Standard-based Grading In Introductory Physics Laboratory Courses

Dr. Yan Wu, University of Wisconsin, Platteville

Yan Wu graduated from Tsinghua University, Beijing, China, in 1996 with a bachelor's degree in Precision Instruments and a minor in Electronics and Computer Technology. She received her M.S. degree in Mechanical Engineering from the University of Alabama in 1998. She received her Ph.D. in Electrical Engineering from the University of Illinois, Urbana-Champaign, in 2005. Her Ph.D. thesis work was in the area of micro-electro-mechanical systems (MEMS) with a focus on effect of space charges on micro- to nano-scale electrostatic actuation. Upon receiving her Ph.D., she worked as a Postdoctoral Research Associate in the Department of Mechanical Science and Engineering in the University of Illinois, Urbana-Champaign, where she worked in multiple projects using scanning probe microscopy to study material properties. In 2009, Yan Wu joined the faculty of the Department of Engineering Physics at the University of Wisconsin, Platteville. From fall 2015 to summer 2016, Yan Wu completed one year of sabbatical as a visiting scholar in the Department of Biomedical Engineering at University of Wisconsin – Madison.

Anna Drazkowski

Standard Based Grading In Introductory Physics Laboratory Courses

Abstract

Introductory physics laboratory courses are often among the first few laboratory courses that all students in engineering discipline take. The learning objectives of introductory physics lab courses, such as understanding experimental error and measurement uncertainty, are the foundation to the more sophisticated laboratory oriented courses in the junior and senior level engineering curriculum. However, as nearly all physics lab instructors can attest, many students fail to master certain lab skills that are related to understanding experimental error and measurement uncertainty, such as error analysis, linear regression, and reporting measurements results with appropriate uncertainty, even though these skills are practiced repeatedly throughout the course. In an effort to address this problem, the authors investigated whether standards-based grading (SBG) is a more effective assessment approach to help students with developing basic laboratory skills. Traditionally, the grades of the introductory physics lab are determined by a series of lab reports. In standards-based grading, students' grades are attached to the specific learning objectives of the course. Instead of receiving a numerical score for each lab report, students are given feedback on the mastery of each of the learning outcomes using the EMRN rubric, where E stands for excellent, M stands for meet expectations, R stands for revision needed and N stands for not assessable. The grade on any particular standard indicates how well a student has mastered the standard at that point in time. This grade may increase over time as the student demonstrates increasing understanding and skill and it may decrease if new evidence reveals previously overlooked flaws in understanding. The hypothesis is that by linking assessments to specific learning objectives and providing students with opportunities to remedy deficiencies, student will understand exactly what they need in order to master the core knowledge and skills in introductory physics lab, thus become personally invested in acquiring the knowledge and skills for advancement. We will present how we implement SBG in introductory physics laboratory courses and the evaluation results using pre and post knowledge assessment tools in sections graded with SBG method and sections with traditional method.

Introduction

Laboratory based courses are an essential part of an engineering education. More and more engineering programs are placing greater emphasis on the importance of hands-on experience offered by laboratory courses. Compared to the efforts and resources invested in purchasing expensive state-of-the-art lab equipment and developing innovative lab modules, little attention is paid to assessing student learning and reflecting on the teaching methods in lab courses. Without careful design and assessment, instructors are more likely to push students into busy work, without achieving real learning and understanding.

Introductory physics laboratory courses are often among the first few laboratory courses that all students in the engineering discipline take. The learning objectives (LO) of introductory physics lab courses, such as understanding experimental error and measurement uncertainty, are the foundation for the more sophisticated laboratory oriented courses in junior and senior level engineering curricula. For example, in my own university, many engineering majors – such mechanical engineering, electrical engineering, and civil engineering – all have their disciplinary

specific measurements and instrumentation courses in their junior or senior years. These courses use skills that should have been built in introductory physics labs; which include error analysis, linear regression, and reporting measurement results with appropriate uncertainty. However, as nearly all physics lab instructors can attest, many students fail to master certain lab skills even though these skills are practiced repeatedly throughout the course. Thus many students are ill-prepared when they enter their junior and senior level instrumentation courses, although they already have passed the introductory physics courses.

One of the causes of this problem comes from a common misunderstanding that both students and instructors make: that physics labs serve only as a reinforcement of the content in the theory or lecture portion of a physics course. The underlying rationale is that students will better understand the physics if they conduct experiments and see for themselves how the physics principles work in "real life." As a consequence, there are no independent learning objectives for the lab portion of a physics course, or it is often unclear to students what they need to learn in labs. However, recent data[1] from three institutions and nearly 3000 students show there was no statistically measurable lab benefit in boosting content mastery of physics. Students perform similarly on questions whose target content was reinforced by a lab activity as on questions whose target content lacked an associated lab activity. This result, although surprising, makes a strong case for reexamining the purpose of labs and how they are taught.

Another cause of the failure in mastering basic lab skills in spite of repeated learning opportunities lies in the method of assessing student learning in physics labs. The traditional method of assessing student learning in physics labs is collecting lab reports. Students turn in a report after each lab, and rubrics are used for grading purposes. The final grade of the lab course or the lab portion of a combined lab/lecture course is the average grade of a series of lab reports. There are several problems of this grading method. First, numerical grades are often inaccurate in assessing the mastery of a set of specific learning objectives. For example, a student may consistently earn a score of 80% from a series of lab reports, with 20% of point loss attributed to propagating error in measurements. Because of this method of assessment, the student can completely fail to achieve the learning objective of being able to propagate error in measurements, yet earn a passing grade for labs. Secondly, there is little incentive for students to learn from mistakes with the traditional method. Making mistakes is a natural and essential part of learning process. As instructors, we should strive to give specific and timely feedback to students about their lab learning, in the hopes that students can learn from their mistakes. Many instructors use some generic lab grading rubrics in order to grade objectively and provide concrete feedback. However, students seldom choose to review the graded report with great care once the grade is given. They aren't given much opportunity to correct their mistakes and don't have motivation to do better the next time if the grade is passable. Finally, with the traditional grading method, students focus on earning points instead of mastering the material. As a lab instructor, I often witness students going to great lengths to perfect the format of the report or simply adding unnecessary words to show their effort. Still, a beautifully formatted report with a lot of words can completely miss the point of the lab.

Standard-based grading (SBG) is an assessment method that relies on students demonstrating mastery of learning objectives (sometimes referred to as standards). In standard-based grading, students' grades are attached to the specific learning objectives (LOs) of the course. Instead of receiving a numerical score for each lab report, students are given feedback

on the mastery of each LO using the EMRN rubric, where E stands for excellent, M stands for meets expectations, R stands for revision needed, and N stands for not assessable. The grade on any particular standard indicates how well a student has mastered the standard at that point in time. This grade may increase over time as the student demonstrates increasing understanding and skill, or it may decrease if new evidence reveals previously overlooked flaws in understanding. One obvious challenge of SBG is that we, as educators, must be deliberate in articulating what we want to achieve through all the teaching and learning activities in a course. In the case of introductory physics labs, goals range from reinforcing content, learning about measurement and uncertainty, practicing communication skills, developing teamwork skills, and, more broadly, learning that physics is an experimental science. It is unrealistic to hit all those goals. A practical benefit SBG for educators is that SBG made assessment easy. Academic programs in higher education are constantly under scrutiny of various organizations, such as the Higher Learning Commission (HLC) and Accreditation Board for Engineering and Technology (ABET). Because assessments and evidence collection are embedded in the day-to-day and lab-to-lab work, it is easier to integrate SBG into course level assessment.

Perhaps the greatest benefit of standard-based grading is for the students.   By linking assessments to specific learning objectives and providing students with opportunities to remedy deficiencies, student will understand exactly what they need in order to master the core knowledge and skills in introductory physics lab, and thus become personally invested in acquiring the knowledge and skills for advancement. In this article, I will present how I implement SBG in introductory physics laboratory courses, the design of study to evaluate the effectiveness of SBG, and the use of pre and post knowledge assessment tools in sections graded with the SBG method and sections with the traditional method.

Method

Overall design of study

In fall 2018, I taught two sections of General Physics I with around 50 students in each section. The course had a lecture/lab format. There were four hours of lecture and two hours of lab for each week. There were a total of 11 labs in the semester, all closely related to the content of the lecture of the week. The lab portion of the course takes up 12% of the total course grade.

With the help of a teaching assistant, I was able to set up and maintain a separate lab grade book that is specific for each section. We used a traditional grading method for one section; this section is referred as 'Traditional Section' in this paper. We used Standard Based Grading (SBG) for the other section, referred to as 'SBG' section. To evaluate the effect of SBG on students' performance, I used two assessment tools and administered pre/post assessments in both sections. Assuming student population are of random nature in each section, comparison of the results of the assessments in the SBG section with those in the Traditional section will supposedly lead to a conclusion of the effect of SBG. The first tool I used was a modified version Measurement Uncertainty Quiz (MUQ), first developed by Duane Deardorff and Robert Beichner[2]. This is a tool that focuses on a threshold concept of physics labs, namely measurement uncertainty. I kept most of the original quiz questions and deleted a few to make MUQ more suitable for entry level physics students. The second tool was 'The Colorado Learning Attitudes about Science Survey for Experimental Physics (E-CLASS).'[3]. E-class

survey assesses students' views about their strategies, habits of mind, and attitudes when doing experiments in lab classes. One of the ways to use E-Class survey is to evaluate if a student's mindset when doing experiments is more like an expert's or a novice's in experimental physics. Both sections were given MUQ and E-Class at the beginning of the semester (pre-assessment) and at the end of a semester (post-assessment). To encourage participation, the pre/post assessments were rewarded with completion points as part of the course grade.

SBG implementation

There are three guiding principles in implementing SBG in order to capture the true spirit or the original intention of SBG.

1. The feedback/grading should be linked to specific learning objectives/outcomes in order to help students know what they need to learn.

2. The grade should communicate students' degree of mastery of the learning objectives, and not be confounded with other variables such as showing up in class, good effort in completing assignment, etc.

3. Students should be permitted to remedy their deficiencies in their learning because mistakes are an important part of the learning process.

In the following, I describe how each of the guiding principles is implemented in my SBG section of physics lab course.

There are several guidelines for setting Learning Objectives (LO) for undergraduate physics labs, including the American Association of Physics Teachers (AAPT) Recommendations for the Undergraduate Physics Laboratory Curriculum[4] and the AP Physics learning outcomes from CollegeBoard. After taking into consideration the physics curriculum of my university and these nationally accepted guidelines, I idenified four major learning objectives, each with specific sub-objectives briefly described in Table 1. The detailed LO with its grading rubrics are listed in Appendix A. There are total 12 specific LOs, denoted as LO1.a, LO1.b, LO4.c as such. Each LO is worth of 1% of the total course grade, as the lab portion is only 12% in my course. All of these LOs are re-assessable using the current lab modules and lab resources in my university.

Table 1 Learning Objectives in Standard Based Grading Section

LO1 Measure real phenomena and apply physical model in observations	a) Describe the purpose of an experiment or a problem to be investigated
	b) Identify theoretical model applicable to the experiment
	c) Construct an experimental set-up, troubleshoot it, and collect data
LO2 Understand experimental error and measurement uncertainty	a) Identify sources of error and estimate measurement uncertainties
	b) Be able to propagate uncertainty in calculations, i.e. given several measured quantities with their uncertainties, find the uncertainty of another quantity that is computed from their values
LO3 Understand how to analyze data	a) Display data in graphical or tabular form with appropriate labels and units
	b) Perform calculations with data
	c) Be able to identify linear functional relationship between physical quantities
	d) Be able to linearize nonlinear function using variable substitution and perform linear regression analysis on the linearized data
LO4 Be able to summarize and communicate results	a) Report measurement results with uncertainties with correct format
	b) Make comparisons between experimental data and theoretical predictions
	c) Draw inferences and conclusions from experimental data

It is import not to set too many learning objectives for several reasons. First, if there are too many LOs, it is difficult to design labs to re-assess all of them over the course of the semester. One import aspect of SBG is to allow students to learn from their mistakes, thus improving their mastery, in the same learning outcome. Secondly, too many objectives make it difficult to keep track of the learning status of every student on each of the standards. The grade book becomes cumbersome and complicated, and thus fails to communicate with students. In his recent work, Zimmerman [5] provides a rule of thumb formula for a reasonable number of learning objectives (LO) in SBG courses, given the limited amount of assessment time available to the students. According to this formula, my choice of 12 LO is a good quantity.

Next, I revised the lab instructions so that the questions and lab report requirements are clearly linked to the specific learning objectives. Students can find what specific learning objectives are being assessed when they answer certain questions or show certain work in a lab report. There are a total of 11 labs in the semester. Each lab assesses only a specific subset of the 12 LOs. Some LOs are accessed more often than the other. For example, LO1.c "Construct an experimental set-up, troubleshoot it, and collect data" is assessed in almost every lab. LO3.d "Be able to linearize nonlinear function using variable substitution and perform linear regression analysis on the linearized data" is assessed in only three labs. For each lab report, students are

given feedback on the mastery of each of the LO assessed in the lab using the EMRN rubric, where E stands for exemplary or exceeds expectations, M stands for meet expectations, R stands for revision needed, and N stands for not assessable. In order to integrate grades from SBG section with the overall grades of the course, a numerical conversion is set up such that E=4, M=3, R=1, and N=0.

A learning mastery gradebook is kept using online Learning Management System (LMS) Canvas. The grade on any particular LO indicates how well a student has mastered the LO at that point in time. To ensure that the most recent assessment result plays a more significant role in a student's grade than the previous, I used a 70/30 decaying average algorithm to calculate the grade after grading each lab report, where the most recent EMRN score takes up 70% of the weighted average and all the previous scores take up only 30%. Mastery level is set to be 3 ('Meet expectations' or M). A screenshot of such a learning mastery gradebook in the middle of a semester is shown in Figure 1. The pie chart shows how the entire study body in the class does in a specific LO. In the figure, the pie chart shows 22% of the students did very well in LO1.a "Describe the purpose of an experiment or a problem to be investigated" and 58% of the students meet the exceptions. Only 20% of the students still need to revisit this LO. Over the course of the lab, it is important to remind students that they should check the learning mastery gradebook for their lab grades. It is also important to point out that occasional bad grade (R or N) will not impact their grade in the long run as long as they can improve the material/skill and earn a good grade (E or M) later.
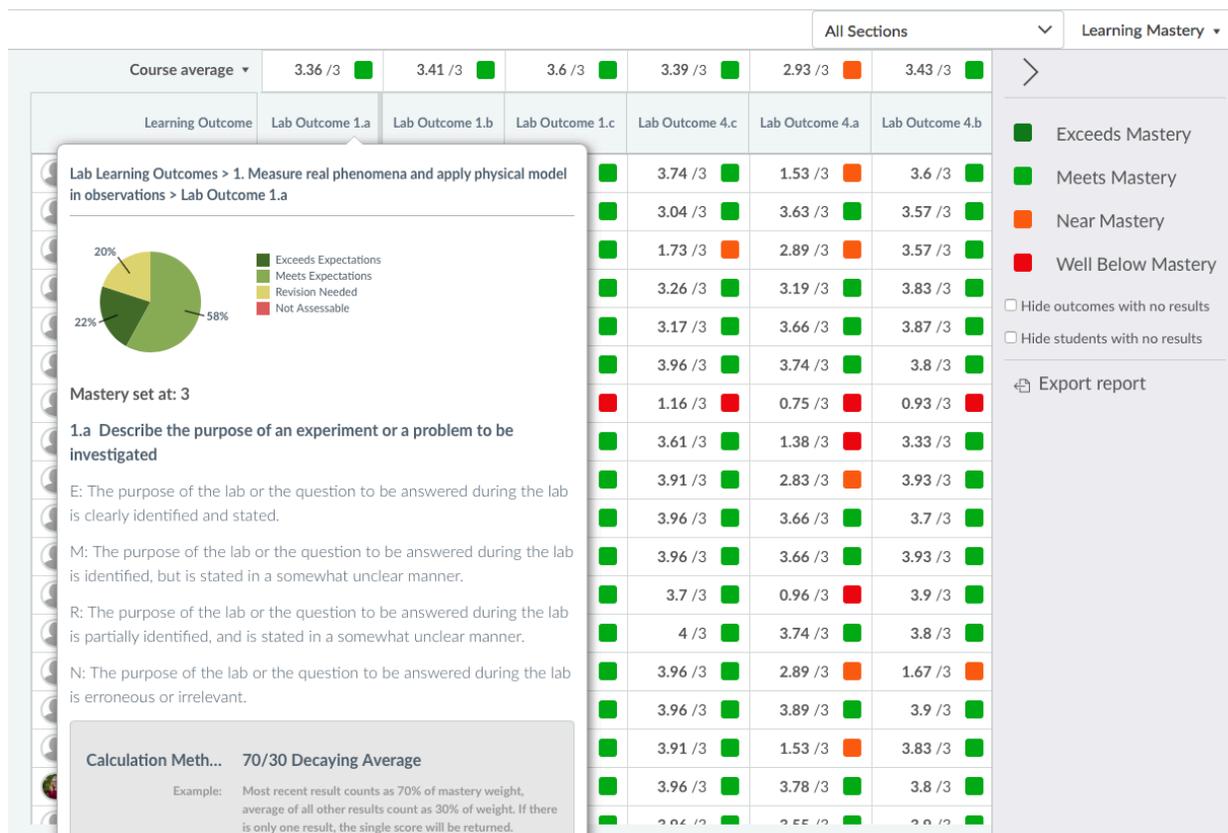

Figure 1 Learning Mastery Gradebook maintained using Canvas in SBG section of General Physics I lab.

<u>Traditional Grading</u>

One advantage of SBG grading is that students can get timely and concrete feedback on their learning status. The goal of giving specific feedback can be achieved in the traditional grading method with clear grading rubrics. In order to be fair in comparison, lab reports in the traditional grading section were graded with a generic grading rubric as listed in Appendix B. This generic grading rubric has many elements that overlap with the 12 LOs in SBG section. It is a generic rubric because it is the same for grading each of the eleven labs, where as in SBG section each lab only assesses a specific subset of the 12 LOs – thus there is a different rubric for each lab.

Results and Discussion

Both the modified MUQ and the E-Class survey were given in the first and the last week of the fall 2018 semester. Responses were first scanned for validity. If a student chose all the same option such as 'A' option in the multiple-choice questions of MUQ or all 'strongly agree' in E-Class survey, the response is considered a 'lazy' response and is not valid. Next, we kept only the data from students who had completed both the pre-assessment and post assessment. In other words, only valid paired responses were analyzed. For MUQ, the number of students who completed both pre and post test and have given valid responses is 26 in the SBG section and 40 in the traditional section. For E-class Survey, the numbers are 25 for SBG and 38 for traditional, respectively.

Pre- and post-test analyses have been widely used as a method of assessment in education and social science. Over the years, many ways to calculate learning gains have been proposed. The most popular might be Hake's gain, or normalized gain, introduced by Richard Hake in 1998 [6]. Recently, David Dellwo [7] has drawn attention to the fact that Hake's gain does not incorporate retention or loss of knowledge during the instructional period. In this article, we use the learning gain defined by David Dellow because it measures the likelihood that a mistake on the pre-instruction test is corrected on the post-instruction test. To analyze the MUQ data, the paired responses can be divided into four groups according to the four possible transitions that can occur between the pre-test and the post-test. These transitions (presented in Table 2) consist of: right to right (RR), right to wrong (RW), wrong to wrong (WW) and wrong to right (WR). Learning gain G defined by David Dellwo measures the percentage of responses that start off wrong in pretest (WR and WW) but turn to right answer in post test (WR). It can be calculated as

$$G = \frac{WR}{WW + WR} \qquad (1)$$

Table 2 Four possible transitions of answers between the pre-test and post-test.

Pre-test	Post-test	Transition
Right	Right	RR
Wrong	Wrong	WW
Right	Wrong	RW
Wrong	Right	WR

Figure 2 shows the percentage distribution of the four transition groups in the SBG section. The learning gain in SBG is 53% based on this data. Figure 3 shows the percentage distribution of the four transition groups in the traditional section. The learning gain in the traditional section is 45% based on this data. The modified MUQ focuses on calculating error from measurements, accuracy, precision, and sources of error in measurement. It is a good assessment on LO2 "Understand experimental error and measurement uncertainty" in the SBG section. Comparing results shown in Figure 2 and Figure 3, a higher learning gain in SBG section indicates that SBG can be an effective method to enhance students' understanding of measurement uncertainty concepts and their mastery in calculating error from measurements. One potential explanation of these results comes from the theory of threshold concept or troublesome knowledge. In nearly every field, there is some concept or piece of knowledge that is counterintuitive and demands learners a transformed way of understanding, interpreting, or viewing something. Measurement uncertainty has many characteristics of the threshold concept as described by Jan Meyer et. al [8]. The learning process to achieve transformative understanding usually takes longer time and has a recursive nature. SBG model gives ample opportunity for students to revisit the threshold concept and gives them the incentive to learn from their mistakes.
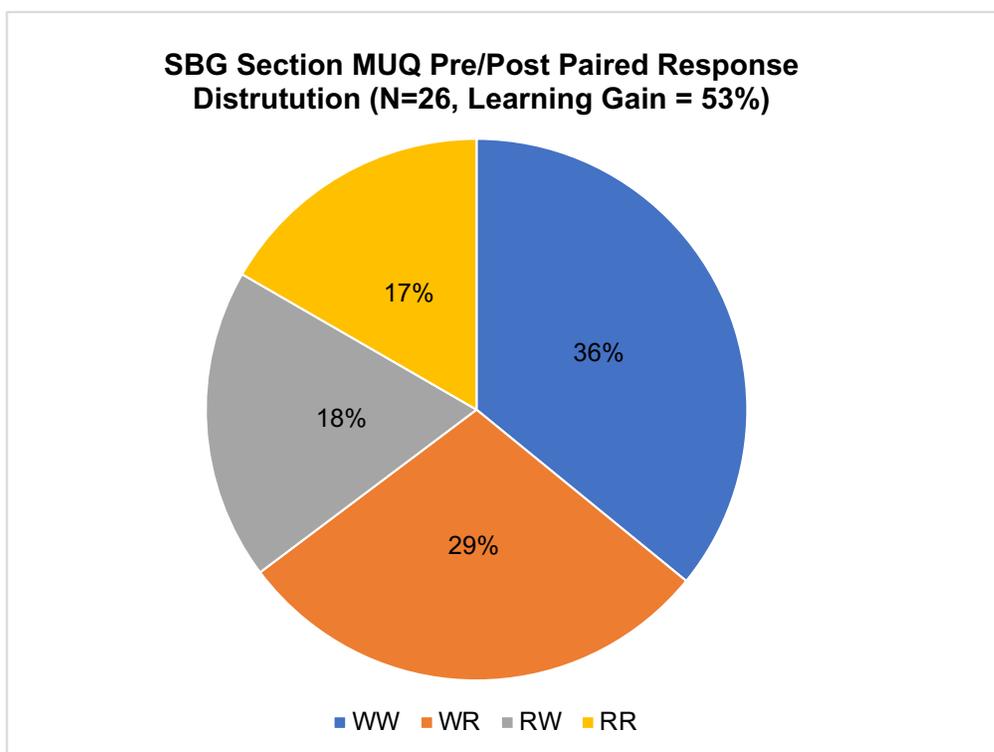

SBG Section MUQ Pre/Post Paired Response Distrutution (N=26, Learning Gain = 53%)

17%
36%
18%
29%

■ WW ■ WR ■ RW ■ RR

Figure 2 Percentage distribution of the four possible transitions of answers between the pre-test and post-test groups in the SBG section. There are 6 questions and 26 paired tests, totaling 156 responses. 36% of the responses were wrong in both pre-test and post-test (WW). 17% were right in both pre and post tests (RR). 29% shows transition from wrong to right (WR), and 18% shows transition from right to wrong (RW). The learning gain measures the percentage of responses that start off wrong in pre-test (WR and WW) turn to the right answer in post-test (WR) and is 53% in SBG section.
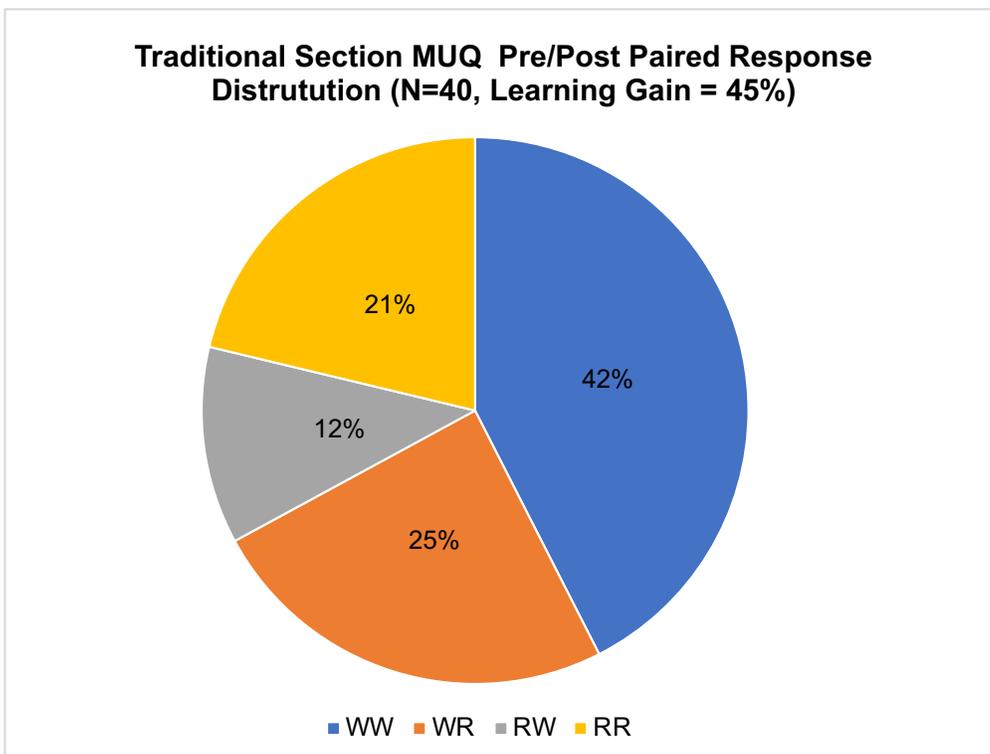
Figure 3 Percentage distribution of the four possible transitions of answers between the pre-test and post-test groups in the traditional grading section. There are 6 questions and 40 paired tests, totaling 240 responses. 42% of the responses were wrong in both pre-test and post-test (WW). 21% were right in both pre and post tests (RR). 25% shows transition from wrong to right (WR), and 12% shows transition from right to wrong (RW). The learning gain measures the percentage of response that start off wrong in pre-test (WR and WW) turn to right answer in post-test (WR) and is 45% in traditional grading section.

The 'Colorado Learning Attitudes about Science Survey for Experimental Physics (E-CLASS).' assesses students' perceptions of the gap between classroom laboratory instruction and professional research. It consists of a total of 30 statements about students' strategies, habits of mind, and attitudes when doing experiments in lab classes. Students choose either agree, disagree, or neutral in response to the statements. The responses then are categorized to 'expert-like response' or 'novice-like response'. For example, statement #16 says "The primary purpose of doing a physics experiment is to confirm previously known results." An expert in experimental physics will disagree with this statement. If a student chose disagree, then this is counted as an expert-like response. For each statement, the number of students who chose expert-like response is divided by the total number of students who took both pre and post surveys and produced valid responses (no "lazy" responses). The fraction of expert-like responses is calculated for each of the 30 statements. The average fraction of expert-like responses and the standard error of the average are calculated among all 30 statements and the standard error of the average is also calculated. Figure 4 shows the comparison between SBG section and traditional section in pre and post E-Class scores. The SBG section has an average expert-like fraction of 0.65±0.05 and 0.66±0.05 in pre-instruction survey and post-instruction survey, respectively. The traditional section has an average expert-like fraction of 0.59±0.04 and 0.53±0.04 in pre-instruction survey and post-instruction survey, respectively. We infer from the

numbers that the fraction of expert-like responses remain relatively the same before and after a semester's lab course for the SBG section. For the traditional section, the fraction seems to drop, but not conclusively, due to the large error bar.
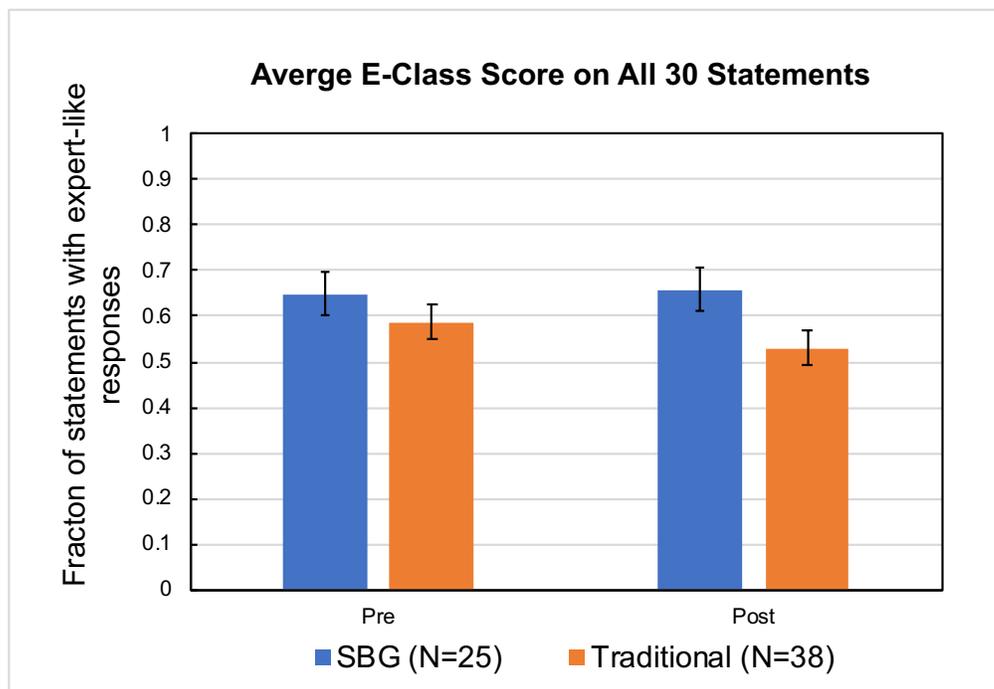

Figure 4 Comparison between the SBG section and the traditional section in pre and post E-Class scores on all 30 statements. The SBG section had an average expert-like fraction of 0.65±0.05 and 0.66±0.05 in pre-instruction survey and post-instruction survey, respectively. The traditional section had an average expert-like fraction of 0.59±0.04 and 0.53±0.04 in pre-instruction survey and post-instruction survey, respectively.

Further examination the data reveals that for statement #4, #16, #17, and #29, both the SBG and traditional section score very low in either pre-instruction survey and post-instruction, with the fraction of expert-like responses to be less than 0.25. For the remaining questions, the fraction is at least more than 0.4 in one of the surveys. Developing expert-like responses to these statements (#4, #16, #17, and #29) is probably out of reach for students in introductory physics labs and is not affected by the grading method of lab report. For example, statement #17 says 'When I encounter difficulties in the lab, my first step is to ask an expert, like the instructor.' The expert-like response is 'disagree' but freshman students in introductory physics labs probably will all agree, no matter if their lab report is graded with SBG method or traditional method. To eliminate the possibility of 'blurred' results due to these questions, we excluded the responses from our analysis and re-calculated the average expert-like fraction. The results are shown in Figure 5. Excluding statement #4, #16, #17, and #29, the SBG section has an average expert-like fraction of 0.65±0.05 and 0.66±0.05 in pre-instruction survey and post-instruction survey, respectively. The traditional section has an average expert-like fraction of 0.59±0.04 and 0.53±0.04 in pre-instruction survey and post-instruction survey, respectively. The results indicate the expert-like fraction remains stable in SBG section yet drops definitely in traditional section. This result is interesting because as physics instructors we certainly hope our instruction will not 'hurt' students' growth in becoming experts in experimental physics. How did the traditional

method of lab grading do something contrary to our wishes? With this question in mind, I examined which statements the traditional section had a significant drop in expert-like responses. There are several statements that show a 'negative' transition, but the most significant one is statement #21, which says "I am usually able to complete an experiment without understanding the equations and physics ideas that describe the system I am investigating". In the traditional section 55% of students disagree with this statement at the beginning of the semester yet after a semester's lab course, only 22% of students disagree. In the SBG section, 56% of students start off with disagree and 48% disagree at the end of the semester. This means the SBG method, at the very least, is doing a better job in retaining students' motivation to understand the physics concepts of the experiment.
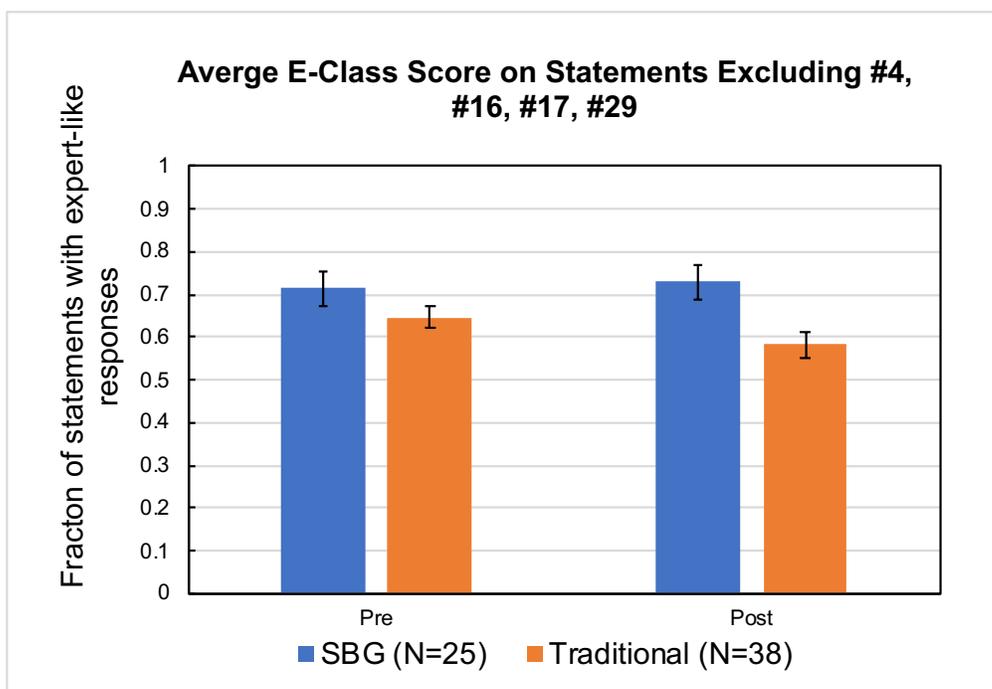

Figure 5 Comparison between the SBG section and the traditional section in pre and post E-Class scores on statements excluding #4, #16, #17, #29. The SBG section had an average expert-like fraction of 0.71±0.04 and 0.73±0.04 in pre-instruction survey and post-instruction survey, respectively. The traditional section had an average expert-like fraction of 0.65±0.03 and 0.58±0.03 in pre-instruction survey and post-instruction survey, respectively.

Conclusions and Recommendations

Results from this study indicate that standard based grading (SBG) has a positive impact to student learning in introductory physics lab courses. Compared to traditional grading methods, it helps students in developing skills that relate to threshold concepts such as measurement uncertainty. It also does a better job at maintaining students' attitudes in doing experiments in an expert-like manner. However, the data presented here is only from one semester and one instructor's class. Further investigation is needed to generalize the conclusion to introductory lab courses. The lab portion of the General Physics I course is only counted as 12% of the overall grade. It would be interesting to see the effects of SBG in stand-alone lab courses or in courses with greater weight for labs.

References

[1] N. G. Holmes and C. E. Wieman, "Introductory physics labs: We can do better," *Phys. Today*, vol. 71, no. 1, pp. 38–45, Jan. 2018.

[2] D. Deardorff, *Introductory Physics Students' Treatment of Measurement Uncertainty*. Ph.D. dissertation, North Carolina State University: Department of Physics, 2001.

[3] B. M. Zwickl, N. Finkelstein, and H. J. Lewandowski, "Development and validation of the Colorado learning attitudes about science survey for experimental physics," presented at the 2012 PHYSICS EDUCATION RESEARCH CONFERENCE, Philadelphia, PA, USA, 2013, pp. 442–445.

[4] D. MacIsaac, Ed., "*Report: AAPT Recommendations for the Undergraduate Physics Laboratory Curriculum* ," *Phys. Teach.*, vol. 53, no. 4, pp. 253–253, Apr. 2015.

[5] T. Zimmerman, "Grading for Understanding – Standards-Based Grading," *Phys. Teach.*, vol. 55, no. 1, pp. 47–50, Jan. 2017.

[6] R. R. Hake, "Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses," *Am. J. Phys.*, vol. 66, no. 1, pp. 64–74, Jan. 1998.

[7] David Dellwo, "Course assessment using multi-stage pre/post testing and the components of normalized change," *J. Scholarsh. Teach. Learn.*, vol. 10, no. 1, pp. 55–67, Jan. 2010.

[8] J. H. F. Meyer and R. Land, "Threshold concepts and troublesome knowledge (2): Epistemological considerations and a conceptual framework for teaching and learning," *High. Educ.*, vol. 49, no. 3, pp. 373–388, Apr. 2005.

Appendix A: SBG Lab Learning Outcomes and Grading Rubrics

1. Measure real phenomena and apply physical model in observations
   a) Describe the purpose of an experiment or a problem to be investigated
      E: The purpose of the lab or the question to be answered during the lab is clearly identified and stated.
      M: The purpose of the lab or the question to be answered during the lab is identified but is stated in a somewhat unclear manner.
      R: The purpose of the lab or the question to be answered during the lab is partially identified and is stated in a somewhat unclear manner.
      N: The purpose of the lab or the question to be answered during the lab is erroneous or irrelevant.
   b) Identify theoretical model applicable to the experiment
      E: Equations or diagrams show accurate and thorough understanding of scientific concepts underlying the lab
      M: Equations or diagrams show accurate understanding of most scientific concepts underlying the lab
      R: Equations or diagrams show limited understanding of most scientific concepts underlying the lab, there are major mistakes in the equation or diagrams
      N: No equations or diagrams shown scientific concepts underlying the lab
   c) Construct an experimental set-up, troubleshoot it, and collect data
      E: Carried out the experiment according to lab procedure completely and accurately

M: Carried out the experiment completely with minor mistake in data collection
R: Did not finish the experiment completely or there are major mistakes in data collection
N: Did not do the experiment
2. Understand experimental error and measurement uncertainty
   a) Identify sources of error and estimate measurement uncertainties
      E: Be able to identify systematic error and random error in a measurement, and correctly assign uncertainty to random error for single or multiple measurements
      M: Be able to assign uncertainty to measurements yet not completely clear about the nature of error
      R: Demonstrate the awareness of uncertainty in measurements but do not know how to assign uncertainty correctly
      N: No evidence shown the awareness that measurements come with uncertainty
   b) Be able to propagate uncertainty in calculations, i.e. given several measured quantities, with their uncertainties, find the uncertainty of another quantity that is computed from their values. *e.g.* $z = c\, x^m y^n$, *find Dz*
      E: Propagate uncertainty accurately
      M: Minor mistake in calculation
      R: Major mistake in calculation
      N: No uncertainty propagation shown
3. Understand how to analyze data
   a) Display data in graphical or tabular form with appropriate labels and units
      E: Professional looking and accurate representation of the data in tables and/or graphs. Graphs and tables are labeled and titled.
      M: Accurate representation of the data in tables and/or graphs with labels and units.
      R: Partial accurate representation of the data, graphs or tables are missing labels or units
      N: Data are not shown OR are inaccurate.
   b) Perform calculations with data
      E: All sample calculations are shown with equations and one set of numbers plugged in, and the results are correct.
      M: Calculations are mostly shown, and the results are correct.
      R: Calculations are partially shown, and there some mistakes in the results.
      N: No calculations are shown, OR results are inaccurate.
   c) Be able to identify linear functional relationship between physical quantities
      1) Fit lines and display linear equations to data points in graphs
      2) Find the slope and intercept with uncertainties using linear regression analysis
         E: Perform both 1) and 2) accurately
         M: Perform 1) and 2) with minor mistake
         R: Perform 1) or 2) with major mistake
         N: No evidence showing either 1) or 2).
   d) Be able to linearize nonlinear function using variable substitution and perform linear regression analysis on the linearized data
      1) Display linearized data in graph and show linear equation
      2) Find the slope and intercept of the linearized data with uncertainties
      3) Relate the slope and intercept with uncertainties to physical variables with uncertainties

E: Perform all three above accurately
M: One mistake in any of the three
R: Two mistakes in above three
N: No evidence or no correct answer shown

4. Be able to summarize and communicate results,
   a) Report measurement results with uncertainties with correct format.
      1) Uncertainties should not have more than 2 significant figures
      2) Determine significant figures of results based on measurement uncertainty
      3) Use appropriate SI units for physical quantities
         E: Perform all three of above accurately
         M: One mistake in any of the three
         R: Two mistakes in above three
         N: No evidence or no correct format shown
   b) Make comparisons between experimental data and theoretical predictions
      E: Be able to make theoretical prediction and make judgment on whether the experimental results agree with theoretical prediction within uncertainty, reasoning shows accurate understanding of the scientific nature of the experiment.
      M: Be able to make connection between experimental data with theoretical prediction, yet the reasoning shows minor misunderstanding of the scientific nature of the experiment.
      R: Be able to make connection between experimental data with theoretical prediction, yet the reasoning shows major misunderstanding of the scientific nature of the experiment.
      N: Unable to relate experimental data with theoretical predictions
   c) Draw inferences and conclusions from experimental data
      E: Conclusion includes whether the findings supported the hypothesis, possible sources of error, and what was learned from the experiment.
      M: Conclusion includes whether the findings supported the hypothesis and what was learned from the experiment.
      R: Conclusion includes what was learned from the experiment.
      N: No conclusion or shows little effort and reflection

Appendix B: Generic Lab Grading Rubrics in Traditional Section

	4	3	1	0
Question/Purpose	The purpose of the lab or the question to be answered during the lab is clearly identified and stated	The purpose of the lab or the question to be answered during the lab is identified, but is stated in a somewhat unclear manner.	The purpose of the lab or the question to be answered during the lab is partially identified, and is stated in a somewhat unclear manner.	The purpose of the lab or the question to be answered during the lab is erroneous or irrelevant.
Error Analysis	Experimental errors, their possible effects, and ways to reduce errors are discussed. Be able to estimate measurement uncertainties and propagate uncertainty in calculations accurately.	Experimental errors and their possible effects are discussed. Be able to estimate measurement uncertainties and propagate uncertainty in calculations with minor mistakes.	Experimental errors are mentioned. There major mistakes in uncertainty calculations.	There is no discussion of errors.
Calculations	All calculations are shown and the results are correct and labeled appropriately.	Some calculations are shown and the results are correct and labeled appropriately.	Some calculations are shown and the results labeled appropriately.	No calculations are shown OR results are inaccurate or mislabeled.
Data	Professional looking and accurate representation of the data in tables and/or graphs.	Accurate representation of the data in tables and/or graphs. Graphs and tables are labeled and titled.	Accurate representation of the data in written form, but no graphs or tables are presented.	Data are not shown OR are inaccurate.

	4	3	1	0
Analysis	The relationship between the variables is discussed and trends/patterns logically analyzed. Be able to use linear regression tool in analysis and know how to analyze the error in regression model.	The relationship between the variables is discussed and trends/patterns logically analyzed. Be able to use linear regression tool in analysis.	The relationship between the variables is discussed but no patterns, trends or predictions are made based on the data.	The relationship between the variables is not discussed.
Scientific Concepts	Report illustrates an accurate and thorough understanding of scientific concepts underlying the lab.	Report illustrates an accurate understanding of most scientific concepts underlying the lab.	Report illustrates a limited understanding of scientific concepts underlying the lab.	Report illustrates inaccurate understanding of scientific concepts underlying the lab.