

# Piecewise Grading Reform Year One: Lab Specifications

Joshua P. Veazey, Grand Valley State University, Allendale, MI 49401  
[veazeyj@gvsu.edu](mailto:veazeyj@gvsu.edu)

Traditional, points-based grading systems make it challenging to communicate that learning is a complex process, and that mistakes are opportunities for growth. In fact, evidence suggests that students *knowingly* prioritize maximizing points over maximizing learning<sup>1</sup>. Over the past several years, different forms of alternative grading have grown in popularity in higher education. Alternative grading can take the form of standards-based mastery grading,<sup>2,3</sup> specifications grading,<sup>4,5</sup> ungrading,<sup>6</sup> or hybrids of these.<sup>7</sup> While specifics within alternative grading systems vary among instructors, successful grading systems share certain features. Clark and Talbert describe the ‘Four Pillars’ of alternative grading<sup>7</sup>: (1) clearly defined standards, (2) helpful feedback, (3) marks indicate progress, and (4) reassessment without penalty. Believing in the advantages of alternative grading systems is easier than implementing them. Much of the advice in published accounts of alternative grading systems suggests a significant investment of time. The wide scope of issues to consider in course design<sup>8</sup>, not to mention the need to anticipate and address student resistance to unfamiliar expectations<sup>9</sup> could require preparation months before the class begins.

Today, I regularly implement alternative grading systems throughout my undergraduate courses. However, during my first iteration six years ago, I felt somewhat overwhelmed at the idea of a complete overhaul. To manage the transition, I decided at the outset that I would slowly convert different pieces of one of my courses over several years. By the end of the third year, all areas of my course had been converted to a hybrid standards and specifications-based grading system, details of which are described elsewhere<sup>7</sup>. In this paper, I discuss the details of Year One of the conversion, where I assessed the lab sections with specifications-based grading. Splitting the full conversion over three years allowed me to spend time facing the learning curve associated with transitioning from grading with points to fairly grading with specifications and allowing for revisions. In addition, the time helped me focus on evaluating one set of specifications for one course category at a time.

The course I describe here is a second-semester introductory electricity and magnetism course (with optics) targeted at engineering majors. Enrollment over the five-year span varied from 40-60 students. The course consisted of two 75-minute lectures per week, and one two-hour

hybrid lab/discussion. Each lab/discussion section was capped at 20 students. Prior to the conversion to specifications grading (Year 0), the course used a traditional, points-based grading system with final course grade determined by the weighted average of scores in labs, online homework, metacognitive quiz reflections,<sup>10</sup> three midterm exams, and one final exam. In Year 0, there was little explicit incentive to engage with instructor feedback, even if (perhaps especially if) grades were high. Once students received their graded work with my feedback, I almost never had another conversation about it, unless it was about point grubbing rather than substance. I needed to find ways to incentivize engagement with instructor feedback, shifting the focus from accumulating points to accumulating knowledge and skills.

In the first year of the transition to alternative grading (Year One), I applied a specifications-based grading scheme to just the lab/discussion section of the course. The specifications were crafted in response to my concerns about the quality of work and student affect for these assignments in Year 0. More emphasis was placed on the extent to which the student demonstrated expert-like experimental skills, while correctness was just one component of assessment. Verification labs have been shown to have no added value over in-class demonstrations for learning content<sup>11</sup>, and assessing students on correctness can promote ethically questionable experimental practices<sup>12</sup>. The lab activities during this conversion were still verification labs, but the specifications focused on communication and process skills rather than achieving a predetermined result.

### **Lab Specifications**

The first pillar of alternative grading calls for clearly defined standards. Table I shows the lab specifications implemented for Year One of the transition (the wording matches the most recent iteration of the course, but it was functionally the same). Students submitted a lab packet with experimental data and responses to summary questions. Each lab was essentially graded pass/fail, using satisfactory/progressing terminology.<sup>13</sup> Students needed to meet all three specifications to receive Satisfactory credit for the lab.

The first specification requires clarity and communication in submitted work, regardless of whether work is correct. The clarity specification communicates to the student that their reasoning is just as important as their final answer. It also assures that the instructor intentionally assesses student reasoning as well, rather than neglecting reasoning even when a correct answer

is achieved in an incorrect way.<sup>14</sup> Sometimes the problem will prompt clarity with the reminder, “Explain your reasoning.” Other times, understanding can be communicated with a clear sketch. I suggest that students imagine a classmate who does not understand how to do the problem. The clarity of their work should be such that this hypothetical student could at least understand how the problem was solved, even if they could not reproduce it yet.

An experimenter should be engaged with their work enough that they are assessing how reasonable the data are as they go along. The second specification states that the data and results of analysis must be plausible. For example, in one experiment students used electromagnetic induction data to determine the magnetic field of a small horseshoe magnet. A common incorrect analysis of the data resulted in magnetic fields of 20 T. Students ought to recognize that this is not plausible for a small horseshoe magnet, and that something went wrong. They can often quickly diagnose it while still in lab. For instances where students may not be experienced enough with a particular measurement to know whether the data were plausible, instructors or written instructions can give guidance. In the magnetic field example mentioned above, the lab instructions include a couple of sentences giving students some reference points, like the magnetic field of an MRI (requiring superconducting coils), and the largest human-made DC magnetic field.

The third specification states that work should be “mostly correct.” Typically, this corresponded to 80% correct, however most weeks I identified at least one individual question that was important enough that it ought to be correct to earn a satisfactory mark. Students should be made aware of these key problems in accordance with the first pillar. Finally, if the lab was not turned in or if it was missing any data or responses to questions, it was scored Incomplete.

The second pillar of alternative grading is “Helpful Feedback.” Figure 1 shows a sample of student work for a full feedback and revision cycle. Following a lab activity where students sight a ray pathway through a convex semicircular prism, the summary problem asks students to sketch rays passing through a different semicircular prism—one that is concave. A typical incorrect student response is shown in Fig. 1(b), where rays are drawn bending as if they were also emerging from the convex prism (as in the lab activity). Instructor feedback for this response first clearly cites Spec. 3 (Correctness). The work also fails to meet Spec. 1 (clarity), as no explanation—written or visual (e.g., normal lines)—is given. Instructor feedback is beneficial

when it is growth-oriented, leaving the door open for the student's work to improve upon resubmission. For the student in Fig. 1(b) feedback is phrased as an opportunity for re-learning and practicing the skill. The feedback avoids negative phrasing that could discourage a growth mindset<sup>15</sup>.

Figure 1(c) shows an example of a student revision. Revisions of substance are required to include a metacognitive reflection with two parts,<sup>10</sup> (1) identify the specific error or problem, and (2) generalize beyond the specific problem to demonstrate growth. The student work in Fig. 1(c) meets both requirements. The result is that the student has demonstrated growth. Whereas their initial submission relied on attempting to recognize a surface-level pattern for how light rays bend in prisms, they now better recognize the actual process for analyzing the geometry of light rays in unfamiliar prisms.

At the other end of the feedback spectrum are minor student errors or missed specifications. In those cases, pure corrections of work by the instructor are still rare; rather, the feedback should be used to bring the student's attention to the error so that they can make sense of it themselves. For example, something as simple as forgetting units might be circled with the feedback "Spec 1 (clarity): include units." Another case is student response to a question with no reasoning communicated, with the quick feedback, "Spec 1 (clarity): Explain your reasoning." Students are reminded (especially cases like these) that receiving a grade of "Progressing" doesn't mean they "blew it" on the assignment. I explicitly tell them that it is sometimes a mechanism to call their attention to something important that they can fix in 5 minutes or less.

The opportunity for revision and resubmission created an incentive for students to engage with and internalize instructor feedback on the initial submission. Engagement with feedback from an expert is an important component of learning via deliberate practice. In accordance with the fourth pillar of alternative grading, I assess those resubmissions without any penalty. A student who earned Progressing or Incomplete on the first submission can replace that first mark with a Satisfactory mark. There is no averaging of the two submissions. Growth means the 'Satisfactory' has been earned. In this way, marks on assignments indicate progress (the third pillar of alternative grading). The student in Fig. 1(b) who has not yet learned how to analyze the geometry of a light ray's path through a prism gets full and complete credit for understanding it

by the time they resubmit their revision. Grades should not be punitive for not mastering a topic on the first attempt.

A policy to place some limits on the number of revision opportunities can help with transitioning to an alternative grading system. Limits are especially helpful for instructors assessing large numbers of students. In this course, students were limited to one revision per lab assignment in response to instructor feedback. Revisions had to be submitted within a week after receiving feedback. With experience, I found that grading student work in this way saved time when compared with giving comparable levels of feedback in a points-based system. I did not need to spend painful chunks of time determining how many points to assign to every fine-grained step or computation. Grading resubmissions typically did not take long at all. For the vast majority I could quickly see if missing specifications were addressed or not; it was not necessary to regrade the entire assignment. If experienced instructors begin to be flexible on the number of revisions allowed, it is good practice to be transparent about this to every student in the class to keep the grading system equitable.

### **Final Course Grades**

Since Year One used alternative grading in only labs, I needed a hybrid system for determining final course grades. Table II is a reproduction from the syllabus for how final course grades were determined. Of the 13 lab activities, students needed to reach a threshold of satisfactory marks to be awarded a final course grade. They *also* needed to reach a percent threshold in the weighted average column. The final course grade was set by the lower grade of the two categories in Table II ('Weighted Average' and 'Number of Satisfactory marks on Lab/Discussion').

### **Student Impressions**

Over five years of using the specification system for labs, student feedback is almost universally positive. After Year 1, one student wrote on their anonymous course evaluation:

*"The lab revision policy made me feel less pressured to allow the individual at the table who was the best at physics lead and place my own thoughts and answers on the labs. This, of course, led to a better understanding. The policy encouraged me to understand the material as opposed to just agreeing with someone at the table who is good at physics then just writing down what they have."*

Students understand the reasons for assessing labs this way, and the above comment reflects one of the shifts in student behavior that I observed. Year One did bring one critical comment that could serve as a caution for new adopters:

*“I think the lab revision policy was nice, but could be more clear on how correct the lab needed to be, most of the time it seemed like we needed 100% correctness...”*

This comment stems from the fact that I had internally flagged 1-2 problems or analysis questions on each assignment as “critical,” meaning students needed to understand them to move on. In the years following this Year One comment, I explicitly called these out to students before they submitted their revision. This change has also been more faithful to the first of Clark and Talbert’s Four Pillars: Clearly defined standards. As for standards-based grading of exams implemented later (Years 2-3), student impressions were initially more mixed. In addition, the workflow for grading re-assessments differed. The standards-based system required more refinement and optimization and will be discussed in a future paper.

## **Discussion**

Piecewise specifications-based grading allows instructors to phase in alternative grading for their courses over a few years. This gradual approach can help new adopters navigate the learning curve that comes with new assessment philosophies. In the first year of my own transition, I limited the specifications-based grading to the hybrid lab/discussion section of the course I taught. Others may choose to first adopt specifications to different areas of their course, such as standalone discussion activities or recitations. The four pillars of alternative grading described by Clark and Talbert provide a useful framework for diverse grading systems that can result in improved learning.

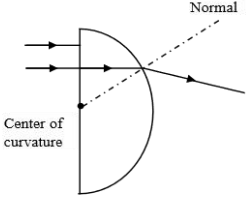
The focus on process and skills in the lab/discussion specifications described here helped create a psychologically safe and growth-oriented classroom environment. Students modified their behavior to engage more closely with instructor feedback. They routinely initiated in-person conversations with me about the feedback they received in an effort to improve. It seemed that they were no longer tossing graded work aside immediately after they got it back. What’s more, conversations with students in this safe atmosphere often led to frank discussions on theories of learning and motivation. Further expansion of standards and specifications-based

grading in later years<sup>7</sup> enhanced the course in similar ways, while keeping a manageable overall workload.

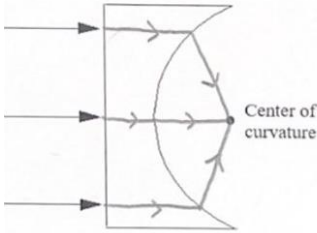
## References

- <sup>1</sup> A. Elby, "Another reason that physics students learn by rote," *Am J Phys* **67**(S1), S52–S57 (1999).
- <sup>2</sup> T. Zimmerman, "Grading for Understanding – Standards-Based Grading," *Phys Teach* **55**(1), 47–50 (2017).
- <sup>3</sup> I.D. Beatty, "Standards-Based Grading in Introductory University Physics," *Journal of the Scholarship of Teaching and Learning* **13**(2), 1–22 (2013).
- <sup>4</sup> L.B. Nilson, *Specifications Grading*, 1st ed. (Stylus, Sterling, VA, 2014).
- <sup>5</sup> K.A. Harper, "Grading Homework to Emphasize Problem-Solving Process Skills," *Phys Teach* **50**(7), 424–426 (2012).
- <sup>6</sup> S.D. Blum, *Ungrading* (West Virginia University Press, Morgantown, WV, 2020).
- <sup>7</sup> D. Clark, and R. Talbert, *Grading for Growth* (Routledge, New York and London, 2023).
- <sup>8</sup> E. Cilli-Turner, J. Dunmyre, T. Mahoney, and C. Wiley, "Mastery Grading: Build-A-Syllabus Workshop," *PRIMUS* **30**(8–10), 952–978 (2020).
- <sup>9</sup> J.S. Kelly, "Mastering Your Sales Pitch: Selling Mastery Grading to Your Students and Yourself," *PRIMUS* **30**(8–10), 979–994 (2020).
- <sup>10</sup> C. Henderson, and K.A. Harper, "Quiz Corrections: Improving Learning by Encouraging Students to Reflect on Their Mistakes," *Phys Teach* **47**(9), 581–586 (2009).
- <sup>11</sup> N.G. Holmes, J. Olsen, J.L. Thomas, and C.E. Wieman, "Value added or misattributed? A multi-institution study on the educational benefit of labs for reinforcing physics content," *Phys Rev Phys Educ Res* **13**(1), 010129 (2017).
- <sup>12</sup> E.M. Smith, M.M. Stein, and N.G. Holmes, "How expectations of confirmation influence students' experimentation decisions in introductory labs," *Phys Rev Phys Educ Res* **16**(1), 010113 (2020).
- <sup>13</sup> R. Talbert, "Specifications grading with the EMRF rubric," <https://rtalbert.org/specs-grading-emrf/>.
- <sup>14</sup> C. Henderson, E. Yerushalmi, V.H. Kuo, P. Heller, and K. Heller, "Grading student problem solutions: The challenge of sending a consistent message," *Am J Phys* **72**(2), 164–169 (2004).
- <sup>15</sup> C. Dweck, *Mindset: The New Psychology of Success* (Random House, 2006).

## Tables and Figures

(a)  **Lab Activity:** Sighting pathways of light rays entering a convex, semicircular dish filled with water. Sketch rays and determine index of refraction.

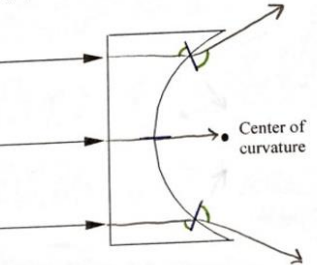
(b) **Summary Problem:** “Based on your lab exercise today, predict how the light rays will bend as they pass through the glass block shaped as shown below. Draw the path of the rays into the glass and back out into the air .”

**Initial Submission:** 

**Instructor Feedback:**

SPEC 3: How to be successful with these types of problems: First always draw the line normal to the surface where the ray is passing through. Then, remember that if the ray travels from a slower medium to faster medium, the ray bends away from the normal. If it travels from faster medium to slower medium, the ray bends towards the normal.

(c) **Revised submission and reflection:**



*The rays bend away from the normal. Before, I did not know how to use the normal to qualitatively find the direction.*

**Figure 1.** Sample revision cycle for a single lab problem. (a) The lab activity that was a precursor to the sample summary problem: sighting and sketching light rays incident on the flat side of a semicircular dish of water. The sketch was used to measure angles and determine the index of refraction. (b) The text of the summary problem at the end of the lab and an initial, incorrect student submission. The instructor feedback cites Specification 3 (correctness) and gives instruction on how to reason through the problem. The feedback is helpful and growth oriented. (c) Representative student revision. This student originally had the error of the type shown in (b). The ray diagram has been corrected, showing the normal line, as well as the incident and refracted angles. The student’s reflection first identifies the specific correction and then reflects on the general difficulty that led to the error in the first submission, demonstrating growth.



**Table I.** Specifications-based rubric for lab/discussion sections.

Score	Specifications
<b>Satisfactory (S)</b>	All three of the following specifications are met: <ol style="list-style-type: none"> <li><u>Clear</u>. All work is clear and legible. Physical reasoning is explained where appropriate.</li> <li><u>Plausible</u>. Experimental data are plausible, or there is an explanation for why they are not plausible, specifically what went wrong.</li> <li><u>Mostly Correct</u>. Most of the work is fully correct. Depending on the lab, "most" may be as low at 70% or as high as 100%. Key problems may be required to be corrected before credit is earned.</li> </ol>
<b>Progressing (P)</b>	At least one of the above specifications is not yet met.
<b>Incomplete (I)</b>	At least one question has not received a good-faith attempt or is unfinished.

**Table II.** Hybrid system to determine final course grades that incorporates specifications. The student must satisfy both columns to earn a grade; in other words they receive the lower of the two grades.

Grade	Weighted Average	Number of Satisfactory Marks on Lab/Discussion (out of 13)
A	> 93%	12 or more
A-	> 90	
B+	> 87	11 or more
B	> 83	
B-	> 80	
C+	> 77	10 or more
C	> 73	
C-	> 70	
D+	> 67	9 or more
D	> 60	
F	< 60	Less than 9