Student preferences for grade weighting in introductory physics

Thomas O'Brien (he/him), W. Brian Lane (he/him), J. Caleb Speirs (he/him), and Naomi Laird (she/her)
Department of Physics, University of North Florida, 1 UNF Drive, Jacksonville, FL, 32224

Grading practices are known to influence student behavior and learning, particularly in introductory physics courses. To study these influences, we must gain an understanding of how students think about grades with respect to the poles of effort and mastery that shape some popular alternative grading strategies. We administered a survey about grading to $N = 88$ introductory physics students, with a combination of free-response and numerical questions. The free-response questions generated feedback about students' experiences with stress related to grades and their perspectives on how grades represent effort and mastery. The numerical question asked about their preferences for how grades in a physics course weight their understanding of course material and their effort toward completing coursework. The numerical responses reveal that just over half our introductory students weighted effort more heavily than understanding. Filtering these numerical responses by students' discussion of stress, effort, and mastery reveals insights into their thinking about how grading is structured.

2025 PERC Proceedings edited by Pawl, Zwolak, and Leak; Peer-reviewed, doi.org/10.1119/perc.2025.pr.OBrien

I. INTRODUCTION: HOW GRADES AFFECT LEARNING

In their recent scoping review, Hackerson *et al*. [1] describe how research into alternative grading practices in STEM education lags behind the popularity of such alternatives. Research that has been published shows promise for the positive impacts alternative grading practices can have on student learning and attitudes, but this research lacks coherence in frameworks, terminology, and methodology. They also found that physics education research lags even farther behind grading-related research in other STEM subjects.

A primary theme in alternative grading practices is the delineation between the *effort* that a student puts into a course and their demonstration of *mastery* over course content. In traditional grading practices, these properties are mixed. For example, an instructor might grade reading assignments and class participation based on completion (effort), exams on correctness (mastery), and homework on a combination of the two. These assignment scores are then combined in a weighted average, such that the final grade is some linear combination of effort and mastery that has been collapsed into a single number. However, much like a vector's magnitude does not contain sufficient information to describe the vector's unique combination of components, this overall percentage is not sufficient to describe how much of its value came from effort, and how much came from mastery.

Some alternative grading strategies are designed under the presumption that mastery is of primary importance. For example, in typical implementations of standards-based grading, the final grade is determined by the number of standards or learning goals that the student has demonstrated mastery of divided by the total number of standards for the entire course [2]. Implementations vary in how this count is determined (e.g., a minimum number of demonstrations, or a rolling average). Typically, assignments do not receive individual scores, deadlines are flexible, and resubmissions are welcome. Because the course grade is directly tied to formative feedback based on the list of standards, this structure focuses the student's attention on learning outcomes and comprehension, supports a growth mindset and collaboration, and can be more equitable for certain groups of students [3]. Standards-based grading in a physics context has been documented as productive [3, 4].

However, while standards-based grading foregrounds learning as a process that students engage in, it might leave students feeling demotivated to engage in learning-related practices that do not directly impact their course grade, like participating in class or reading course material. Other alternative grading strategies foreground these learning practices under the assumption that learning happens when students spend time on appropriately designed learning activities. For example, in typical implementations of labor-based contract grading, the final grade is based on students' completion of assignments like homework, projects, presentations, or feedback conversations with the instructor [5]. A contract specifies how students earn a benchmark grade (usually B) based on completing a certain number of assignments, with subsequent terms specifying the thresholds for raising or lowering one's grade. Instructors still provide feedback on correctness and work quality but, as in standards-based grading, do not assign grades to individual assignments. This approach foregrounds learning processes over content mastery and acknowledges that grades are a form of currency that the instructor awards in exchange for the students' time investment. Typically, the contract leaves students some agency over which assignments they complete, thereby supporting student intellectual risk-taking, self-reflection, and self-regulation. The work required for a higher-grade threshold can take different forms and reflect student interest, which encourages student intrinsic motivation. The terms of the labor-based contract is open to negotiation during the course, which encourages open dialogue and reduces the transactional relationship between an instructor and student. Research has shown labor-based contract grading to reduce students' anxiety and stress, thus supporting mental health [5–8]. However, this approach might encourage students to disengage with the actual learning in favor of completing a checklist. Labor-based contract grading has been implemented in physics classrooms but this implementation has not been studied in the research literature.

Both of these approaches are based on strong pedagogical reasoning, and illustrate how separating effort from mastery is an important goal in the alternative grading movement. Even instructors who adopt a differently framed alternative grading strategy or who retain a traditional grading strategy could benefit from considering how these poles of grading are represented in their strategy. This framework can also help instructors and researchers parse student perspectives of grading in terms of effort and mastery. An understanding of how students think about effort and mastery can be used as a lens for understanding their responses to alternative grading strategies, how these strategies impact learning outcomes, and how these strategies support student well-being.

To help develop this framework, we administered a grading preferences survey to $N = 88$ introductory physics students. We use their responses to a set of numerical questions about grade weighting to answer the following research questions: How do physics students express their grading preferences in terms of understanding and effort? How might these preferences be informed by the way students think about grades in terms of stress, effort, and mastery?

II. METHODOLOGY

We administered the following survey about grading to $N = 88$ introductory physics students enrolled in algebra- and calculus-based Physics I and II at a mid-sized regional state university with R2 status. We collected responses near the end of the semester, from courses that used traditional grading strategies. The survey included a combination of free-response and numerical questions. The free-response questions (1-5, 7-8) solicited feedback about students' experiences with grades and the numerical question (6) asked students to provide their answers on three sliding scales.

1. When you think of grades, what comes to mind?
2. What emotions do you think grades make students feel in a physics course? Explain.
3. Think of a course you took where you felt the grading system was unfair and/or did not support your learning. What stands out to you about this course's grading system?
4. Think of a course you took where you felt the grading system was fair and supported your learning. What stands out to you about this course's grading system?
5. In a physics course, what do you think a course grade of B means? What do you think it should mean?
6. Suppose you had a choice in how physics instructors weighted students' understanding of the material (for example, the degree to which you "got the right answer") and their effort in the course (for example, the degree to which you "did all the work") in calculating the final letter grades. How would you prefer the instructor to weight understanding and effort? If you think something else should also be factored in, tell us what that is.
7. Suppose you were in a physics class where the assignments, tests, etc. had comments and feedback from the instructor but no numerical scores. How would this impact how you study in a physics course? What sort of comments and feedback would you like the instructor to provide so that you would know how you are doing in the course as well as what you needed to do differently?
8. Suppose you were in a physics class where the assignments, tests, etc. had comments and feedback from the instructor but no numerical scores. What would be a good way for your instructor to determine your final letter grade at the end of the semester?

Note that in question 6, we used the term "understanding" instead of "mastery" since that seemed more accessible to students. Our survey platform did not include an option to enforce normalized responses on question 6, such that not all students' percentages summed to 100%. We normalized the total percentages afterwards so that the understanding and effort percentages added to 100.

We carried out content analysis on the students' free responses to questions 1-5 and 7-8 to identify themes in their responses. These themes summarize what qualities students said grades represented (e.g., "Grades represent effort," "Grades represent mastery") and what effects students said grades produce (e.g., "Grades produce stress," "Grades produce a negative learning environment"). For the themes about qualities that grades represented, we further identified whether each student presented this idea in a positive light, a negative light, or in mixed terms.

Here, we focus on three important themes from students' free-response answers: "Stress is a negative outcome of grades" (Stress-Expressing). This theme indicates that a student referred to grade-induced stress or anxiety as a hindrance to the learning process. "Grades represent the effort a student puts in, and this is a good thing" (Effort-Affirming). This theme indicates that a student made a positive comment about a grade reflecting good-faith effort to learn material or complete an assignment. "Grades represent a student's mastery of the subject matter, and this could be a good thing or a mix of good and bad" (Mastery-Affirming, Mastery-Agnostic). This theme indicates that a student refers to a grade reflecting their demonstration of sufficient understanding of a topic or fluency with a practice, and they might frame this positively or in mixed terms. We carried out this thematic coding collaboratively by hand, reaching consensus by discussion. We used the student as our unit of analysis, coding whether a student mentioned each theme, not how many times or in response to which question. This process is further discussed in [9].

We choose to focus on these themes because they each disaggregate the student population nearly evenly (53% Stress-Expressing and 47% Stress-Silent; 47% Effort-Affirming and 49% Effort-Silent; 30% Mastery-Affirming, 29% Mastery-Agnostic, and 40% Mastery-Silent) and because they are important considerations when formulating a grading strategy.

III. RESULTS

Here, we discuss students' responses to question 6 about their preferences for how effort and understanding are weighted in a physics class. We study how these weightings were reported by the students overall and how these weightings vary between students who expressed ideas related to stress, effort, and mastery in their free responses. We use Kruskal-Wallis H-tests (a non-parametric alternative to ANOVA for comparing 3 or more groups) to evaluate whether each of these thematic groupings shows a significant difference in the students' weighting preferences, and we use Cohen's d to evaluate the effect size of significant differences.

All students. Figure 1a shows the weighting distributions reported by all students. There is a slight preference for weighting effort more heavily than understanding, with 51% of students weighting effort more heavily than understanding, 14% of students weighting effort and understanding equally, and 35% of students weighting understanding more heavily than effort.

Theme 1: Stress. In the free responses, we found that 53% of the students described grade-related stress as an explicitly negative experience. We disaggregated the grade weighting responses to investigate whether these Stress-Expressing students indicated different weighting preferences. Figure 1b shows the weightings from Stress-Expressing students (solid curve) and Stress-Silent students (dashed curve). The distributions are statistically indistinguishable, with a Kruskal-Wallis H-test revealing no statistically significant difference ($p = .50$). Therefore, it does not seem that a perception of negative grade-related stress explains how students prefer

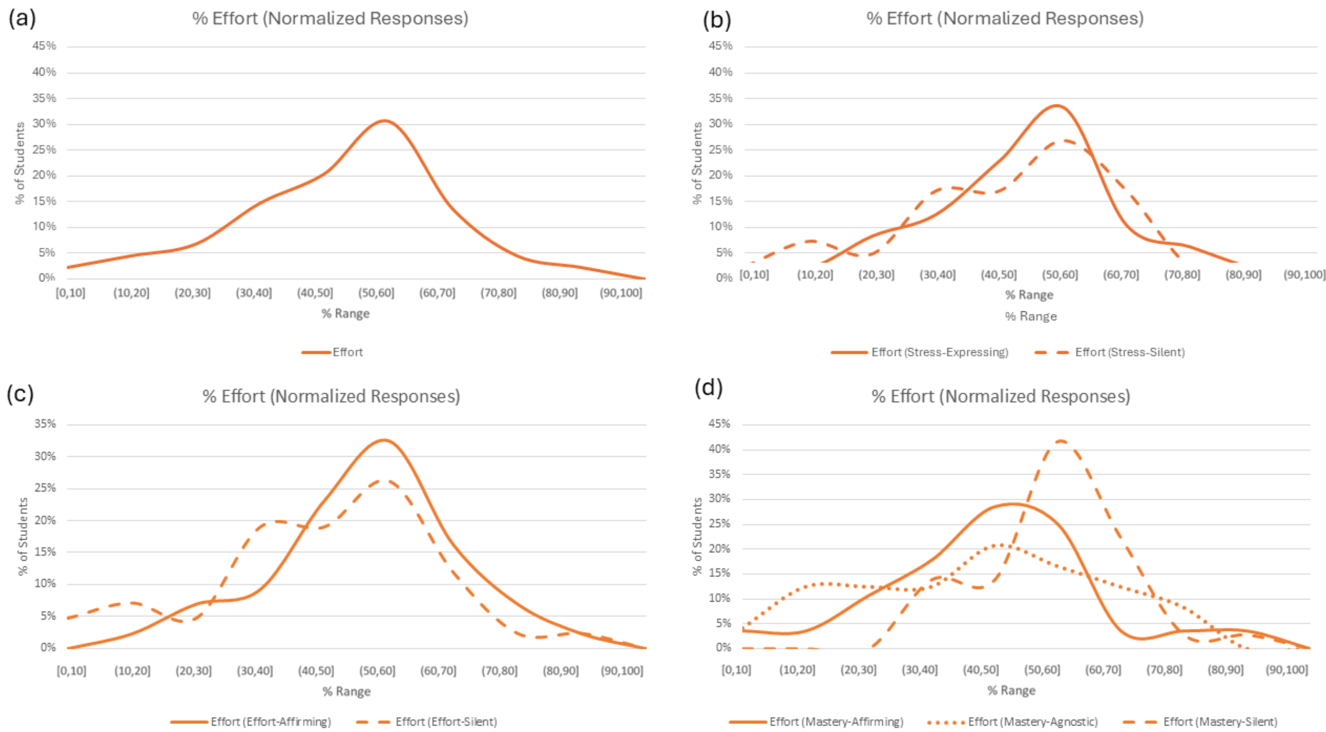

FIG. 1. Distributions of students' preferred weightings for effort when determining a physics grade. Responses are normalized, such that the remaining percentage is assigned to understanding. Responses are reported for (a) all $N = 88$ students and disaggregated by (b) whether students mentioned grade-related stress as a negative experience, (c) whether students positively mentioned grades as representing effort, and (d) whether students mentioned grades as representing mastery, either positively or in mixed terms.

their grade to be weighted.

Theme 2: Effort. In the free responses, we found that 47% of the students described grades as representing their effort, and framed this representation as a good thing. We disaggregated the grade weighting responses to investigate whether these Effort-Affirming students indicated different weighting preferences. Figure 1c shows the weightings from Effort-Affirming students (solid curve) and Effort-Silent students (dashed curve). The distributions are statistically indistinguishable, with a Kruskal-Wallis H-test revealing no statistically significant differences ($p = .08$). Therefore, it does not seem that a positive perception of grades representing effort explains how students prefer their grade to be weighted.

Theme 3: Mastery. In the free responses, we found that 30% of students described grades as representing their mastery of the subject matter, and framed this representation as a good thing (Mastery-Affirming students). We also found that 29% of students also described grades as representing mastery but framed it in mixed terms (Mastery-Agnostic students). We disaggregated the grade weighting responses by these categories. Figure 1d shows the weightings from Mastery-Affirming students (solid curve), Mastery-Agnostic students (dotted curve), and Mastery-Silent students (dashed curve). These distributions show a statistically significantly higher-effort peak for Mastery-Silent students than for Mastery-Affirming and Mastery-Agnostic students ($p < .025$ for each comparison). These differences have medium effect sizes of $d = 0.53$ between Mastery-Affirming and Mastery-Silent and $d = 0.64$ between Mastery-Agnostic and Mastery-Silent. It therefore seems that there is some association between a preference for weighting effort higher in a physics grade and whether a student perceives mastery as important enough to a grade to mention it explicitly in these free response questions.

Comparing Figures 1c and 1d reveals an interesting contrast. Effort-Affirming students did not weight effort significantly higher than Effort-Silent students. It therefore seems that these students prioritized effort roughly equally in their weightings regardless of whether they expressed effort as a component of grades in their free responses. This similarity could indicate that students value effort roughly equally whether they express this value or not. However, both Mastery-Affirming and Mastery-Agnostic students weighted effort significantly lower than Mastery-Silent students. This difference could mean that students' preferences for grade weighting are shaped by the degree to which they have internalized the nature of grades as representing content mastery.

IV. DISCUSSION AND CONCLUSIONS

It is perhaps unsurprising that these students overall showed a preference for higher weighting toward effort. With grades operating as a currency in classes, students likely develop a perspective of grades as a compensation for the time they invest. Additionally, physics is often perceived as a difficult subject, and students might feel that progress toward completion is more attainable than a certain threshold of understanding subject to the instructor's assessment.

It is interesting, though, that the peak for effort in Figure 1a is not far to the right of 50%, and only 7% of these students went so far as to weight effort greater than 70% (a ratio greater than 2:1 between effort and understanding). These students, it seems, do not solely value effort as a more attainable goal, but do see a place for grades to incorporate their understanding of physics content. We find this interpretation consistent with students' free responses.

It is also interesting how students' mention of grade-related stress has no significant impact in how they would prefer effort and understanding weighted. One might think that students who are more concerned with grade-related stress would prefer an effort-heavy weighting, to avoid the stress of "getting the correct answers," but this is not supported by the data.

Our observations suggest a few implications for instruction:

1. Students want grades to reflect both their effort spent on course tasks and their understanding of course material. As discussed in Section I, this tension sits at the heart of designing an equitable grading strategy, and grading strategies that emphasize solely one or the other could induce student resistance. Indeed, student buy-in to alternative grading strategies is an important factor in a successful implementation [10]. This is not to say that polar extremes like standards-based grading or labor-based contract grading (which address a significant philosophical concern) should not be implemented, but our findings do suggest that instructors should address these mixed preferences when introducing alternative grading strategies.
2. Students' preferences regarding the balance of effort and understanding seem independent of their reported experiences of grade-related stress. Since neither of these groups reported stress more often than the other, it would seem unlikely that student stress could be reduced by appealing more to one of these preferences in a grading strategy. While a well-designed grading strategy might reduce student stress, we are unlikely to eliminate stress from a reformed design alone. The way in which we communicate our grading strategies to students should emphasize our intent to promote their mental well-being while motivating them to learn.
3. Students' preferences of how effort and understanding are weighted might be malleable by expanding their considerations of how grades represent content mastery. Given how these students demonstrated mixed responses to grades as content mastery (as a good thing, in mixed terms, or not expressed at all), instructors might benefit from discussing their own perceptions of what mastery means and how that is actualized in the grading process. This malleability could be leveraged when orienting a new class of students to an alternative grading strategy.

Our approach carries a few limitations. First, administering the survey at the end of the semester means our sample does not include students who withdrew from their physics course earlier in the semester. Second, some students used the "something else" option in Question 6 to include items that could be attributed to effort or understanding (e.g., class attendance). We did not add the weighting for these other options back into effort or understanding so as not to presume on what the students meant. Finally, while we normalized the percent weightings that students reported, it is possible we missed some important meaning when a student's total exceeded 100%. We do not know, for example, whether an excesses was the result of a numerical mistake, or whether a student intended to communicate something about how they think grades should be computed (e.g., with extra credit built in).

To support research about alternative grading strategies in physics education, we surveyed students about their perceptions of grades in physics using free-response and numerical questions. Their free responses provided a basis of disaggregating the students based on their expressions of grade-related stress, grades as representative of effort within the course, and grades as representative of content mastery. We found that, overall, students preferred a higher weighting toward effort than understanding when determining a physics course grade. These preferences showed no significant differences based on free-response expressions of stress or effort. We found a statistically significant difference in weighting preferences with a medium effect size between students who mentioned grades as representing content mastery (whether in positive terms or mixed terms) and those who did not mention content mastery at all. This difference seems to indicate that students' preferences for grade weighting are shaped by their perspective of whether grades represent content mastery. We reflected on the implications of these findings for future studies of alternative grading strategies in physics and how physics instructors formulate and present alternative grading strategies to students.

ACKNOWLEDGMENTS

This work was supported by the University of North Florida Office of Undergraduate Research and Summer Undergraduate Research Fellowship.

[1] E. L. Hackerson, T. Slominski, N. Johnson, J. B. Buncher, S. Ismael, L. Singelmann, A. Leontyev, A. G. Knopps, A. McDarby, J. J. Nguyen, et al., Disciplinary and Interdisciplinary Science Education Research **6**, 15 (2024).
[2] M. A. Muñoz and T. R. Guskey, Phi Delta Kappan **96**, 64 (2015).
[3] D. J. Webb and W. H. Potter, arXiv preprint arXiv:2102.10451 (2021).
[4] I. D. Beatty, Journal of the Scholarship of Teaching and Learning pp. 1–22 (2013).
[5] A. B. Inoue, *Labor-based grading contracts: Building equity and inclusion in the compassionate writing classroom* (Wac Clearinghouse Fort Collins, CO, 2019).
[6] K. Kryger and G. X. Zimmerman, Journal of Writing Assessment **13** (2020).
[7] E. Ward, Assessing Writing **48**, 100508 (2021).
[8] J. O. Inman and R. A. Powell, College Composition & Communication **70**, 30 (2018).
[9] W. B. Lane, J. C. Speirs, and N. Laird (forthcoming).
[10] A. B. Carter, Ph.D. thesis, Walden University (2016).