>
</table>

Every objective requires an equation, drawing, description, or some type of mathematical manipulation. In other words, the student must be able to do something associated with each
mastery objective. Some of the objectives involve more work than others, but each one is an important enough aspect of the solution process to be classified as its own objective. Some objectives warrant the creation of sub-objectives, as in objectives A.1, A.2, A.3 (problem setup) or E.1, E.2 (equilibrium) in Table 2 to get a desirable granularity. Each unique set of objectives frames a solution process for problems in that course. In Statics, almost all the objectives can be assessed for every statics problem. In the other two courses, not all objectives are activated in every single problem. For example, in Dynamics the necessary objectives for the problem depend on the solution method, whether through energy methods or balance of momentum. The students are asked to organize their solutions around the mastery objectives. In doing so, they become familiar with the necessary information required for each one.

**What is mastery?**

The goal of each course is to master all the course objectives which thereby demonstrates that the student is competent in all aspects of the course. Of course, neither time nor student motivation allows every student to master every objective during a semester, so a threshold for mastery must be set. There were many factors considered for what this threshold should be, the most important consideration being what a student needs to do to convince the instructor that they know what they are doing.

Our definition of mastery is redundant demonstration of the key concepts for different problems done at different times. The exact number of repeated demonstrations varies between the courses based on the number of times each objective is assessed in a typical semester. As mentioned, Statics problems assess almost all objectives on every assessment problem for a total of 11 opportunities to respond to each objective. In Dynamics and Deformable Solids, the problems will include many of the foundational objectives but some of the later objectives only show up on three to five of the assessment problems. These problem-type differences set the mastery level of an objective at five correct showings in Statics and three correct showings in Dynamics and Deformable Solids. A showing or demonstration is a correct response to the objective on a single assessment. The mastery progress of each student is tracked throughout the entire course. The general idea of mastery demonstration is shown in Fig. 1.

![Mastery Objectives](image)

**Fig. 1. Demonstration of Mastery.** Each module assessment problem has some (usually not all) of the mastery objectives available. Mastery is a redundant demonstration of each objective over multiple problems. This figure shows six of ten objectives are available in Problem 3. The student has demonstrated mastery of Objective 6 with complete and correct solutions to Problems 2, 4, and 5.
The style of mastery assessment problems is similar to the problems that the students solve in recitation and homework. In particular, the assessment contains a single problem that activates some (but usually not all) of the mastery objectives. A typical assessment problem in *Dynamics* is shown in Fig. 2. Note that most of the objectives are available in most of the problems and there are multiple assessments. The student builds a case for mastery, objective by objective, across multiple mastery assessments.

![Fig. 2. Typical mastery assessment format.](image)

Since there are more than enough opportunities available to each student to achieve mastery in the course, the weight and stress surrounding each assessment has decreased. The homework problems and practice exam problems are also solved in the same objective format, so the assessments model very closely what the students do every day in the course. This grading system has departed from the idea that a bad grade on a bad day will bring down your entire course grade. Having a bad day is acceptable, and if a student has a bad assessment performance then the only consequence is that it does not contribute to their mastery score. They missed out on an opportunity to increase their mastery score, but their mastery score will not decrease because of a bad or missed assessment. The continuous assessment cycle means there will always be more opportunities to demonstrate mastery in the course (up until the final). The students are encouraged to take advantage of all available mastery opportunities, but the grading system no longer penalizes a student for misunderstanding a question.

Achieving mastery in the courses has been described as repeated correct demonstrations of work for an objective. A correct demonstration is based on the grading rubric that was developed to determine the level of correctness assigned to a student response. The goal of the new evaluation system was to provide more meaningful feedback about each student’s learning in the course, so a binary grading rubric did not suit this. The different levels created for the grading rubric differentiate between varying degrees of mistakes that students make (from a mathematical error to conceptual error) but is concise enough that multiple graders are able to arrive at the same grade if individually grading the same assessment. The grading rubric is provided in Table 2.

Each student’s work is assessed with the rubric in Table 2 for every available objective on each module assessment. Mastery is evaluated on an index that goes from 0 to 12. We acknowledge a mastered objective as being an index greater than 10 on that scale. The index is the product of the assessment factor (see the third column of Table 2) with the weight assigned to the problem.
Problems with greater difficulty are assigned a higher weight, but most objectives have a weight in the neighborhood of 4. What that implies is that you need roughly three ‘Complete and correct’ showings on a given objective to prove mastery. The weights for problems in one semester in Deformable Solids are shown in Table 3. You can see that the objectives are not always available and that the weights vary. You can also see that most objectives far exceeded the 12 needed to gain mastery. Providing more opportunity than is required is an important part of the system and allows students to maintain a positive outlook throughout the semester.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Complete and correct</td>
<td>1.00</td>
</tr>
<tr>
<td>b</td>
<td>Significant understanding, but with minor execution</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>errors</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>Some understanding, minor conceptual errors</td>
<td>0.25</td>
</tr>
<tr>
<td>d</td>
<td>Major conceptual errors</td>
<td>0.00</td>
</tr>
<tr>
<td>e</td>
<td>No evidence shown</td>
<td>0.00</td>
</tr>
<tr>
<td>g</td>
<td>Response was largely irrelevant to this problem</td>
<td>0.00</td>
</tr>
</tbody>
</table>

In Statics, the weights tend to be lower to force more demonstrations to get to mastery. It is theoretically possible to accumulate enough showings of ‘Significant understanding, but with minor execution errors’ to reach mastery. However, these performances only have a factor of 0.5, so it is very difficult to make a demonstration of mastery with only this level of performance. It is not possible to get to mastery with ‘c’ level performances. The last three rubric items do not add anything to a student’s mastery score but provide differentiation in the feedback given for each objective. At the end of the semester, if enough mastery points have been achieved equivalent to three or five ‘Complete and correct’ showings then the student has mastered that objective.

**Feedback**

There are no grades given in these three courses anymore, apart from the final grade, but after every module assessment the students receive two forms of detailed feedback about their performance. The first form of feedback is a Mastery Assessment (MA) letter that provides a summary of the mastery objectives and how the student did on each objective, based upon the rubric in Table 2. The letter includes the student’s progress on each objective over the semester up to that point and gives an overall assessment of progress on each objective. An excerpt from such a letter is provided in Fig. 3. The student can use this with the posted solution to reflect on their mistakes or successes on each assessment problem.

The second form of feedback is called the dashboard. An example dashboard for a specific student is shown in Fig. 4. The dashboard is a visual document that captures the student’s performance in all aspects of the course. The dashboard contains visual elements for the computing projects, homework, and other required coursework. The mastery bar chart is the element on the dashboard that shows how a student’s mastery is developing over a semester. The chart has a bar for each objective in the course. In each assessment, the activated objectives have the potential to increase the length of the mastery bars. The amount that each objective bar
increases depends on the weight given to that problem and the factor associated with the rubric score.

Table 3. Weighting of objectives. Each problem for each module assessment is weighted in accord with the difficulty of the problem, relative to that objective. This table shows the actual weights for the problems in Deformable Solids for one semester. The total is the column sum and represents how much opportunity was available for that objective.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Mastery Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A.1</td>
</tr>
<tr>
<td>MA 1</td>
<td>4</td>
</tr>
<tr>
<td>MA 2</td>
<td>4</td>
</tr>
<tr>
<td>MA 3</td>
<td>2</td>
</tr>
<tr>
<td>MA 4</td>
<td>4</td>
</tr>
<tr>
<td>MA 5</td>
<td>4</td>
</tr>
<tr>
<td>MA 6</td>
<td>2</td>
</tr>
<tr>
<td>MA 7</td>
<td>4</td>
</tr>
<tr>
<td>Final 1</td>
<td>2</td>
</tr>
<tr>
<td>Final 2</td>
<td>4</td>
</tr>
<tr>
<td>Final 3</td>
<td>2</td>
</tr>
<tr>
<td>Final 4</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>32</td>
</tr>
</tbody>
</table>

C. Free-body Diagrams. Construct accurate and complete Free-Body Diagram(s) (FBD). Treat distributed forces (e.g., pressure and weight) and point effects (e.g., reactions and point loads) appropriately. Designate actions on FBDs as vectors (use vector notation). Implement known constraints on forces (e.g., normal forces have a known direction, etc.

MA1 : (b) Significant understanding, but with execution errors
MA2 : (.) This item was not tested in this task
MA3 : (.) This item was not tested in this task
MA4 : (c) Some understanding, but with minor logical errors
MA5 : (b) Significant understanding, but with execution errors
Overall mastery of this objective (so far) : DEVELOPING

F. Strain-displacement relationships. Characterize the description of motion. Derive or use the relationships that connect strain to the derivative of the displacement field through the kinematic hypothesis.

MA1 : (a) Understanding complete and correct
MA2 : (b) Significant understanding, but with execution errors
MA3 : (a) Understanding complete and correct
MA4 : (.) This item was not tested in this task
MA5 : (a) Understanding complete and correct
Overall mastery of this objective (so far) : GOOD

Fig. 3. Excerpt from feedback letter. After each module assessment, each student gets a feedback letter that includes a description of each mastery objective along with the outcome from that assessment (and all previous ones). The overall assessment corresponds with the bar on the dashboard. This figure shows two of the objectives for a student after MAS5 in Deformable Solids.

Figure 5 shows an example of how the mastery bars changed throughout a semester for a specific student. Note that the mastery bars can only increase in length and will never decrease. There is no penalty associated with doing something wrong; it is just a lost opportunity to make a demonstration of mastery. Mastery of a certain objective is achieved when that bar reaches the green line on the chart, which corresponds to ‘good’ mastery. We associate ‘strong’ mastery with the cyan line at the top of the chart.
The feedback for the course includes the dashboard, which is a complete representation of all work done by the student up to the date of the dashboard. This dashboard shows the final result for a student in Deformable Solids.

The students use the dashboard to track their progress in the course. It provides very detailed feedback about the features of the course they are doing well in and the areas that they need to work on. In Fig. 5 you can see that the example student has achieved ‘good’ mastery of objectives C, E.1, and G.2 and ‘strong’ mastery of objectives A.1, F, and I after Mastery Assessment 6. It is not possible for the student to improve on objectives with ‘strong’ mastery, but they might still need to do them to get to other objectives on subsequent exams. While some of the objectives are mastered, many of the objectives are still below mastery. Some well below. The student knows which areas to focus on. After the final, the example student has mastered 10 of the 15 objectives in the course, with two or three fairly close to mastery (although we do not count those in the final grade assessment).

The students are provided with an MA letter and an updated dashboard following every assessment, so the feedback is frequent and formative. This is the only feedback that we provide on their assessment performance, unless the student seeks individual audience with the instructor. No grades or percentages are given throughout the semester. The possibility of mastering all objectives remains feasible practically to the end of the course because the final exam provides four additional opportunities in addition to the seven associated with the module assessments.

Note that is not possible to have mastered any objectives after the first exam. This fact helps to change the conversations about achievement in the class. Although, perhaps out of habit, students will still sometimes ask what their grade is in the course and if they will pass. The instructors explain that it is not possible to assign a grade since the student’s performance on the remaining assessments is unknown. The conversation can then be refocused to the areas of the mastery bar.
chart that are low and try to resolve the issues the student is having with those objectives. These conversations are easy to tailor to each individual student by quickly reviewing their dashboard.

![Chart example](chart.png)

Fig. 5. Example of mastery bar chart progression in Deformable Solids. The evolution of the mastery bar chart for a student in Deformable Solids. The black dots are the class average for that objective.

We have written MATLAB programs to generate the dashboards and the MA letters automatically, reading from a master spreadsheet where the scores are recorded. We have adjusted the programs over time to reflect back to the student as much information as possible. The development of the programs is aimed at reducing the effort required of the instructor to provide the feedback. The programs could be easily adapted to other courses and are available from the authors upon request.

**Translating mastery into a grade**

There is a time, at the end of the semester, when a final grade must be assigned. The final grade comes directly from the number of objectives that have been mastered. Even the mastery bar chart can be broken down into numbers and correlated to a percentage. Although, the mastery portion is only one part of the course grade for these courses. The other components of the course grade include computing projects and engagement. The grade ‘formula’ for the course is

\[
\text{grade} = (\text{mastery} + \text{projects}) \times \text{engagement}
\]

Of course, this equation does not result in a grade, but rather an index that we use to sort performances prior to giving the final course grade. In fact, the S-index that is reported on the dashboard represents this calculation based upon the work completed at that time. The portion of the S-index that comes from mastery is the M-index, which is also given on the dashboard. The M-index is basically a weighted sum of the mastery bars and is designed to come out roughly
equal to the number of objectives mastered at the end of the semester. It does not represent that early in the semester but provides relative measure for students to gage their progress relative to their peers in mastery alone.

No matter what angle you take, it is not possible to determine course grade from mastery alone. But that is true for any course with multiple components contributing to the final grade. Since the \( S\)-index is based upon the performance of other students, it is not possible for a student to do the computation. We assert that the \( S\)-index does not provide the absolute grade, but no student can earn a higher grade than another student with a higher \( S\)-index. That allows us to look at absolute performance in the assignment of final grades. For example, we consider the actual number of objectives mastered in the final grade (and not just the estimate provided by the \( M\)-index).

We have elected not to provide the details of grade calculations to students because students in possession of a grading formula are very likely to form their decision processes around it. What students do with grading formulas is often at odds with good learning practices. When a student is concerned about grades, we sometimes publish the grade outcomes of a few previous semesters so that they can see, for example, what the failure rate tends to be or how many students tend to get ‘A’ grades. Most student interest in grades has less to do with their learning and more to do with the tyranny of how grades form self-concept. We articulate the principle that the final grade comes from the number of objectives mastered and to achieve the highest grade for the mastery portion they must master as many objectives as they can (hopefully all of them). With that simple principle, the most obvious actions the student can take are those that increase mastery. Our focus is on measuring their learning and successes in the course, not assigning them a score. The bar charts are very effective at providing students a clear visual and motivation of the goal they are striving for.

Another benefit of the mastery system is the encouragement for students to focus on their areas of weakness and less about achieving a certain score. Once mastery is achieved in an objective, they do not receive extra credit for continuing to show that objective (they may still have to complete it to be successful on the assessment), but they can focus their efforts on the objectives with low bars. This is a natural progression towards the end of the semester as some of the objectives are getting mastered, the students will start using their remaining assessment opportunities to grow those low bars.

After working in this grading system for several years, the instructors have a good idea of how the various levels of mastery correlate to the traditional letter grades. We have built in a safeguard for mastery in the grading system by asserting that mastery of fewer than five objectives will result in automatic failure of the course, independent of performance of computing projects and independent of the level of engagement. There is a similar safeguard on the computing projects. Each semester, a few students fail to meet these levels of performance. It is possible to fail the course due to a combined low performance in all areas, but the most common incidences of failure are due to not meeting the minimal performance levels of either mastery or projects. If a student masters more than minimal number of objectives, then they have enough knowledge in the course to move on (but still need to make up the gaps to assure success in post-requisite courses).
Results from mastery-based grading

In *The Mechanics Project* we changed both pedagogy and content. The extent of the change makes it difficult to isolate the effects of various changes on performance outcomes. While it might be possible to get final grade outcomes for students prior to the implementation of *The Mechanics Project*, it would not be possible to determine what factors led to those grades. Even if one looked at the results of a common test like the FE exam, the measured outcome would be biased toward courses that delivered traditional content because the exams are tuned to certain types of problems and are written by people who teach or learned in that environment.

What we can do is look at the performance of the approach over an extended period of use. For this purpose, we will consider the Dynamics course, where we have 12 semesters of data on learning outcomes associated with 829 students who took and passed the course. Figure 6 shows the objective availability and mastery performance over the period from the SP14 semester through the FA19 semester.

The cyan line on the objective availability chart shows how well we achieve the target of providing roughly twice the number of opportunities needed to master the objective. If the students were to earn complete and correct scores on that much opportunity, then they would master the objective. The open circles show the average availability over the 12 semesters. It is evident that the opportunity varies (sometimes quite a bit) from one semester to another. One of the most highly variable objectives is *J—Compute Relevant Dynamical Quantities*. This objective is really about answering the specific thing asked in the problem statement. Over time we have learned how to write problem statements better to incorporate this objective.

![Figure 6: Results of mastery-based assessment](image)

It is also evident that some of the objectives are sometimes under-offered. In those cases, it is usually that the mastery objective shows up late in the course (e.g., H.1 and H.2) or it is simply...
difficult to devote an adequate number of problems to the objective (e.g., G). We do not worry much about under-offering objective \(G—\text{Conservation of momentum}\), because it is the subject of two of the six computing projects in the course. Observe also that balance of angular momentum is sampled less than balance of linear momentum. This result is due to the first two modules being associated with particle mechanics (in which balance of angular momentum is not used) and the increased emphasis on energy methods (which provide an alternative to balance of momentum methods).

Figure 6 also shows the variation in student mastery performance over the 12 semesters. Each objective shows the quartiles associated with mastery. Note that we consider ‘good’ mastery in the course grades as passing the green line; ‘strong’ mastery is getting to the cyan line. It is evident that the objectives with the most opportunity get mastered more thoroughly than those that are less available. That shows that the development of mastery is gradual and that a student who has not mastered an idea early in the course might succeed if given enough chances.

You can see how the level of mastery correlates with the final grade outcome in Fig. 7. Over the same 12 semesters, we break down the performances into those who get A grades (i.e., A+, A, and A-), B grades (i.e., B+, B, and B-), C grades (i.e., C+ and C), and failing grades (D and E). The D grade is considered failing because students must earn a C or better to advance to the next class. Hence, if they get a D they will need to repeat the course. The rate of failure in this course is about 7%, which confirms the general observation that active learning tends to significantly reduce failure rates [13].

![Fig. 7. Variation of mastery performance with course grade. Over the 12 semesters (SP14 through FA19), the mastery performance associated with A (n=222), B (n=381), C (n=226), and failing (n=62) grade outcomes is shown for Dynamics.](image)

One thing that is clear from the results associated with final grades is that there is a fairly wide range in level of mastery of some of the objectives, even among students who earn essentially the same final grade. Part of that is the normal spread across a grade range, but also note that the mastery performance only constitutes half of the final grade (computing projects make up the other half and the engagement multiplier can shift the grades up or down by half a grade level). It appears that A-students typically master at least 11 of the 16 objectives. B-students master about 9, and you can see that they tend to miss by a little on important ideas of balance of angular momentum and application of energy methods. The C-students master about 6 objectives. It is
difficult to judge the failing students because many of those students fail because they have not engaged the course and that might include not taking some of the mastery assessments.

Figure 8 shows how success has changed over time. The chart on the left shows how mastery looked for all passing students in the first two semesters of the project. The chart on the right shows the most recent two semesters. There has been a general increase in mastery with time, which can be attributed to improvements in the administration of the class, better training of the undergraduate teaching assistants, and increased proficiency with programming (freeing up more time to focus on exam preparation).

Further, student comments from anonymous course evaluations give some insight into how the students see the mastery-based grading system. For example, it is evident from the following comment that students are often not familiar with this type of grading system.

“...The mastery-based objective grading system was something new to me but at the same time something that I liked.”

It takes students some time to adapt to the system. Early comments from students tended to be along the lines of “I don’t know how I am doing in this course…” because they were looking for conventional scores. In reality, they were getting much more detailed feedback than most had ever gotten before. But they needed to learn how to interpret it. Over time the course documents to do that were improved.

You also find that students really make the connections with the benefits of the mastery-based grading system. For example, one student wrote:

“The mastery-based grading system is easily the best part of this course due to the fact that it lowers stress levels for students. Rather than seeing a F on a test score we just saw our bars go up or stay put and that was much less stressful than seeing a bad grade on a test or assignment. Mastery-based grading systems are excellent.”

Two others wrote:

“I liked the grading system of the course; it is really encouraging the students to perform well by improving the specific categories which you are not doing too well.”
“There were a lot of Mastery Assessments which made it hard to study for them since there was so many other things going on, but that aspect can also be good because it gives students more chances to perform well.”

Of course, not all student comments on mastery-based grading are positive, but most are, and the number of positive comments has increased with time as we have improved on our execution with the system.

The change in the grade conversations with students is indicative of how the students have changed the way they think about their learning in the course. It is no longer common to hear a student ask a question like, “What do I need to do on this assessment to get a B in the course?” Now, the questions tend to be more like, “I am struggling with free body diagrams; can I go over previous assessments with you to see where I am going wrong in that objective?” The conversations related to grades are much more constructive and have decreased the number of angry grade discussions that happen with students. The grading rubric is straightforward, and the students even have an opportunity to practice it during rehearsal exams where we do peer grading, which leaves very little doubt in their minds about what they need to do.

Lessons learned

Changing from a traditional grading system to the mastery-based grading system is hard to do incrementally. In fact, we fully implemented the new grading system in a single semester in Dynamics. Student reactions were quite negative because they did not understand the new system (and feared the worst). The angst was exacerbated by the fact that we had changed several elements of the course, leaving students with few avenues to compare with students who had already completed the course. One lesson that came through very clearly is that students rely very heavily on information about courses that are handed down from one class to the next. We also learned that the length of this information chain is not very long. In a semester or two most students were getting their information from students who had been in the course under the new implementation and expressed fewer concerns, especially about grading.

The second lesson that we learned, and one that many who implement pedagogical change learn, is that it is important to articulate why you are making the change and how it affects the students. Further, it is not enough to make that case once at the start of the semester. You need to rehearse the value propositions of any unusual course elements (and probably should the usual ones, too) periodically throughout the semester. We developed a set of documents that cover grading that we share with the students on the course website and suggest that they revisit when they get their first dashboard or any other feedback. We frequently point out how mastery-based grading is designed to support students by reducing exam stress, that it provides multiple opportunities to demonstrate mastery, and the measure of mastery is non-decreasing. Student response to mastery-based grading is generally viewed as positive by students, and the student-to-student communication on grading now favors this system over the traditional one.

Because the mastery-based grading system involves relatively coarse performance judgements made many times, the grading can be farmed out to a graduate student or professional grader. The inter-rater reliability of the system is quite high. With grading done by someone else, the instructor can focus more on the role of coach and advocate for learning. Still, there is quite a lot of data generated and managing it is not trivial. We have developed a set of computer programs that generate the dashboards and send feedback individually to students (most learning
management systems are not capable of providing the information on the dashboard). The third lesson that we learned was about the importance of having a good data management approach and how important it is to provide quick turnaround of feedback. Automating the feedback process has allowed this grade system to scale up with increasing course sizes.

Conclusion
The primary reason for moving to mastery-based grading was to provide more and better feedback, while at the same time disrupting some of the worst tendencies of students that have formed around traditional grading strategies. The process of finding sound mastery objectives was deceptively difficult, but once found they have proven to be a great context for thinking about how and what to teach and how to organize assessment. The change to mastery-based grading has achieved the primary objective, but it has also engendered a culture shift of students who experience this system. Conversations about grading with students are more focused on authentic learning issues than they were with the traditional system and students have shown that they understand and embrace the values associated with mastery-based grading.

References
