

Ungrading in a third-year biology laboratory indicates positive student results without loss of outcomes

Celeste Leander,^{1,2} Tessa Blanchard,¹ Maia V. Palka,² Amy S. Kim,² Christine Goedhart^{2,3}

AUTHOR AFFILIATIONS See affiliation list on p. 10.

ABSTRACT Traditional grading practices in undergraduate science education have been criticized for emphasizing compliance and performance over authentic learning and scientific inquiry. Although ungrading has gained attention as an alternative assessment strategy, evidence from upper-level STEM laboratory courses remains limited. Here, we evaluated a fully ungraded assessment model in a third-year project-based biology laboratory course at a large Canadian research university. Grades and point-based assessments were replaced with formative feedback, student-generated learning goals, reflective practice, and end-of-term self-assessment. Qualitative responses from 93 students were analyzed using inductive thematic coding. Students most frequently reported decreased stress, increased enjoyment, greater focus on learning rather than grades, increased motivation, and a safer environment for experimentation and risk-taking. Comparisons with traditionally graded laboratory courses similarly emphasized reduced anxiety and greater engagement with learning. To assess academic outcomes, 20 final manuscripts from the ungraded cohort and 20 from previously graded cohorts were evaluated blindly using a 16-variable scientific writing rubric. No significant differences were detected in manuscript quality or final course grades between graded and ungraded cohorts. These findings suggest that ungrading can improve student experience and foster intrinsic motivation in upper-level biology laboratories without diminishing learning outcomes.

KEYWORDS STEM, higher education, ungrading, alternative grading, biology laboratory

Despite its ubiquity in formal education, traditional grading has long been criticized for its limitations in accurately reflecting student learning (1–4). Compounding the problem of inaccurate measurement is the mounting evidence that traditional grading systems decrease motivation and shift focus away from learning (2, 5). Since their relatively recent advent, conventional letter and percentage-based grades have evolved into a system that privileges correctness and compliance over collaboration, critical thinking, creativity, or risk-taking, all of which are skills that are central to scientific inquiry. These privileges of understanding correctness and compliance are nuanced, favoring language and cultural natives, neurotypical students, and those that perform well under standardized conditions. Furthermore, these grading practices favor the use and assessment of lower-order cognitive skills and do not adequately assess student understanding (3, 4, 6, 7). Additionally, traditional assessments that have a strong emphasis on grades may inadvertently diminish the value of experiential learning. This is particularly worrisome in fields where hands-on experience is critical to student development. As these criticisms become more obvious, there is increasing interest in alternative grading models that aim to better align equitable assessment with student learning.

Editor Stacy Ochoa, University of California San Diego, La Jolla, California, USA

Address correspondence to Celeste Leander, celeste.leander@botany.ubc.ca.

The authors declare no conflict of interest.

Received 15 May 2026

Accepted 20 August 2026

Published 11 September 2026

Copyright © 2026 Leander et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Several categories of approaches have emerged as viable alternatives to standard grading practices, including labor-based grading, standards-based (or specifications-based) grading, binning or bucketing, and ungrading. Although similar, these systems differ in their underlying philosophies and should be aligned with instructor or learner values (8). Many of these assessment strategies avoid the use of summative assessments and have been shown to promote self-efficacy, foster the use of higher-order cognitive skills, and help support persons typically excluded due to ethnicity, gender, race (PEERs), or neurodivergence. Similarly, these approaches are also consistent with Indigenous principles of education, which emphasize holistic and relational learning that supports personal growth, reflection, and reciprocal relationships rather than performance (9).

Labor-based grading

Labor-based grading is rooted in critical pedagogy and aims to shift the focus of assessment from summative outcomes to rewarding effort by measuring the time and effort students invest rather than the quality of the final product (10, 11). By removing the emphasis on performance or product, labor-based grading acknowledges that students bring diverse lived experiences and cultural perspectives to their work. In practice, labor-based grading can involve setting expectations for reading, writing, doing, or participation and rewarding students who meet those benchmarks, regardless of “perceived quality” as judged by the instructor. While this method can democratize the classroom and validate multiple forms of engagement, it may be less suited for contexts where mastery of specific content or technical accuracy is essential. Instructors who choose labor-based grading typically value effort as a measure of learning.

Standards-based and specifications-based grading

Standards-based grading (SBG) and specifications-based grading assess students based on their ability to demonstrate specific learning objectives (12, 13). These sometimes involve creating contracts with students beforehand, and in doing so, can be considered aligned with contract-based grading. These systems articulate clear, granular outcomes (e.g., “Can explain the role of ribosomes in protein synthesis”) and evaluate student work against predefined criteria. Rather than earning points toward an aggregate score, students receive binary or tiered evaluations (e.g., “met,” “not yet met,” or “exceeds”) and are often allowed to revise their work until proficiency is demonstrated (14). This approach supports transparency between instruction and assessment (12, 15, 16). However, it may inadvertently reduce assessment to checklists, overlooking growth, or reduce the opportunity for students to practice the process of inquiry inherent to authentic scientific work. A further consideration with this approach is that the time and energy for retakes are a luxury not all students can afford (17). Instructors who implement standards or specifications-based grading methods value mastery of specific outcomes.

Binning or bucketing

Binning aims at a holistic approach to measuring student understanding. Rather than the granular assignment of points or fractions of points based on a fine scale rubric, binning zooms out in an attempt to categorize understanding of a concept or content and lumps scores into pre-determined categories, which are scored numerically (e.g., 30%, 50%, 75%, or 100%). Binning works well for educators hoping to measure mastery of content or concepts in a more holistic yet nuanced way than is found in the binary system of standards-based grading.

Ungrading

Ungrading represents a more radical departure from traditional systems and encompasses many disparate approaches (3). Rather than replacing grades with another metric, ungrading removes them entirely during the learning process, allowing students to

focus on feedback, reflection, and self-assessment (18–21). Instructors typically provide detailed formative feedback throughout the term and may offer summative commentary, but numeric or letter grades are not assigned until required by institutional policy—typically through a final student self-assessment or portfolio. Ungrading is grounded in constructivist and humanistic theories of education, emphasizing trust, intrinsic motivation, and learner autonomy. Proponents argue that ungrading fosters deeper learning, metacognition, and reduces the stress and competition often associated with grades (3, 22). Critics, however, question whether ungrading can be applied rigorously in STEM contexts or whether students, particularly those from marginalized backgrounds, may struggle without the clarity and structure that grades provide (21). Instructors who adopt ungrading may value many things, including measuring student growth. In this study, we choose to further emancipate students from our control of grading by measuring student success based on how the student chooses to define it.

These alternative grading practices are not mutually exclusive and may be blended or adapted depending on course context, instructor values, and student needs. Indeed, a key challenge facing educators is not the lack of alternative assessment models, but the abundance of them, each with its own pedagogical assumptions and practical considerations. Instructors must decide not only what they want to measure (e.g., mastery, effort, growth) but also how those measurements align with their course and values.

Assessment and learning in the life sciences

In the life sciences, assessment often serves dual purposes: to evaluate students' understanding of foundational content (e.g., metabolic pathways, genetic principles) and to prepare them for future scientific practice, which is not fundamentally about content. In fact, undergraduate biology education is shifting towards a concept-oriented approach, following the creation of the *Vision and Change for Undergraduate Biology Education* report (23). Laboratory courses, in particular, demand that students engage in both conceptual learning and skill acquisition. Effective assessment in this context should capture not only what students know but also how they learn, the choices and decisions they make, and how they communicate scientific knowledge (24, 25). Yet, traditional grading in laboratory settings often rewards compliance with protocols and formatting rather than genuine inquiry or reflective thinking.

There is growing awareness of the need for assessment practices that promote equity and inclusion in life science education (26). In our course, we also strive to encourage risk-taking and resilience in scientific inquiry. Traditional grading systems can exacerbate disparities by penalizing students who lack prior exposure, resources, or confidence (3). These factors are often tied to systemic or inherent inequities, particularly in persons excluded due to ethnicity or race (PEERs). In contrast, alternative approaches that value growth, feedback, and student agency may help dismantle barriers to participation and success in the life sciences (3).

Our context and rationale

Informed by these broader pedagogical conversations, we sought to implement and evaluate a complete ungrading approach in a project-based third-year biology laboratory course at a large Canadian research university. By “complete ungrading,” we mean that we kept no scores, grades, or spreadsheets throughout the term. Our goals were twofold: first, to create a learning environment that emphasized curiosity and collaboration rather than performance and competition; and second, to explore whether ungrading could maintain academic standards in a scientific setting while improving student experience. Although artificial intelligence (AI) was not readily available when this study took place, we do recognize parallels in the goal to design meaningful student work where cheating provides no benefit. At the heart of our teaching philosophy is a commitment to student agency and indigenous principles of relational, holistic assessment (9, 27). These principles emphasize assessment as a relational and continuous process that supports the whole learner, values reciprocal feedback and reflection,

and recognizes numerous ways of demonstrating learning, rather than relying solely on numerical and individualistic measures of achievement (9). We understand that students, when given structured opportunities and support, are capable of assessing their own learning with depth and honesty (3).

In the fall of 2021, as students returned to campus after prolonged remote learning during the COVID-19 pandemic, we replaced traditional grading with a fully ungraded model. Although ungraded components of a course have become more common, fewer instructors and courses embrace ungrading as the full assessment model. Over the term, we gathered both qualitative and quantitative data to examine how students responded to this change, how the quality of their final projects compared with those from traditionally graded cohorts, and whether their self-assigned grades aligned with instructor-assigned grades from previous years.

This study contributes to the growing literature on alternative assessment by offering empirical data on the implementation of ungrading in an upper-level biology lab. It explores not only outcomes, but also student perceptions, and addresses the tension between innovation and accountability in postsecondary STEM education.

METHODS

Qualitative student feedback

Students took part in a fully ungraded laboratory course. This course consists of three large student projects, including a student-designed and implemented independent research portion. No changes were made to course content as we moved to ungrading. With our model, we asked students to monitor and justify their own marks at the end of the term. To further emancipate students from our control over their learning experience, we structured the course so that students had access to the procedural lab manual well in advance but without top-down learning objectives. Before each lab, students read the lab manual and set their own individual learning goals in their lab notebooks. In this way, students with more experience at a certain task or concept could focus on another aspect of the lab. Following each laboratory, students reflected back on their learning goals. At the end of the term, students used their individual learning goals in addition to reflections of feedback on their projects to help justify a course grade. Although we did provide substantial feedback on all projects and laboratory notebooks, no marks were allocated by the teaching team, and nothing was graded throughout the term. At the end of the term, 93 students (out of 96 enrolled) provided responses to the prompt:

"Please comment on what this 'ungrading' experience was like for you this term. (If you've taken a lab course that was graded traditionally, how was your experience different in this ungraded course?)"

These unedited responses were anonymized and independently coded by two researchers who were unfamiliar with the course using NVivo 14 (Lumivero, version 14.23.3) (28). Thematic codes were cataloged and analyzed to identify patterns in student perception (Table 1).

Coding of student responses

Student responses were independently reviewed by two individuals using an inductive thematic analysis approach to establish interrater reliability. All responses were coded independently, and emerging analytic codes were agreed upon by both individuals to ensure consistency. Any discrepancies between coders were reviewed, discussed to reach consensus, and assigned a final code, thereby strengthening the reliability of the analysis.

Quality assessment of final projects

We randomly selected 20 final manuscripts from ungraded terms and 20 from previous traditionally graded terms, taught by the same instructional team. Identifying

TABLE 1 Example unedited student comments on ungrading and NVivo coding for each comment^a

Unedited student comment	NVIVO coding
It was a strange experience but a good one! I was not used to not being graded traditionally. I found that it reduced a lot of my stress and anxiety and perhaps one of the reasons why we chose a final project that was a bit unconventional was because I felt like I could explore different options that may not work without being penalized for it. I really liked this course especially because it really encouraged us to explore and be brave with our final projects by letting us know that it's okay if it doesn't work! I really appreciated this because in my other lab courses, I definitely felt anxious about doing something that was unconventional and uncertain.	Reduced stress and increased enjoyment Safe learning environment Alternative grading is unfamiliar
Originally, I thought that ungrading would make me be lazy in this course. I thought that I wouldn't feel motivated without a grading system. It was the opposite. I was fueled by my enjoyment and without grades, I never felt discouraged. In the previous labs I have taken, getting good grades was all I cared about, and I never really absorbed any of the material. When I got bad grades, it made me dislike and anxious about going into the lab and doing any work related to the course. In this lab, I feel like I learned way more than I normally do in labs.	Comparison with other courses Reduce stress/increase enjoyment Focus on learning over grades More effort or motivation for learning Alternative grading is unfamiliar
I'm very grateful that this class is ungraded! I think knowing that I was responsible for my own mark kept me motivated and furthermore, made sure that I was held accountable for how hard I worked on our experiments and projects. It felt nice knowing that when I came into class, my instructors weren't there to mark what I was doing (like in other labs), but instead encourage learning and participation. In comparison to other lab courses, I felt much more relaxed when it came to this class and I think the upgraded system is something that should continue to be considered!	More effort or motivation for learning Focus on learning over grades Reduce stress/increase enjoyment Comparison with other courses
I appreciated being in a learning environment that encouraged exploration and did not penalize exploring things we were not already perfect at. I felt truly heard and valued as a student. At first, I was concerned that the lack of grades would also mean a lack of feedback on assignments, but the teaching team definitely proved me wrong.	Feedback for learning Safe learning environment
This "ungrading" experience was awesome! It really kept me motivated to try new things and add risk factors to my learning. I think this positively affected my learning and I will be able to remember everything I learned from this course because I did not bear the stress of doing an assignment for a letter grade. I also overall had a lot of fun and felt very relaxed during class. I've never looked forward to a lab course until I took this class.	More effort or motivation for learning Focus on learning over grade Reduce stress/increase enjoyment

^aSuccinct examples were chosen to demonstrate depth and breadth of coding.

information from all manuscripts was removed. An independent educator unfamiliar with the course evaluated the manuscripts using a 16-variable rubric adapted from Vishwanathan et al. (29), with each category scored from 0–2, with 0 indicating a variable was absent or very poor, 1 indicating presence and marginally adequate inclusion of the variable, and 2 indicating presence and adequate inclusion of the variable (Table 2). Average cumulative scores between graded and ungraded manuscripts were compared using an unpaired *t*-test.

Grade comparison

At the end of the term, students are asked to submit a substantial self-assessment modified from Susan Blum (18). Final course marks assigned by students in the ungraded Fall 2021 term were compared with instructor-assigned grades from Fall 2018, the last traditionally graded, pre-pandemic offering by the same teaching team. A Welch's *t*-test was used to assess statistical differences between group means.

RESULTS

Student feedback about the effects of ungrading

Thematic coding of 93 open-ended responses yielded 23 unique themes and 379 coded instances (Fig. 1). The most frequent theme that students report was decreased stress and/or increased enjoyment (78 mentions), followed by a shift from an emphasis on learning over grades (67), and student comparisons with traditional labs that were

TABLE 2 Variables used to score student manuscripts^a

Abstract	Introduction	Methods	Results	Discussion	Overall
Includes objectives, methods, results, and conclusion succinctly	Aims/objectives clearly articulated	Study design (will it answer the questions?)	Description complete	Relates back to intro	Structure of paper (clear and appropriate)
	Acknowledges appropriate related work by others	Describes what (not how)	Figures complete (title, legend, error bars)	Explains results	Language clear and correct
	Hypothesis or prediction (or goal if observational)	Adequate and correct statistics or measurements	Presentation (figures look good, succinct)		References relevant
					No unnecessary duplication of info

^aEach variable was scored 0–2 based on the presence and quality of the work.

traditionally graded (60). Additional recurring themes included increased motivation (35) and the perception of a safer learning environment (32). Less frequent responses (<20 mentions) included praise for flexibility, appreciation for autonomy, and real-world applicability of self-assessment. A small number of students (4) reported that the lack of grades reduced motivation, while an equal number argued traditional grading hindered their learning (not shown in figure).

Student feedback when asked to compare the ungraded lab course to other traditionally graded lab courses

Thematic coding of 56 open-ended responses yielded 13 unique themes and 188 coded instances (Fig. 2). The most frequent theme that students reported was decreased stress, anxiety, or pressure (43 mentions), followed by an emphasis on learning over grades (31), and having a more enjoyable experience overall (25). Additional recurring themes

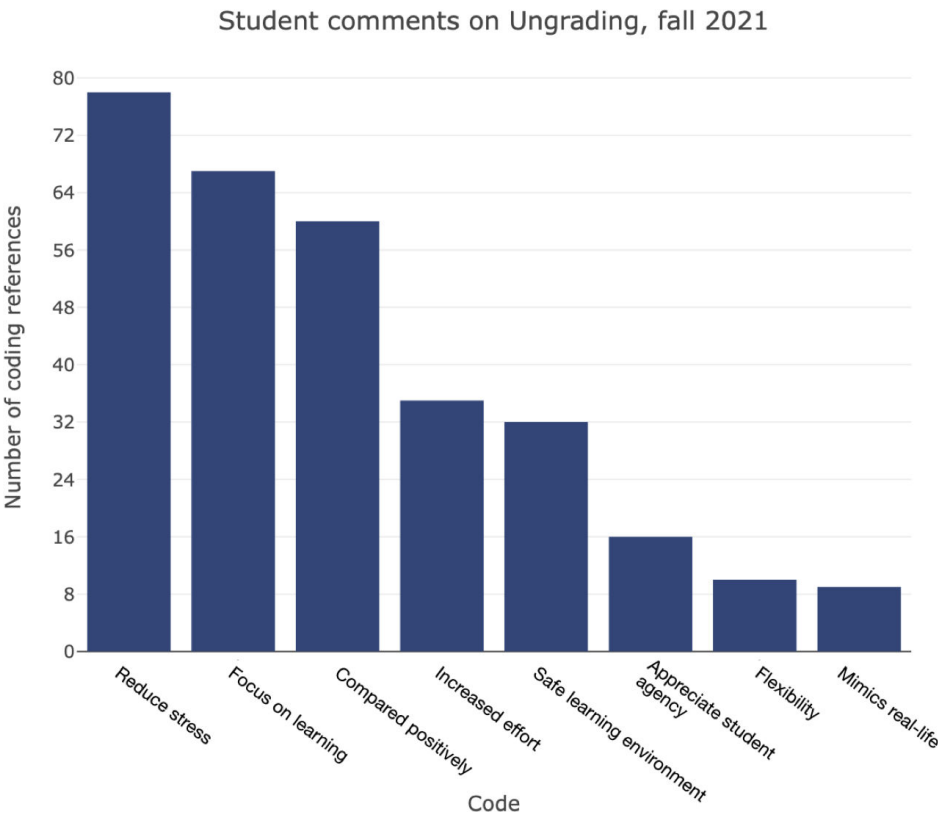


FIG 1 Most frequent thematic codes identified in open-ended responses to the effects of ungrading.

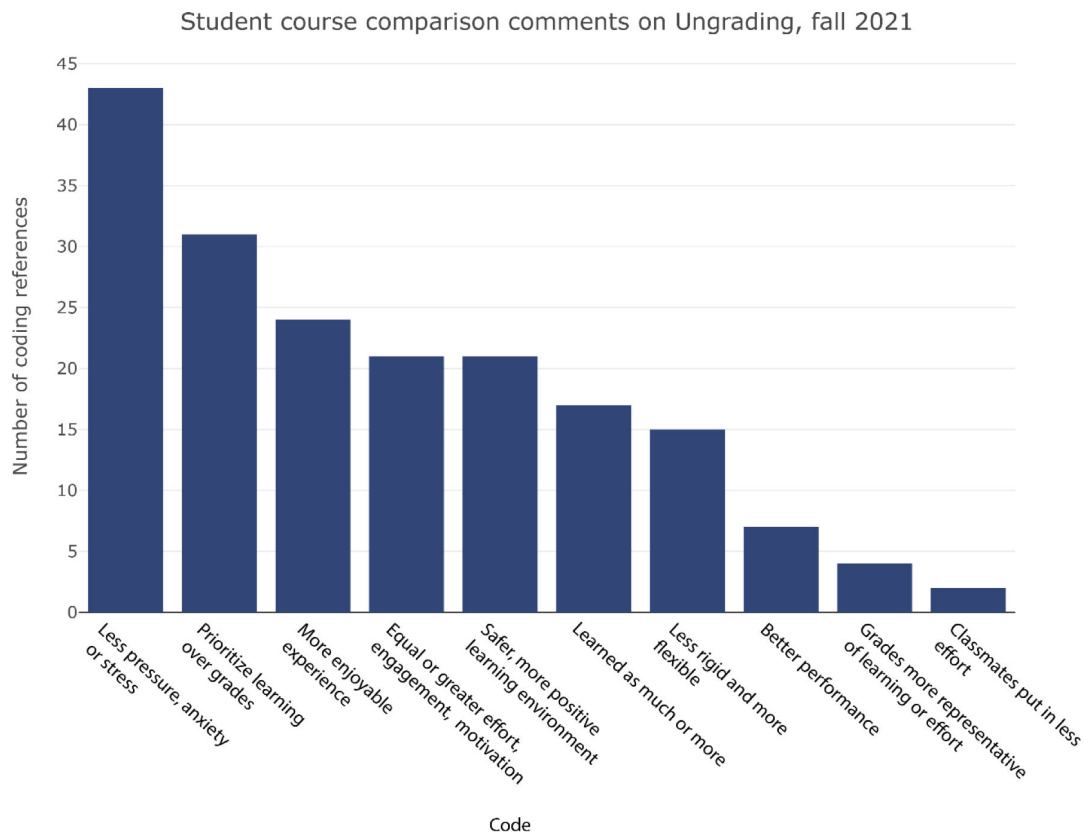


FIG 2 Most frequent thematic codes identified in open-ended responses to the comparison of this ungraded lab course with other lab courses.

included increased motivation (22) and the perception of a safer learning environment (22). Less frequent responses (<29 mentions) included praise for flexibility, learning as much or more as in a traditionally graded lab, and perception that ungrading is more representative of students' actual learning/effort. A small number of students (2) reported that they felt as though classmates put in less effort compared with a traditionally graded course. Three codes were omitted from Fig. 2, as only one response was recorded for each. Sample unedited student responses and correlated coding can be found in Table 1.

Quality of student work

There was no significant difference in the overall quality of final manuscripts between the graded and ungraded groups ($P = 0.35$), suggesting comparable academic rigor (Fig. 3).

Final course grades

The comparison of final grades from the ungraded 2021 term ($N = 94$) and the graded 2018 term ($N = 97$) revealed no statistically significant difference (Welch's t -test: $P = 0.851$) (Fig. 3).

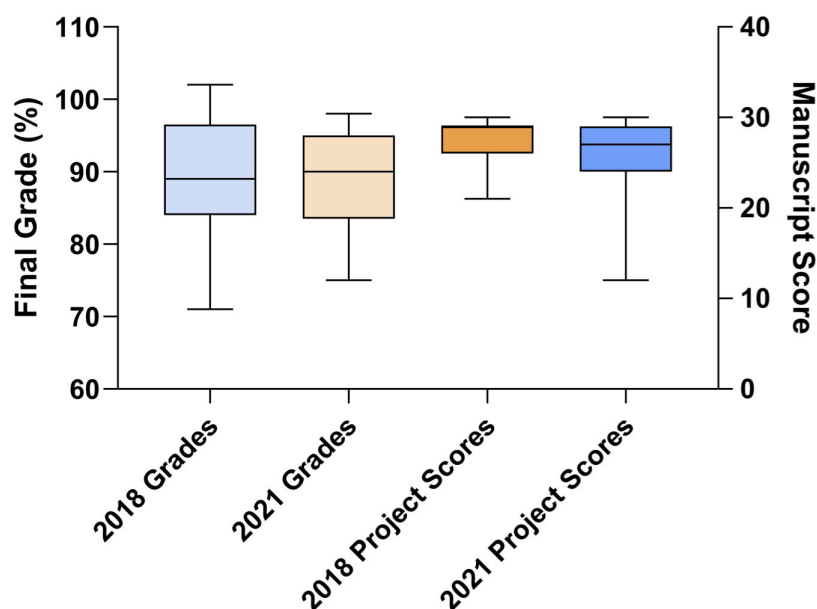


FIG 3 Comparison of final course grades and quality of final projects in 2021 (ungraded) and 2018 (traditionally graded). No significant difference was observed in either measurement.

DISCUSSION

This study suggests that ungrading is a feasible assessment strategy in a third-year biology laboratory course that can enhance the student learning experience without sacrificing academic rigor or grade accuracy. Ungrading invites students to engage more deeply and authentically with course material. By doing so, students demonstrated that when given the opportunity and support to take ownership of their learning, they rose to the occasion. We suggest that ungrading is not a permissive or lax alternative, but is instead, an approach that encourages students to reflect on and honor their learning trajectory with pride of ownership.

A noteworthy outcome was the consistency in the quality of final projects when comparing traditionally graded cohorts with those assessed under the ungrading model. Initial concerns that removing formal grades would lead to diminished effort or inflated self-assessments were not borne out in our data. In fact, we observed the opposite. Most students described feeling more motivated, not less, when they were freed from the pressure of numerical evaluation. Instead of working to please the instructor or optimize their GPA, students were able to center their own learning goals and curiosities. This shift—from extrinsic to intrinsic motivation—is consistent with self-determination theory and other well-established frameworks in educational psychology (30). The evidence here points to a reorientation of purpose: students were not working less; they were working differently and often more meaningfully. As an instructional team, our relationship with students became one of true allies in learning rather than as adversarial authority figures.

Student feedback underscored this transformation. As expected, most students described reduced stress levels and increased enjoyment of their lab work. The ungrading environment created space for experimentation, failure, and revision—hallmarks of real scientific inquiry that are often overshadowed by high-stakes grading systems. We saw students make more brave choices around term projects—trying things that were risky and more likely to fail. Prior to ungrading, we would occasionally see students design “brave” projects (i.e., those where the teaching team had little or no expertise to assist, or were likely to fail). Across four sections of lab, we would see one or two such projects. With the adoption of ungrading, we see students propose these “brave” projects more frequently, typically four or five per term. Interestingly, students

sometimes recognize this choice and comment on this in course reflections (see Table 1, comment 1). Rather than being penalized for not getting something right, students felt empowered to take risks, knowing they would receive feedback and the chance to possibly iterate. This process mirrors the actual practice of science more closely than a one-shot, high-stakes lab report.

Equally important, the ungrading framework aligned closely with Indigenous principles of education, which emphasize holistic evaluation (27). In centering student voice and narrative, ungrading helps create a learning environment that values multiple ways of knowing and being. This approach supports a broader range of learners—including those who may not excel in traditional, high-pressure academic structures but who thrive when given autonomy and encouragement to reflect on their own progress. The ethical and cultural implications of this alignment are important. By incorporating relational assessment practices, ungrading moves beyond mere pedagogical innovation and toward a decolonizing approach to science education.

Still, ungrading is not a panacea. We emphasize that a single grading system is not ideal or appropriate for all courses. Ungrading poses real challenges, particularly for students who are accustomed to using grades as a benchmark for success. Although few, some participants expressed uncertainty or discomfort in the absence of clear numerical indicators. Others admitted that without the external pressure of a grade, their motivation initially waned. We do not discard these students and suggest refinement of our practice to include more robust support systems that help students track their own progress regularly rather than one self-assessment at the end of the term.

Instructors who wish to employ ungrading must be prepared to shift their roles from judges of quality to facilitators of growth. This demands time and a willingness to tolerate ambiguity. It may also require institutional buy-in, particularly in large or multi-section courses where consistency and comparability are key concerns.

Looking forward, there is a clear need for further research. What kinds of students thrive in ungraded environments and which ones struggle? How do variables such as prior academic success, cultural background, or neurodivergence influence student experiences of autonomy and feedback? Could hybrid models, such as courses that allow students to choose between graded and ungraded pathways, offer a middle ground that balances structure with agency? And how might ungrading intersect with broader efforts to improve equity and inclusion in STEM education?

Ultimately, this study contributes to a growing body of evidence that suggests assessment can be reimaged to better serve learners without compromising educational standards. Ungrading invites us to trust our students, to value process over product, and to see learning not as a race to a grade, but as a shared, evolving journey. If done thoughtfully, it holds the potential to transform not only how we teach, but how students come to see themselves as scientists—and as capable, self-directed learners.

Limitations and situational awareness

The participants of this study are upper-year Combined Major in Science (CMS) and biology students who successfully completed their first- and second-year degree requirements. As a result, this study consists of a self-selected population likely well-adjusted to traditional academic structures and may have been performing well in the absence of ungrading. Consequently, the findings may not be generalizable to students earlier in their academic programs or those who are still adapting to university-level coursework.

The cohort in 2021 contained 96 students in four lab sections consisting of 24 students, and the cohort in 2018 contained 97 students in four lab sections consisting of 25 students or less. Small class sizes have been shown to enhance academic performance by providing more personalized feedback, greater opportunities to participate, and increased individual attention (31). Ungrading is also easier to implement successfully in smaller class sizes. Therefore, the positive outcomes observed in this study may have

been influenced not only by ungrading but also by the benefits associated with smaller class sizes.

Another limitation to this study is that the quality assessment of final projects was conducted by only one evaluator, although she was not familiar with the course. While this likely reduced potential bias associated with past knowledge of the course, relying on a single assessor limits the ability to evaluate inter-rater reliability and may have introduced subjective bias into the assessment process.

ACKNOWLEDGMENTS

This research and the writing of this article were conducted on the traditional, ancestral, and unceded territory of the xʷməθkʷəy̓əm (Musqueam) people.

This work was funded by the Departments of Zoology and the Department of Botany at the University of British Columbia and UBC ISO TL Linkage Grant, “Alternative Grading in Higher Education,” 2023.

AUTHOR AFFILIATIONS

¹Department of Zoology, Beaty Biodiversity Research Centre and Biological Sciences Building, University of British Columbia, Vancouver, British Columbia, Canada

²Department of Botany, Beaty Biodiversity Research Centre and Biological Sciences Building, University of British Columbia, Vancouver, British Columbia, Canada

³Skylight: Science Centre for Learning and Teaching, Faculty of Science, University of British Columbia, Vancouver, British Columbia, Canada

AUTHOR ORCID*s*

Celeste Leander  <http://orcid.org/0009-0006-8946-606X>

Maia V. Palka  <http://orcid.org/0009-0005-8743-4438>

Amy S. Kim  <http://orcid.org/0000-0002-2623-4118>

Christine Goedhart  <http://orcid.org/0009-0006-5490-3266>

AUTHOR CONTRIBUTIONS

Celeste Leander, Conceptualization, Data curation, Funding acquisition, Methodology, Project administration, Visualization, Writing – original draft | Tessa Blanchard, Data curation, Methodology, Software, Writing – review and editing | Maia V. Palka, Visualization, Writing – review and editing | Amy S. Kim, Visualization, Writing – review and editing | Christine Goedhart, Conceptualization, Formal analysis, Software, Visualization, Writing – review and editing

REFERENCES

1. Feldman J. 2019. Beyond standards-based grading: why equity must be part of grading reform. *Phi Delta Kappan* 100:52–55. <https://doi.org/10.1177/0031721719846890>
2. Schinske J, Tanner K. 2014. Teaching more by grading less (or differently). *CBE Life Sci Educ* 13:159–166. <https://doi.org/10.1187/cbe.14-03-0054>
3. Sorensen-Unruh C. 2024. The ungrading learning theory we have is not the ungrading learning theory we need. *CBE Life Sci Educ* 23:es6. <https://doi.org/10.1187/cbe.24-01-0031>
4. Uminski C, Burbach SM, Couch BA. 2024. Undergraduate biology lecture courses predominantly test facts about science rather than scientific practices. *CBE Life Sci Educ* 23:ar19. <https://doi.org/10.1187/cbe.23-12-0244>
5. Koenka AC, Linnenbrink-Garcia L, Moshontz H, Atkinson KM, Sanchez CE, Cooper H. 2021. A meta-analysis on the impact of grades and comments on academic motivation and achievement: a case for written feedback. Edited by L. Yang, M. M. Chiu, and Z. Yan. *Educ Psychol (Lond)* 41:922–947. <https://doi.org/10.1080/01443410.2019.1659939>
6. Boatright-Horowitz SL, Arruda C. 2013. College students' categorical perceptions of grades: it's simply “good” vs. “bad”. *Assess. Eval. High. Educ* 38:253–259. <https://doi.org/10.1080/02602938.2011.618877>
7. Brookhart SM, Guskey TR, Bowers AJ, McMillan JH, Smith JK, Smith LF, Stevens MT, Welsh ME. 2016. A century of grading research: meaning and value in the most common educational measure. *Rev Educ Res* 86:803–848. <https://doi.org/10.3102/0034654316672069>
8. Lewis D. 2024. Alternative grading handout. QUBES Educational Resources. Available from: <https://doi.org/10.25334/92PB-EZ25>
9. Letendre A, Bennetch RJ. 2022. A new-old way of doing assessment: indigenous ways of knowing and Ungrading in a pandemic. Poster at the Society for Teaching and Learning in Higher Education.
10. Inoue AB. 2017. Writing assessment as the conditions for translingual approaches: an argument for fairer assessments, p 119–134. *In* Horner B, Tetreault E (ed), *Crossing divides: exploring translingual writing pedagogies and programs*. Utah State University Press, Utah, USA.
11. Sims ML. 2023. Shifting perceptions of socially just writing assessment: labor-based contract grading and multilingual writing instruction. *Assessing Writing* 57:100731. <https://doi.org/10.1016/j.asw.2023.100731>

12. Howitz WJ, McKnelly KJ, Link RD. 2021. Developing and implementing a specifications grading system in an organic chemistry laboratory course. *J Chem Educ* 98:385–394. <https://doi.org/10.1021/acs.jchemed.0c00450>
13. Fernandez TM, MartinkM, Mangum RT, Bell-Huff CL. 2020. Whose grade is it anyway?: transitioning engineering courses to an evidence-based specifications grading system. 2020 ASEE Virtual Annual Conference Content Access Proceedings, Virtual on Line. , p 1–13 ASEE Conferences
14. Nilson LB. 2014. Specifications grading: restoring rigor, motivating students, and saving faculty time. 1st ed. Routledge, New York, USA.
15. Katzman SD, Hurst-Kennedy J, Barrera A, Talley J, Javazon E, Diaz M, Anzovino ME. 2021. The effect of specifications grading on students' learning and attitudes in an undergraduate-level cell biology course. *J Microbiol Biol Educ* 22:1–8. <https://doi.org/10.1128/jmbe.00200-21>
16. McKnelly KJ, Howitz WJ, Thane TA, Link RD. 2023. Specifications grading at scale: improved letter grades and grading-related interactions in a course with over 1,000 students. *J Chem Educ* 100:3179–3193. <https://doi.org/10.1021/acs.jchemed.2c00740>
17. Supriya K, Bang C, Ebie J, Pagliarulo C, Tucker D, Villegas K, Wright C, Brownell S. 2024. Optional exam retakes reduce anxiety but may exacerbate score disparities between students with different social identities. *CBE Life Sci Educ* 23:ar30. <https://doi.org/10.1187/cbe.21-11-0320>
18. Blum S. 2020. Ungrading: why rating students undermines learning (and what to do instead). West Virginia University Press, Morgantown, WV.
19. Sackstein S. 2020. Shifting the grading mindset, p 74–81. *In* Blum S (ed), *Ungrading: why rating students undermines learning (and what to do instead)*. West Virginia University Press, Morgantown, WV.
20. Stommel J. 2020. Ungrading: an FAQ. Jesse Stommel. Ungrading: an FAQ (jessestommel.com)
21. Stommel J. 2021. Ungrading: an introduction. Jesse Stommel. Ungrading: an introduction (jessestommel.com)
22. Osterhage JL, Usher EL, Douin TA, Bailey WM. 2019. Opportunities for self-evaluation increase student calibration in an introductory biology course. *CBE Life Sci Educ* 18:ar16. <https://doi.org/10.1187/cbe.18-10-0202>
23. American Association for the Advancement of Science. 2011. Vision and change in undergraduate biology education: a call to action: a summary of recommendations made at a national conference organized by the American Association for the Advancement of Science. July 15–17, 2009. Washington, DC
24. Indorf JL, Weremijewicz J, Janos DP, Gaines MS. 2019. Adding authenticity to inquiry in a first-year, research-based, biology laboratory course. *LSE* 18:ar38. <https://doi.org/10.1187/cbe.18-07-0126>
25. Beck C, Butler A, da Silva KB. 2014. Promoting inquiry-based teaching in laboratory courses: are we meeting the grade? *CBE Life Sci Educ* 13:444–452. <https://doi.org/10.1187/cbe.13-12-0245>
26. Theobald EJ, Hill MJ, Tran E, Agrawal S, Arroyo EN, Behling S, Chambwe N, Cintrón DL, Cooper JD, Dunster G, et al. 2020. Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proc Natl Acad Sci USA* 117:6476–6483. <https://doi.org/10.1073/pnas.1916903117>
27. Cajete GA. 2016. Indigenous education and the development of indigenous community leaders. *Leadership* 12:364–376. <https://doi.org/10.1177/1742715015610412>
28. Lumivero. 2025. NVivo (version 14.23.3) [NVivo 14]. Available from: <https://lumivero.com/products/nvivo/>
29. Vishwanathan K, Kambhampati SBS, Maini L. 2022. Reviewer's checklists for evaluating scientific manuscripts. *JOIO* 56:175–182. <https://doi.org/10.1007/s43465-022-00602-x>
30. Deci EL, Ryan RM. 1985. Intrinsic motivation and self-determination in human behaviour. Springer Science & Business Media, Boston, MA.
31. Dee TS, West MR. 2011. The non-cognitive returns to class size. *Educ Eval Policy Anal* 33:23–46. <https://doi.org/10.3102/0162373710392370>