Consistency Matters: Practical Grading Tips for Large Undergraduate Labs

Dr. Christopher D. Schmitz, University of Illinois at Urbana - Champaign

Christopher D. Schmitz received his Ph.D. in Electrical and Computer Engineering from the University of Illinois in 2002.

Dr. Jessica R TerBush, University of Illinois at Urbana - Champaign

Jessica received her B.S.E, M.S.E., and PhD in Materials Science and Engineering from the University of Michigan, Ann Arbor. After graduation, she worked as a post-doc for approximately three years at Monash University in Clayton, Victoria, Australia. She then spent three years working as a Senior Research Specialist at the Missouri University of Science and Technology in Rolla, Missouri, where she trained users on the focused ion beam (FIB), scanning electron microscope (SEM), and transmission electron microscope (TEM). In 2016, she moved to the University of Illinois, Urbana-Champaign, to serve as a lecturer in the department of Materials Science and Engineering. Here, she is responsible for teaching the junior labs as well as providing instruction on writing in engineering.

Katherine Ansell, University of Illinois at Urbana - Champaign

Katie Ansell is a Teaching Assistant Professor of Physics at the University of Illinois, Urbana-Champaign. Her teaching and research activities focus on the practical and social aspects of the classroom that contribute to the development of student expertise in Introductory Physics Laboratories.

Elaine C. Schulte, University of Illinois at Urbana - Champaign

Elaine C. Schulte is the Assistant Director of Undergraduate Programs for the Department of Physics and Senoir Lecturer at the University of Illinois, Urbana-Champaign. Her teaching interests include formalized problem-solving technique in non-major introductory physics courses and encouraging open-ended exploration in introductory physics laboratory. She also develops and manages graduate TA training and professional developent activities.

David Mussulman, University of Illinois at Urbana - Champaign

IT professional newly transitioned to teaching faculty. Engineering and Information Sciences.

Chandrasekhar Radhakrishnan, University of Illinois at Urbana - Champaign
Siva Nalla, University of Illinois at Urbana - Champaign

Siva Nalla is a Graduate Student Teaching Assistant in Bioengineering at the University of Illinois Urbana-Champaign (UIUC), where he assists with courses in Quantitative Physiology and Cell and Tissue Engineering. He earned his B.S. in Materials Science and Engineering at UIUC, M.S. in Teaching Chemistry at UIUC, and is currently a final year Ph.D. candidate in Bioengineering at UIUC. His master's thesis focused on inquiry based learning in general chemistry courses. Prior to his Ph.D. program, he has 12 years of lead TA instruction experience in General Chemistry and Biomaterials Engineering courses at UIUC. He also developed and instructed a Science and Engineering Physics course at The Hong Kong University of Science and Technology for a Center for Talented Youth summer program from 2011-2019.

Prof. Caroline Cvetkovic, University of Illinois Urbana-Champaign

Caroline Cvetkovic is a Teaching Assistant Professor of Bioengineering at the University of Illinois Urbana-Champaign, where she instructs courses in quantitative physiology, neural engineering, biofabrication, and heat transfer. She earned her B.S., M.S., and Ph.D. in Bioengineering at the University of Illinois Urbana-Champaign. She then completed a postdoctoral fellowship in the Center for Neuroregeneration and Department of Neurosurgery at the Houston Methodist Research Institute.

Dr. Rebecca Marie Reck, University of Illinois at Urbana - Champaign

Rebecca M. Reck is a Teaching Associate Professor of Bioengineering at the University of Illinois Urbana-Champaign. Her research includes alternative grading, entrepreneurial mindset, instructional laboratories, and equity-focused teaching. She teaches biomedical instrumentation, signal processing, and control systems. She earned a Ph.D. in Systems Engineering from the University of Illinois Urbana-Champaign, an M.S. in Electrical Engineering from Iowa State University, and a B.S. in Electrical Engineering from Rose-Hulman Institute of Technology.

Consistency Matters: Practical Grading Tips for Large Undergraduate Labs

Introduction

Practical considerations of teaching large undergraduate lab courses often call for the use of instructional teams that work under the supervision of a course director to teach, grade, and otherwise support the success of a class and its students. Grading and feedback are frequently challenging to align within an instructional team of diverse backgrounds. Much of the grading in undergraduate lab courses includes evaluating subjective answers and student writing, a challenging, but important, skill. Many teaching assistants and new instructors have little experience providing feedback and grades for these types of assessments. This also leads to unique challenges in providing consistent feedback and grades between instructors of different experiences and expectations in the same course.

Our community of engineering lab and design course instructors have found that challenges with consistent and timely grading are a recurring theme in both formal and informal conversations. Both new and experienced course directors bring different challenges forward and exchange ideas for how to address those challenges. Common challenges mentioned by the group include how to ensure that grading is consistent across sections, how to provide timely feedback, and how to train teaching assistants. We received feedback from students on these courses about inconsistencies in grades, application of course policies, and related support.

In this tips and tricks paper, we will summarize knowledge gained from these discussions so that others may benefit from our community's work. We will focus on rubrics, team grading calibration activities, grading tools, and ways to improve efficiency and consistency. Each of these areas have had a positive impact on consistency and timeliness of grading and feedback in our courses. We observed that adopting new tools and practices opens the door for collaborative learning and continuous improvement within instructional teams.

Grades are often a topic of conversation in many contexts around a university. Grades and feedback are a form of communication, often cryptic, between a faculty member and the student as well as future employers, coaches, funding sources, and advisors. In student conversations, consistency of grades and fairness are common topics. Grades may also have an impact on students' mental health [1]. Additionally, grades should be valid, reliable, and equitable [1], [2].

There are a variety of grading schemes that are used in courses to translate scores on assessments into the standard letters or grade points used in many institutions in the United States. Standards-based grading focuses on students' achievement of learning objectives in the course [3]. Specifications grading sets measurable and achievable specifications for course assignments [4], [5]. In many grading schemes, rubrics are often used to define more detailed expectations for earning a grade to the graders and possibly the students [1], [5], [6]. Recently, there has also been a push for more automation in grading schemes as well [7], [8].

In large courses, there are often multiple people responsible for grading and providing feedback on student work. The way the workload is split and how the group collaborates once the grading varies significantly [9]. There are many examples about how to train teaching assistants (TAs) to grade student work in STEM courses [10], [11], [12]. Each of these examples provides a framework that has worked for the author in their course or institution. What is generally acknowledged is that TAs have a direct impact on the course and students [13], [14], [15].

Since the previous work and our experiences have shown that there are many effective ways to approach grading in large laboratory courses, we have presented several case studies of grading schemes that have been adapted and improved over time. Each of these cases provides context of the course, tools that are used, details about training TAs, and tips for improving equity, efficiency, and consistency of grades. In the discussion, we provide some common themes from the cases.

Case Studies

In our discussions, we have found that there are a variety of solutions that may work in large lab courses to create a manageable workload and provide consistent feedback to the students. We have not identified a "best" solution, and not all of these solutions will work in every context. As such, we have provided descriptive case studies of working grading strategies in a variety of contexts: introductory physics labs, introductory circuits labs of different contexts, a design of database course, and required upper-level labs in materials and biomedical engineering. Each case study is summarized in Table 1, followed by details of each case study.

Table 1 - Overview of Case Studies

Case study	Approach	Course Context
Introductory Physics Lab	• TA training for GradeScope • TA Rubric training • Regular staff discussion with online chat	• Grader calibration • Student resistance to standards-based grading
First year design lab in ECE	• Team projects assigned mentors • Rubric feedback given by mentor; grades assigned by a single graduate TA • Rubric grading in GradeScope	• Mentors provide feedback in tune with the project details • Having a single grader for all projects makes the process fair, consistent, and efficient
Second year circuits lab	Lab reports with rubric based grading that emphasizes knowledge, structure, process, and reflection	Rubric with quality criteria, clear specifications, and samples
Information Sciences database lab	• Rubric tied to learning objectives, standards-based grading • Feedback with revision and resubmission	• TA empowerment as mentors • Calibration across assignments more than across graders, who may oversee very different projects
Upper-level Biomedical Engineering lab	• Perspectives on auto-graded, TA-graded rubric-based grading, and peer review • Peer review as pre-feedback before more formal grading	• Some tools perform better for granular rubric scoring • Grader calibration

Case study	Approach	Course Context
Materials Engineering Upper-level lab	• TA or team of 2-3 undergrad staff grading a single assignment per semester • Structure of the class allows for various assignments active all semester long	• TAs grade the lab assignment they taught so aware of challenges during lab • Grader calibration • Grading turnaround time

Introductory Physics Lab
This introductory laboratory explores electrodynamics and wave phenomena through direct experimentation. Labs are designed to develop student shared decision-making and collaborative writing skills. This is a large (about 300-400 students) service course. The laboratory staff consists of 5 to 7 graduate teaching assistants (TAs), a faculty coordinator, and one or more undergraduate learning assistants (LAs). Each TA is responsible for facilitating 2 or 3 lab sections of 30 or fewer students. Because group reports are submitted for grading, each TA grades 10 or fewer reports per lab section. The faculty coordinator is responsible for training the TAs in the lab exercises, course tools, and grading standards.

Lab reports are submitted using GradeScope, a commercially available, collaborative grading tool. Each semester, the TAs receive course-specific training on GradeScope. Student work is evaluated with a rubric identifying the skills and standards for the course. Standards are matched to a point score in each of three categories. These categories are: scientific proposal completeness (following the instructions); scientific report completeness (following the instructions); mastery of the specific skill highlighted in the lab. Students are graded only against the standard in the rubric. The grading turn-around goal is one week after the lab is completed.

The coordinator builds all rubrics and enters them into GradeScope. Once all lab sections are complete, the entire team grades their students' reports. The team uses the following grading rules: only the faculty coordinator can change the rubric; only the rubric can be used to assign points; explanatory feedback must be tagged **on the student report at the point of error** if credit is reduced from full. Issues perceived as not fitting the rubric are brought to the attention of the team for discussion. Once a decision is made, the coordinator may update the rubric. When a group's work appears to fall between two rubric elements, the larger of the two options should be given unless the TA makes a compelling argument for the lower score.

Regular discussion is needed for the technique to work. The instructional team makes liberal use of Microsoft Teams, particularly the group chat. TAs rapidly communicate with each other and the faculty coordinator. This reduces grading delays. More experienced TAs have near-real-time opportunities to mentor novice TAs as part of their professional development.

Using a rubric and GradeScope, together with rapid, near synchronous communication, improved grading consistency across the team. This does not mean that every grader has the same average grade per section. What it does mean is that similar skill deficits are identified and graded with a more consistent outcome. Students can review their work with comments, and the coordinator can review all work for consistency.

Students should take the feedback given on their prior reports and apply it to the next report. Unfortunately, it can also take three or four reports before feedback, action, feedback scheme begins to show marked improvement in student reporting. This can frustrate student groups, particularly those who are struggling.

The improved consistency and collaboration among the grading team members does not relieve all student frustration. We observe students can be resistant to standards identification as the metric of score determination. Interpreting a skill standard is challenging for some students, particularly if they are highly socialized to understand correctness or incorrectness against specific errors (e.g., calculation errors, spelling errors, etc.) rather than mastering an important skill (e.g., following instructions, working together, showing their work, clarity of presentation, etc.).

First-year design laboratory in Electrical and Computer Engineering

This first-year course is a design *elective* for first-year students co-enrolled in either of our two first-year *core* ECE courses. Enrollment is typically 50-60 students each semester. The students work in teams of two or three students and the majority of graded submissions are team submissions using the GradeScope online grading platform. Each team is assigned a mentor who is often an undergraduate assistant member of the staff but might also be a graduate TA. The course meets for two hours each week where the instructor first reminds the students of upcoming course deadlines, rubric details, and/or a short tutorial on a technical topic or skill. After the instruction, the students move into the laboratory space where they continue circuit construction and testing. Outside of class time, students complete journal entries where they document the progress, pivots (either in project goals or design), measurements, and validation of sub-features of their project. Weekend open labs are popular for teams seeking more time for construction and testing under staff supervision, especially as the semester and the projects progress.

In any semester, the number of team projects typically exceeds 20, and the open-ended design nature of the course makes it difficult for any one person to be intimately familiar with the details or progression of each-and-every course project. For this reason, it is most efficient for the mentor to make a first pass on the GradeScope grader, visiting their own projects to insert feedback in text form. The feedback is based on the rubric for that assignment which was made available to the students in advance and will also be used for grading. Expectations for the current assignment are discussed in a one-hour weekly staff meeting to help the mentors focus their feedback on the progression the team should have demon. The mentor feedback will highlight the positive progress of the team as well as alert the team to any deficiencies in their approach, be it in the technical details of the solution, in the process of test and validation, or in the reporting of their activities. The mentor does not, however, enter or comment on grades. One graduate TA (typically deemed the "head TA") then uses the GradeScope grader to read the comments left by the mentor on each project, search for attributes of the submission possibly missed by the mentor and then assign grades via the rubric built into that assignment. By leveraging the feedback provided by the mentors, the TA can more quickly and confidently review the work done by the team and assign grades across the entire course that maintain consistency.

The rubric is designed to provide clear expectations to the students and guidelines for the TA responsible for grading to assign the score for each rubric item. For the weekly journal report, the rubric has categories for Progress, Challenges, Plans, and Teamwork. Each category is defined in the rubric as well as examples provided for what achieves a score of Promising, Within Expectations, or Falling Behind which are mapped linearly to points on the grading scale.

Category	Explanation	Promising	Within Expectations	Falling Behind
Progress (8pts)	Has the project progressed in a satisfactory pace since the last entry? Evidence of progress should include research, schematics, circuit simulation, circuit construction, and/or circuit testing. Artifacts such as data, photos, and videos should support the details of the work. Have you reached any milestones? - Block diagram (2pts) - Circuit diagrams (2pts) - Bi-weekly Verification (2pts) - Detailed Description (2pts)	Progress has been made that supports evidence that the students are working as a team both during their scheduled lab time *and* outside of the scheduled lab time. Project is either ahead of schedule or has pivoted in ways that have made the project much more viable. The journal includes all the required components (block diagrams, circuit diagrams and verification graphs)	Progress has been made that supports evidence that the students are working as a team both during their scheduled lab time *and* outside of the scheduled lab time. The journal only included partial required diagrams/graphs.	Progress is slow suggesting that the team or some team members are not taking the project seriously. For students/teams who struggle to comprehend the task, additional time may be required to meet satisfactory progress.
Challenges (4pts)	When substantial progress has been made, it almost guarantees that the "easier" tasks have been accomplished to the point that a momentary roadblock has been reached. - Clear explanation of the problem (2pts) - Steps taken to tackle the problem and future actions (2pt)	A meaningful roadblock has been identified and possibly solved through research or consultation with the course staff.	The challenges mentioned are vague, futuristic, or trivial, suggesting the project goals are weak or that slow progress may have resulted in a delay in encountering such challenges.	Challenges remain unidentified suggesting some serious issue with the project. This can be quite unique, and the staff will provide feedback.
Plans (2pts)	Given where you are, what are your plans for the next two weeks? How does it compare with your original timeline? Have you adjusted any milestones?	A plan has been laid out that adjusts for the team's pace and for the pivots from the original timeline and milestones. Both the original timeline and adjusted timeline are clearly presented.	A plan has been suggested, but it is unclear how the plan compares to the original and/or it is unclear how that plan has taken the team's pace and abilities into account. Other details lacking, such as milestones.	What's a plan? [Answer: It's that thing that keeps you from failing!]
Teamwork (2pts)	Each teammate should have a primary responsibility each week. That should be evident in both the work just finished and in the plan for the upcoming weeks. [Solo projects may skip this section.]	Individual contributions and responsibilities are clearly established.	Individual contributions are enumerated, but unclear as to why or what. Example of weak statement: *Teddy will work on the amplifier and Sally will work on the logic.*	Sometimes, a weak teamwork statement is followed by clear evidence that the teammates are actually working solo…That is, sometimes one person doesn't really understand what the other is doing. That's not good.

Figure 1. Rubric for ECE first-year design

This method of grading is efficient in that the staff with the closest connection with the team can provide feedback to both the students *and* the grading TA accelerating the grading process. Having a single person assign grades makes the process fair and consistent. The weekly staff discussions assist in training such that the feedback and assigned grades match the instructor's expectations for the course.

Second-year circuits lab for non-majors

Next, we consider a sophomore-level service course designed to introduce fundamental circuit concepts to non-ECE majors. The course carries 3 credit hours and includes a laboratory component that meets every other week. Lectures meet three times per week for 50 minutes each.

Given the course's credit-hour structure, its role as a service course, and its sophomore-level standing, it is important that students do not feel undue pressure and are provided with appropriate guided support as they build foundational knowledge and problem-solving skills.

In order to earn full credit for a laboratory, students must (i) successfully complete the in-lab work and receive sign-off from the teaching assistant, and (ii) submit a lab report that includes all elements specified in the report instructions. All submissions are made on Canvas.

The lab report is highly structured, with clearly identified sections that guide students through the expected content. In addition, students are provided with a detailed grading rubric that clearly outlines expectations and point allocations for each section, as shown below.

- **Pre-lab (0.75 points)**
- **Title, Names and Dates (0.25 points)**
  - the title may be copied from the title of the lab, provided to you
  - all of the names of the team members who worked on the challenge must be listed
    - if your name is not on a report, you will <u>not</u> receive credit.
    - You must put a date of the report reflecting the date that the lab was done, as well as the date that the report was written....
  - Statement of Purpose
    - this section may be copied from the description of the lab, provided to you in the lab "challenge" (literally, copy+paste, if you wish)
- **Plan (1 point)**
  - this section should include schematic diagram(s), simulation(s), and any other information relevant to put the rest of the report in context
- **Execution (1 point)**
  - this section must describe what you did, and what happened
- **Results/conclusions (3 points)**
  - This section must answer the question(s) posed in the lab "challenge", and explain your understanding of what happened and why
  - For every report, you should demonstrate your understanding by answering the following:
    - What did you do?
    - What did you learn?
    - Why was it important?

In addition to including all required sections, each laboratory report is evaluated using quality criteria that apply to the document as a whole. These criteria are also given to students. Reports must be readable, with clear, typed text and legible figures; coherent, with logical organization and a clear flow of ideas; complete, containing all required elements and sufficient detail for another student to reproduce the work; and correct, free from mathematical errors and inaccurate or misleading technical statements. A sample lab report is provided to assist students in understanding the expected structure, level of detail, and quality of technical communication.

Detailed feedback is given to help students identify areas of improvement. Samples are given in Table 2 below.

Table 2. Examples of feedback to students

Plan (-0.25) • No diagram of setup Overall the plan needs more work, please describe the included figure and explain your plan a bit more clearly. The last sentence of the plan doesn't make sense, kind of a run on sentence. Good results otherwise, make sure you are answering all report question though.	**Missing pre-labs (-0.25)** Good results overall! Focus on the formatting of your report for next time, for example keep the font used consistent. You should also preferably typeset your equations instead of showing a hand drawing (fine to handwrite in the prelab, but not the report). Also, for your conclusion you don't need to write as much as you did, a short paragraph that answers what you did, what you learned, and why the enrolment was important will suffice.
Results (-0.2) Did not answer why the dropout frequency is higher for sine than square wave.	**Conclusion missing (-0.2)**

Information Sciences database lab

IS 206 is a sophomore-junior targeted required course with around 130 students enrolled, an instructor and 4 graduate student Teaching Assistants. IS 206 has a multi-week group project where students use Human-Centered Design and database design to create a database. The students describe the narratives, user stories, and then go through database design processes to eventually develop and populate a small database that addresses their user needs.

The weekly assignments build on each other as the students develop and revise their design. This project uses specification-based grading where each assignment in the project has around 4 learning objectives to be demonstrated in the work. Each learning objective has a set of criteria that further describe components of the objective. An example of an objective is:

Translate ERD to relational schema

- Every component of the ERD is included in the schema
- Includes 5+ entities, attributes, and 3+ relationships
- Schema is text in the format taught in the class, follows the flow chart
- 1-2 mistakes: Approaching, 3+: Beginning, ERD or schema missing: No evidence

An example rubric is shown in Table 3, below [16].

Table 3 - Example Specifications Grading Rubric

Rubric Level	Score	Accompanying feedback	Next actions
Satisfactory	4	Encouragement or acknowledgement of specific components.	None

Approaching	3	Feedback that describes the error or omission that's stopping the group from getting a Satisfactory.	Student should be able to resolve from the feedback without further assistance
Beginning	1	Feedback that points back to the assignment requirements or other course materials, including references to office hours or other support channels.	Student should contact the course staff to ensure understanding and objectives.
No Evidence	0	Minimal	Either the work should be submitted, or the student should seek clarification from the assignment on why this was not credited.

Grader feedback is required for each assignment. In Canvas SpeedGrader, this is done in annotations of the student documents or in the whole assignment feedback. This is important to identify the grader as these marks in Canvas give student-visible names connected to the feedback (in-rubric comments do not). Students then know who to follow-up with about their grading.

After feedback is released to the students, they have a one-week window to revise and resubmit (to the same assignment in Canvas, with extended due dates) to have their work reevaluated by the same original grader. The grader reviews the changes, updates the score, and adds a comment to acknowledge the regrade.

The simplified rubric speeds up grading. By looking at holistic outcomes and not necessarily per-point components, students and graders are focused on demonstrating learning objectives. The complexity of calculating "is this an 87% or an 90%" from traditional grading, for example, goes away. There are no "-0.5% for each mistake" scores. If the grader understands the work to be at a particular holistic level, they can comment on what earned at that level and score it. Graders calibrate based on examples and discussions in course staff meetings.

Consistency of graders across student submissions and across assignments in the project puts the responsibility and impact of feedback on the graders. If the grader does not feel the work (yet) demonstrates satisfactory levels, it is the grader's responsibility to provide instructive feedback to help the student improve a future revision. This changes the nature of grading from something that seems punitive or reductive in traditional grading to something more akin to mentoring. Students may inquire about their feedback in Canvas comments, in-class, or during the grader's office hours. The communication and iteration involved in this approach leads to relationship-building between the student and grader in a way that is often missing from traditional feedback methods. Students benefit from this empathetic connection, both in the quality of the final work and in the sense of connectedness that they can achieve satisfactory scores when someone is there to help them.

Revise and resubmit also empower the TA to give authentic feedback without concern about punishing the students' grade. In the authors' experience, graduate students struggle with wanting thoughtful critical engagement with students but are concerned about grade impacts. By framing the scoring around feedback, this empowers the TA as a mentor whose goal is to get the project to the satisfactory level.

Using a revise and resubmit window encourages communication between the students and course staff, focuses on feedback and what's needed to (still) reach objectives, and focuses on the growth and output more than the score. This works well for project-based assignments where different groups may be difficult to compare to each other (i.e. across grader calibration) versus scaffolding a fixed grader through the development of the project emphasizing growth mindset for the students and graders as a mentor for the TAs.

Required Upper-Level Biomedical Engineering Labs
This case study is split into two perspectives: a teaching assistant and the instructor of record. These are important viewpoints to consider when picking tools and working toward consistency.

A Teaching Assistant's Perspective on Quizzes and Lab Reports
Canvas was used in a third-year core Quantitative Physiology lab course to promote uniform grading across homework, quizzes, and lab reports. In-class quizzes were auto-graded using Canvas Quizzes, with full/no credit for single-answer questions and prorated scoring of questions with multiple correct responses. Time limits and accommodation-based extensions were implemented through the same feature. The Quiz feature also allows the instructor to designate time limits and add time-based accommodations.

Lab reports and projects were graded in Canvas SpeedGrader using structured, rubric-driven assessments to improve transparency and inter-grader reliability among Teaching Assistants (TAs). Assignments were broken down into standardized sections aligned to learning objectives (e.g., introduction, methods, results), with further sub-items such as results represented through figures and textual analysis of outcomes. This reduced ambiguity and enabled consistency across graders. Rubrics were reusable across courses and semesters within Canvas, allowing returning TAs to build grading intuition over time while reducing retraining overhead.

To further support consistency, TA calibration sessions were held before major grading events. The instructor and TAs reviewed representative submissions to align expectations around rubric interpretation, point allocation, and examples of partial credit. This process helped surface discrepancies, establish shared norms for subjectivity (e.g., quality of interpretation or clarity of conclusions), and reduce score variance once grading began. Calibration was particularly important for lab report sections with variable point distributions (e.g., a range of 5–25 points), where scorer discretion is inherently higher. For example, TAs frequently encountered variation in how students reported and interpreted statistical outcomes in the Results and Discussion sections. During this discussion, TAs jointly defined criteria for full versus partial credit responses (e.g., appropriately labeled figures with error bars, accurate reporting of test statistics, and conclusions supported by data). This helped to standardize deductions, ensuring that similar quality work was rewarded consistently across sections and graders. (Continuity of TAs across semesters, when possible, also alleviates these concerns.) After grading, section averages were compared to ensure no statistically significant differences.

Several practical lessons were learned through iterative use of digital grading tools. First, SpeedGrader was most effective for low-point, question-based assignments, as it enabled TAs to annotate directly on the submissions and utilize a shared comment library that could be accessed by all graders. Second, rubrics that explicitly define full, partial, and minimal credit reduced drift in grader interpretations. Finally, for complex, high-point lab reports and project, other platforms such as GradeScope occasionally proved better suited due to its rubric-based partial credit system that supported consistent application of deductions across multiple TAs and semesters. Canvas assignments can be linked to GradeScope when needed.

An Instructor's Perspective on Peer Review in Canvas

Canvas was used again during a class period in the same course for a group-to-group Peer Review of each other's draft lab reports. This allowed time to revise and resubmit before being graded, and prepared students for expectations before submitting later assignments with heavier weights.

A draft was uploaded to Canvas by one person in the group. Peer Reviews can be assigned between groups manually by the instructor or automatically by Canvas, for online or paper assignments. The instructor assigned the reviews manually and individually (within each section) so as to maximize the number of peers reviewing each report. For example, each report (for one group) was assigned and reviewed by 3 non-group members in that section, though each reviewer could not identify the others. Assignment rubrics were visible to students, with point selection and feedback tools available. Finally, students formed groups to provide brief in-person feedback. Four days were given for revisions. Graders were able to compare initial versus final submissions within SpeedGrader.

Students responded positively to the Canvas Peer Review exercise, with 82% stating "I liked it, and I think it works well." Reported benefits included reduced anxiety around feedback from peers and standardization of expectations across students as they were forced to critically evaluate peer work, eventually leveling out consistency in grading. However, a larger class size may prove more time-intensive for assignment of reviews, and some students had trouble locating reviewer comments within Canvas.

Collectively, these practices improved grading consistency, reduced TA uncertainty, and supported scalable course delivery. The strategies described here reflect practical insights gained through repeated implementation and may be readily adopted or adapted by other BME programs managing multi-TA instructional environments.

Materials Engineering Required Upper-Level Labs

MSE 307 and 308 are the junior-level labs in Materials Engineering. In the context of other Materials Science departments, these are large classes with between 80-100 students, although this is small compared to other departments within the College of Engineering (e.g., Civil, Computer Science, Electrical, or Mechanical). The combination of both classes fulfills the Advanced Composition requirement for graduation for our students, so students are required to complete multiple lab reports over the course of the semester.

Given the class size and the complexity of the experiments – students learn to use equipment such as the differential scanning calorimeter and universal load frames – experiments run in parallel, with about one third of the students rotating through a given experiment at a time. At any given point in the semester, lab reports are being submitted on three different topics simultaneously. This is more than one person can handle, so grading is divided between members of the teaching staff (instructor, graduate teaching assistants, and senior undergraduate teaching assistants). Grading is usually done in either the native Canvas interface or with a program such as Adobe Acrobat to mark up the reports and provide comments.

As has been emphasized multiple times, it is important to maintain consistency with multiple graders. To do so, several policies are in place within MSE 307 and 308. First, each graduate TA is assigned only one experiment to grade for the semester, which also happens to be the same experiment they taught in lab. That way, they have firsthand knowledge of any issues that might have arisen, and the same standards are applied to all reports on a particular topic.

With undergraduate graders, they usually work in groups of 2-3 to grade a particular experiment, so it's a little more challenging to ensure complete consistency for a given report. In this case, it is critical for the graders to coordinate, agreeing on the number of points to deduct for a given issue, what to look for in a given rubric item, etc. Some of this onus lies on the graders, although the instructor does her best to encourage both an open discussion of the rubric prior to grading as well as group grading of the first few reports.

Training and resources are also key to successful, consistent grading. Teaching assistants are provided with a copy of the rubric, a sample breakdown of points used previously (which they are welcome to modify), sample graded reports from previous semesters (high, medium, and low scores), as well as access to lecture slides and videos. The instructor is also available at any time to answer questions, particularly when faced with a grading issue the teaching assistants are unsure how to handle.

With all grading done manually in the class (and recognizing that each grader may be reading up to 30-35 reports at a time), it can be difficult to ensure timeliness of grading and feedback, which is critical for student growth and improvement. Graders are strongly encouraged to provide feedback within 1-2 weeks of submission, giving the students opportunity to make changes before their next report is due and ensuring that reports do not stack up too much for the grader (new reports are submitted every two weeks). However, it depends on the instructor to enforce such a timeline.

It is important to recognize that this method is not perfect, and some issues still arise. The most common is inconsistency between reports. For example, although all the Optical Microscopy reports were graded by GRADER 1, and all the Thermal Stress and Strain reports were graded by GRADER 2, GRADER 1 and GRADER 2 may have different philosophies, particularly for the "technical writing" portion of the grade (grammar, use of language, etc.). Grading the writing tends to be more subjective than the technical data portion and is the hardest portion to successfully calibrate between graders. These differences are usually reflected by differences in average scores; however, *for a given report,* there was still consistency. When questions or concerns do arise about grades, students are instructed to first reach out to the specific grader for an explanation of why they lost points, but if they still have questions, the course instructor is also available for mediation.

Discussion
On our campus, Canvas is the primary supported learning management system (LMS). GradeScope is also supported and integrates with Canvas. Therefore, these are the two most common tools used in these cases. The implementation details among the courses, however, are important. GradeScope is popular because it helps large teams easily apply the same comment for multiple students, apply rubrics consistently, and allow team grading supervision. However, for this paper, rubric-based grading in Canvas's SpeedGrader is also sufficient.

Facilitating instructional team communication is also important. Group chat tools are commonly used for this purpose. Popular choices are Discord, Slack, and Microsoft Teams. These platforms are generally selected based on ease of use and availability. These tools provide an easy way for the course staff teams to communicate and ask questions between preparation meetings and class sessions.

Graduate TAs on our campus are required to attend at the very least a one-time workshop prior to their first semester teaching; this workshop typically has sessions related to teaching and grading. In some departments, they are also required to take an additional 1-credit seminar course on teaching and leadership. TAs thus start with at least a limited background on pedagogical techniques and grading, but often require additional resources and training. For those with more extensive preparation and experience, it's possible to start with course-specific training from the instructor

In general, the electronic grading tools listed above contributed to efficiency because they offer reusability of comments among graders and consistently applying rubrics to student work. These tools allow for simultaneous grading and sharing of graded work within the team. Familiarity with the student work is also mentioned to help with efficiency in a couple of cases, although in each case the grader is familiar with the student work through different types of interactions.

Consistency and coordination from a faculty member or other lead instructor are important for consistent, equitable grading as described in these cases. Communication within the team occurs through TA calibration activities, group grading, meetings and chats, or shared feedback databases or a combination of these methods. Asking for clarification and discussing deviations as soon as possible provides ways for teams to reach consensus. However, if that is not possible the final decision should rest with the faculty coordinator. Documented, detailed feedback is a means of communicating within the team and to the students. The feedback provides more context that can help explain why points were taken off and what should be improved in the future.

All cases described use rubrics. Rubrics provide a framework for expectations and feedback about the assignments and can be tailored to the assignment goals or kept generic to course level learning objectives. The interpretation of rubrics by students and graders can lead to confusion, inconsistent scoring, and conflict. All cases included a process for calibrating the graders to understand the rubric for consistent outcomes, but different cases have different recommendations for what worked for them and different methods of reconciling feedback with the students.

These cases review the supportive, collaborative techniques used to facilitate consistent, timely and equitable grading in their courses. This is an important goal, however, it does have its issues. Collaborative grading and rubric usage can be confusing at first to TAs. Some feel that their autonomy is compromised by using the standards set by the lead instructor. Training for grading consistency also takes time and can be frustrating. Collaborative grading tools also don't solve issues related to submission inconsistencies by students nor do they resolve all student-related complaints about grading.

Conclusions

We have presented several cases of successful grading schemes for large undergraduate laboratory courses in STEM. The instructors in each case have continuously improved their schemes to work for them, their staff, and the students in their courses. One tool that is used in every case is a detailed rubric with training processes to make it effective. Our goal in sharing these cases is to highlight these successes so that other instructors can learn from our schemes so that they can adapt and apply them in their courses.

References

[1] M. D. Svinicki and W. J. McKeachie, *McKeachie's Teaching Tips: Strategies, Research, and Theory for College and University Teachers*, vol. 14. Wadsworth Cengage Learning, 2014.

[2] Joe Feldman, *Grading for Equity: What It Is, Why It Matters, and How It Can Transform Schools and Classrooms*. Thousand Oaks, CA: Corwin, 2019. [Online]. Available: https://crescendoedgroup.org/grading-for-equity/

[3] S. A. Atwood, M. T. Siniawski, and A. R. Carberry, "Using Standards-based Grading to Effectively Assess Project-based Design Courses," presented at the 2014 ASEE Annual Conference & Exposition, Jun. 2014, p. 24.1345.1-24.1345.10. doi: 10.18260/1-2--23278.

[4] T. M. Fernandez, K. M. Martin, R. T. Mangum, and C. L. Bell-Huff, "Whose Grade is it Anyway?: Transitioning Engineering Courses to an Evidence-based Specifications Grading System," presented at the ASEE Annual Conference and Exposition, Conference Proceedings, 2020. doi: 10.18260/1-2--35512.

[5] R. M. Reck, "Adopting Alternative Grading in an Upper-Level Laboratory Course in Bioengineering," in *2022 IEEE Frontiers in Education Conference (FIE)*, Oct. 2022, pp. 1–5. doi: 10.1109/FIE56618.2022.9962684.

[6] A. Powe and J. Moorhead, "Grading Lab Reports Effectively: Using Rubrics Developed Collaboratively By Ece And Technical Writing Instructors," presented at the 2006 Annual Conference & Exposition, Jun. 2006, p. 11.669.1-11.669.17. Accessed: Jan. 21, 2026. [Online]. Available: https://peer.asee.org/grading-lab-reports-effectively-using-rubrics-developed-collaboratively-by-ece-and-technical-writing-instructors

[7] M. Christopher, "Use of an Automated Grading Circuit for a Lab-based Course," Columbus, Ohio, Jun. 24, 2017.

[8] A. F. C. Pinzón, J. J. Ramírez-Echeverry, and F. Restrepo-Calle, "Automated grading software tool with feedback process to support learning of hardware description languages," *Res. Pract. Technol. Enhanc. Learn.*, vol. 19, pp. 015–015, Jan. 2024, doi: 10.58459/rptel.2024.19015.

[9] Brittany Ann Kos and Sarah Miller, "Grade-a-thons and Divide-and-Conquer: Effective Assessment at Scale," Columbus, Ohio, Jun. 24, 2017.

[10] S. Ciston, C. Cerretani, and M. S. Went, “Teaching with Graduate Teaching Assistants: Tips for Promoting High Performance Instructional Teams,” presented at the 2016 ASEE Annual Conference & Exposition, Jun. 2016. Accessed: Jan. 21, 2026. [Online]. Available: https://peer.asee.org/teaching-with-graduate-teaching-assistants-tips-for-promoting-high-performance-instructional-teams
[11] M. A. Verleger and H. A. Diefes-Dux, “A Teaching Assistant Training Protocol for Improving Feedback on Open-Ended Engineering Problems in Large Classes,” presented at the 2013 ASEE Annual Conference & Exposition, Jun. 2013, p. 23.121.1-23.121.12. Accessed: Jan. 21, 2026. [Online]. Available: https://peer.asee.org/a-teaching-assistant-training-protocol-for-improving-feedback-on-open-ended-engineering-problems-in-large-classes
[12] J. Brocato, J. Harden, and B. Chapman, “Improving The Writing Evaluation Abilities Of Graduate Teaching Assistants In ECE Labs,” presented at the 2005 Annual Conference, Jun. 2005, p. 10.734.1-10.734.6. Accessed: Jan. 21, 2026. [Online]. Available: https://peer.asee.org/improving-the-writing-evaluation-abilities-of-graduate-teaching-assistants-in-ece-labs
[13] S. Franey, A. Gregerson, and M. T. Braun, “Playing the TA Lottery: A Study of How Teaching Assistants Impact Grades in Engineering Courses,” presented at the 2012 ASEE Annual Conference & Exposition, Jun. 2012, p. 25.1046.1-25.1046.10. Accessed: Jan. 21, 2026. [Online]. Available: https://peer.asee.org/playing-the-ta-lottery-a-study-of-how-teaching-assistants-impact-grades-in-engineering-courses
[14] C. Dylan, B. Daniel, B. Maeghan Marie, M. Paula Ann, A. N. Kevin, and N. M. Audra, “Student and Teaching Assistant Perspectives on Characteristics of an Effective Teaching Assistant,” Seattle, Washington, Jun. 2015.
[15] R. M. Reck, “The influence of teaching assistants in an undergraduate engineering laboratory course,” in *2017 IEEE Frontiers in Education Conference (FIE)*, Oct. 2017, pp. 1–6. doi: 10.1109/FIE.2017.8190679.
[16] Robert Talbert, “EMRN Rubric,” Robert Talbert, PhD. Accessed: Jan. 20, 2026. [Online]. Available: https://www.rtalbert.org/emrn