



Implementing Ungrading in Undergraduate Exercise Physiology: Student Perceptions of Learning

ABSTRACT

Ungrading or alternative grading has gained popularity among undergraduate educators. The purpose of this study was to compare students' perceptions of learning and two methods of grading in an upper-level exercise physiology course with third- and fourth-year undergraduate students. Two ungraded (UG1 and UG2) and two traditionally graded (TG1 and TG2) sections were compared. A weighted grading scheme was used in the TG sections. In the UG sections students received a final letter grade which was co-decided between the student and instructor. To evaluate the grading methods between the four sections, the researcher used deidentified student self-reflections and anonymous student opinion surveys (TG1: $n=18$, TG2: $n=13$, UG1: $n=19$, UG2: $n=12$). Significant differences between sections centered on perceived learning ($F_{3,45}=4.18$; $p=0.01$; $\eta^2=0.22$). Perceived learning was significantly lower in TG2 than UG1 ($p=0.007$), UG2 ($p=0.005$), and TG1 ($p=0.008$). Significant differences also existed between courses for rating of the evaluation and grading techniques ($F_{3,45}=10.66$; $p<0.001$; $\eta^2=0.42$). Evaluation and grading techniques rating was significantly lower in TG2 than UG1 ($p<0.001$), UG2 ($p<0.001$), and TG1 ($p=0.006$). 87.1% ($n=27/31$) of the students indicated they agreed or strongly agreed that their learning was enhanced by UG. Qualitative analysis of open-ended reflection questions revealed students felt they understood content better with less grade-based anxiety. Students in the ungrading sections more regularly reflected on their learning and appreciated feedback and opportunities to revise mistakes. Overall, ungrading was effectively implemented in an undergraduate exercise physiology course.

KEYWORDS

ungrading, exercise physiology, alternative grading, undergraduate

INTRODUCTION

Faculty in higher education are increasingly experimenting with alternate and more inclusive modes of assessing student learning. Blum (2020) defines ungrading as a “practice which eliminates or greatly minimizes the use of assigned points or letter grades in a course, focusing instead on providing frequent and detailed feedback to students on their work, in relation to the course learning goals”; that definition is grounded in the work and philosophy of earlier scholars (Butler and Nisan 1986; Kohn 1999; Milton, Pollio, and Eison 1986). Many forms of ungrading use a student-centered approach focused on student self-reflection, metacognition, and decreasing the effects of potential instructor bias (Blum 2020; Stommel 2023). Blum (2020) and Newell-Caito (2024) describe a variety of forms of ungrading, and underlying all approaches is the notion that interval letter grade systems, the “traditional” A–F grading systems, as they are currently practiced, do not serve the intended purpose and can, in the worst cases, be harmful to students (Blum 2020; Chamberlin, Yasué, and Chiang 2023; Malouff, Emmerton, and Schutte 2013; Stommel 2023).

Ungrading may refer to both summative final grades and/or formative assessment within a course. Minimal grading used in many graduate and professional programs such as medicine, nursing, pharmacy, and veterinary studies refers to assignment of final grades as incomplete/complete or pass/fail (Frank and Sutherland-Smith 2021; Melrose 2017; Soric, Robinson, and Ulbrich 2019; West et al. 2020). Since most post-secondary institutions require final grades, ungrading approaches often relate to formative feedback on assignments and activities within the class as opposed to minimal gradation used for final course grades. Some common examples of ungrading include: minimal grading in which instructors use fewer gradations (e.g. complete/incomplete or strong/satisfactory/weak), grade-free zones with defined periods or assignments which are not graded, contract grading which outlines what students should complete to earn specific grades, specification grading in which students demonstrate understanding of specific concepts, and reflective self-assessment grading methods, such as process letters, which ask students to reflect on their learning and completed work (Blum 2020; Newell-Caito 2024; Stommel 2023). Use of student designed assessments or collaboratively developed rubrics are other forms of student-centered assessment included in ungrading (Blum 2020; Newell-Caito 2024; Stommel 2023).

Overall, the use of traditional grading systems, such as the A–F system, inherently do not promote student learning or center the student in the learning process (Gorichanaz 2022). Ungrading approaches emphasize a more complete understanding of student learning by using a combination of pedagogical practices implemented out of a desire to refocus on learning rather than grades. Deemphasizing grades reduces grade anxiety and allows students to focus on deeper understanding of course content (Chamberlin, Yasué, and Chiang 2023; Marcus and Tomasi 2020; Pulfrey, Buchs, and Butera 2011). Additionally, ungrading ignites curiosity, and makes learning rewarding, genuine, and effective (Blum 2020).

Grades have historically been understood as communicating feedback, sorting and ranking students, and motivating or providing accountability for students (Blum 2020; Kohn 2011; Schinske and Tanner 2014). Communication of progress and providing feedback are important for learning, but a single grade on an assignment does not provide enough feedback for students to understand where they need improvement. Additionally, feedback is often overlooked when given alongside a number or letter grade (Kohn 2011). Sorting and ranking provides a powerful avenue for instructor bias and often disproportionately disadvantages minoritized students (Malouff, Emmerton, and Schutte 2013). When used to sort or rank students, grades may also create a competitive relationship among students (Farias, Farias, and Fairfield 2010; Hayek et al. 2017), undermining possibilities for beneficial collaborative learning and encouraging a transactional relationship between teacher and student.

Especially in high achieving students, grades may also promote a fear of failure and reduce intellectual risk taking (Butler 1988; Pulfrey, Buchs, and Butera 2011). Pressure to earn and maintain high grades is a leading cause of stress and anxiety, among other factors which influence student mental health (Chamberlin, Yasué, and Chiang 2023). Spring and colleagues (2011) found that alternative student evaluation techniques can improve student health and wellbeing, which is positively related to academic performance (Bas 2021).

Grades may provide accountability or motivation for some students; however, the motivation is likely extrinsic, meaning task completion and satisfaction comes from extrinsic rewards or punishments (Deci and Ryan 2002). Extrinsic motivation may create a preference for the easiest possible task, which discourages the opportunity for deep understanding that comes from intrinsic motivation, or actively seeking meaningful learning opportunities beyond the mandatory minimum (Deci and Ryan 2002; Harackiewicz, Abrahams, and Wageman 1987; Kohn 2011). Additionally, traditional grading schemes tend to reduce student interest in learning, especially if the student

receives a low grade (Butler 1988; Harackiewicz, Abrahams, and Wageman 1987; Kohn 2011). While extrinsic motivation may be appropriate for shallow processing tasks such as memorization, students completing tasks which require deep processing, such as critical thinking, application, and synthesis of knowledge, will perform better if they are intrinsically motivated to understand and apply concepts (Deci and Ryan 2002; Walker, Greene, and Mansell 2005). Thus, students studying in causal reasoning-based disciplines, such as physiology and other STEM fields, philosophy, law, economics, and psychology, would benefit from enhanced intrinsic motivation.

Similar to other causal-reasoning based disciplines, exercise physiology requires an understanding of causal relationships and an understanding complete enough to allow for application in different situations (Michael 2007). Students often try to memorize course content which is then quickly forgotten after a learning assessment. Student perceptions that they should memorize facts is problematic in causal reasoning-based disciplines because the application of reasoning skills is more advanced than simple memorization (Michael 2007). Ungrading can potentially be useful in these disciplines by encouraging and rewarding intrinsic motivation and attention to understanding instead of memorization (Butler 1988). The added benefit of reducing grade anxiety also allows students in any discipline to focus on understanding and applying course content instead of on missed points. If students are focused on understanding and application, it is possible they will perform better (Chamberlin, Yasué, and Chiang 2023).

While many instructors have shifted to forms of ungrading, it is not yet common, perhaps due to perceived challenges with extensive time commitment, large class sizes, limited time for content delivery, and institutional mandates for assigning final grades (Patton and Johnson 2021; Schinske and Tanner 2014). Despite these perceived obstacles, ungrading is possible in causal reasoning-based disciplines. In two different semesters, ungrading was implemented in an upper division undergraduate exercise physiology course. More specifically, the instructor used a contract and minimal gradations-based evaluation system which also incorporated self-reflection. The rationale for transitioning to ungrading was to reduce grade anxiety and to give students more input into aligning their effort and perceived learning with their grade in the course. Before implementing ungrading, the instructor extensively researched different versions of ungrading and spoke with other faculty who were experienced with different ungrading styles. In previous sections of the course, traditional weighted grading was implemented.

The purpose of this study was to compare students' perceptions of learning, students' perceptions of the grading and evaluation technique used, and final grades between ungraded and graded sections of an upper division undergraduate exercise physiology course. Deviating from traditional grading schemes raises questions about grade inflation, the phenomenon of awarding higher grades over time for the same quality of work. Faculty starting with ungrading may face suspicions about grade inflation, thus, a comparison of final grades was included in this study. The second purpose of this study was to describe one instructor's process implementing ungrading. The researcher hypothesized students would report greater satisfaction with the ungrading method and greater perception of learning compared to traditionally graded sections. Additionally, the researcher hypothesized students would earn higher final grades in the ungraded sections of the course compared to the traditionally graded sections.

METHODS

Participants

The researcher collected data from students enrolled in an undergraduate exercise physiology course at a small liberal arts college in North America. Students were upper division pre-healthcare

(pre-physical therapy, pre-medicine, pre-occupational therapy, pre-physician assistant, and pre-athletic training) and exercise science students. Human anatomy and physiology, and an exercise science lab skills course were prerequisites for the course. Exercise physiology was not a required course for any major or minor course of study but is a prerequisite for some graduate programs and is instrumental for several exercise-based certification exams. The Institutional Review Board of the College of Saint Benedict/Saint John's University determined that this study was exempt.

The course took place once per academic year with relatively small class sizes. The same instructor taught all sections of the course referenced in this study and all sections of the course for the 10 years prior to this study. In traditionally graded section 1 (TG1) took place in spring 2018 with 18 students enrolled in the course. Traditionally graded section 2 (TG2) took place in spring 2019 with 13 students enrolled. In ungraded section 1 (UG1) offered in spring 2021, 12 students were enrolled, and 19 were enrolled in ungraded section 2 (UG2) during fall 2021. All students were enrolled in the course sections during the preceding semester and all students enrolled completed the course.

Procedures

The researcher used anonymous student opinion surveys required and collected by the institution for all course sections. Data were also collected from de-identified student self-reflections for the ungraded sections. Data from the institutionally required student opinion surveys (TG1: $n=10/18$; TG2: $n=13/13$; UG1: $n=15/19$; UG2: $n=11/12$) helped highlight student perceptions of the grading technique and perceived learning. The researcher also analyzed data from the final self-reflection survey in the ungraded sections (UG1: $n=19$; UG2: $n=12$) for student perceptions of ungrading specifically and their perceived learning. Questions analyzed from each survey are included in Table 1. Additional Likert scale and open-ended questions about how ungrading affected learning were added to the final self-reflection survey for UG2 (Table 2).

Table 1. Questions analyzed from institutionally required student opinion surveys for traditionally and ungraded sections of an undergraduate exercise physiology course

Institutional student opinion survey	Response options
The amount I learned in the course was. . .	(0) Very poor (1) Poor (2) Fair (3) Good (4) Very good (5) Excellent
The evaluative and grading techniques were. . .	
Was this class intellectually stimulating? Did it stretch your thinking? Why or why not?	Open-ended questions (Response not required)
What aspects of this class contributed most to your learning?	
What aspects of this class detracted from your learning?	
What suggestions do you have for improving the class?	

Table 2. Questions analyzed from the final student self-reflections from the ungraded sections of an undergraduate exercise physiology course

Final student self-reflection	Response options
I learned a lot in this course.	Strongly disagree Disagree Agree Strongly agree
I valued my learning in this course.	
My instructor valued my learning in this course.	
[Overall I felt. . .] my learning was enhanced with ungrading.	
Use this space to write any additional comments or feedback.	Open-ended questions (Response not required)
Final student self-reflection: Additional questions added in UG2	Response options
[Overall I felt. . .] ungrading motivated me to do my work more than traditional grading.	Strongly disagree Disagree Agree Strongly agree
[Overall I felt. . .] ungrading did NOT detract from my learning.	
[Overall I felt. . .] my understanding of the course content was enhanced with ungrading more than traditional grading.	
[Overall I felt. . .] my ability to remember course content was enhanced with ungrading more than with traditional grading.	
[Overall I felt. . .] I liked receiving feedback on my work better than a letter or number grade.	
How did you feel about ungrading this semester? For example, did it produce or reduce grade-based anxiety and why? Did you feel as though it enhanced or decreased your learning and why? Other comments?	Open-ended questions (response required)
Compared to traditional grading, HOW did ungrading help your learning?	
Compared to traditional grading, HOW did receiving feedback rather than a letter or number grade help your learning?	

TG1 and TG2 took place over 16-week semesters. The class met three days per week for 55-minute sessions. The lab associated with the course met once per week for 80 minutes. UG1 was taught primarily online in a four-week block with 16 total class days. Each class day consisted of three hours of morning class time and two three-hour afternoon lab periods per week. UG2 was taught face to face during a 16-week semester-long course in the same format as TG1 and TG2. Students received both qualitative feedback and a number grade for each assignment, a traditional grade approach, in TG1 and TG2. Table 3 shows the weighting system used to calculate final course grades. The

alignment of course percentages to letter grades found in Table 4 were standardized by the institution.

Table 3. Course requirements for traditionally graded and ungraded sections of exercise physiology and assignment weighting used in the traditionally graded sections

TG course requirements	UG course requirements
Reading assignments/class participation (15%)	Reading and in-class assignments
Group research project (all associated assignments, including peer evaluation) (25%)	Group research project (all associated assignment, including satisfactory peer evaluation)
Problem sets 3 at 10% each	Problem sets and self-evaluations (x3)
Lab reports 5 at 6% each	Lab reports*
	Midterm and final self-reflection surveys and instructor meetings

* Because UG1 occurred during the height of the COVID-19 pandemic, the group research project was performed during lab sessions and students were not required to write additional lab reports to earn a grade of B. During UG2, students performed three multi-week lab activities and were required to write three lab reports in addition to the other requirements to earn a grade of B.

Table 4. Letter grades and course percentages used in the traditionally graded course

Letter grade	Course percentage
A	93-100%
A/B	89-92%
B	83-88%
B/C	79-82%
C	73-78%
C/D	69-72%
D	60-68%
F	<60%

In UG1 and UG2, the instructor employed a modified grading approach. Three days before the sections began students received a link via email to an introductory video for the course in which the

instructor described the course in general, the concept of ungrading, and the ungrading procedure for the course. Short informational readings (Gibbs 2019; Kohn 2011; Stommel 2017) about ungrading were also included in the syllabus and on the course learning management site. The ungrading description included all learning assessments that students were asked to complete satisfactorily to earn a grade of “B” (3.0) in the course. Learning assessments are included in Table 3. The primary learning assessments in the TG sections are the same as those required for a “B” in the UG sections. The instructor gave qualitative feedback on all assignments to students, similar to the TG sections, but no number or letter grade. If a student earned an “incomplete” for an assignment indicating the work was not fully satisfactory, they were given the opportunity to revise their mistakes and resubmit until they earned a “complete.”

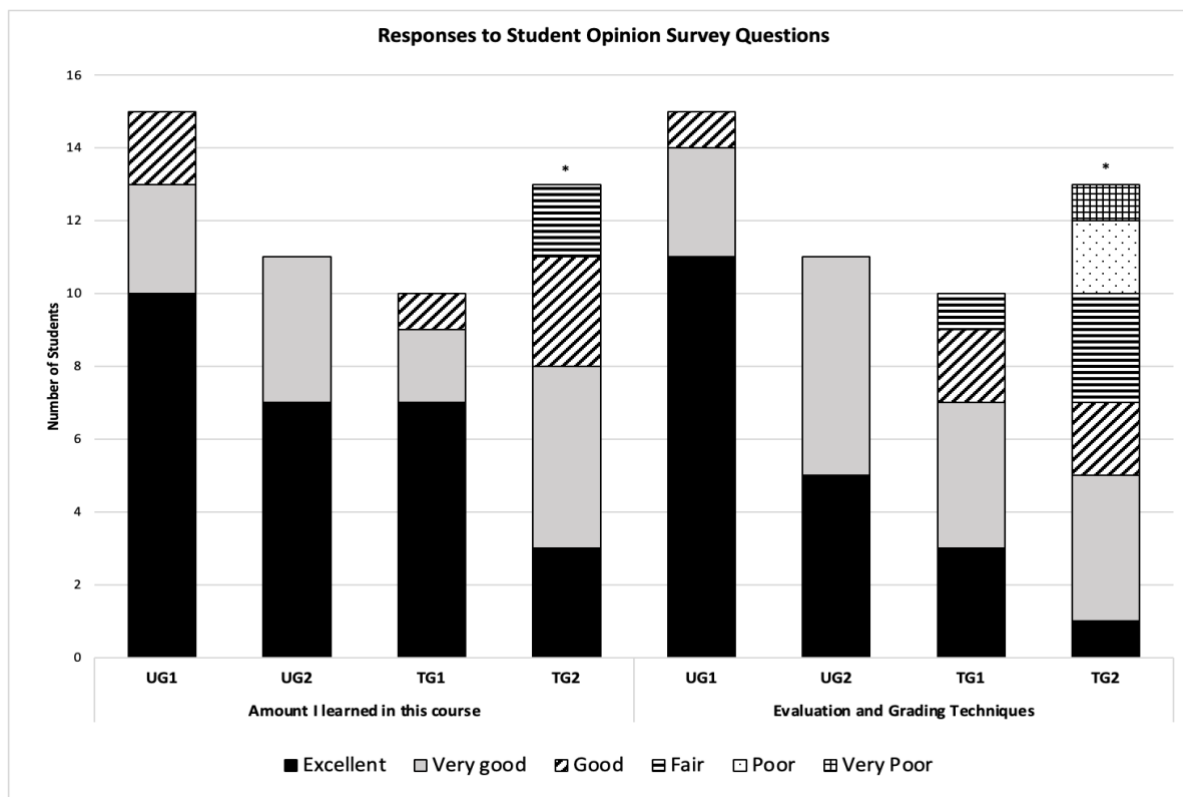
Additional evidence of understanding, engagement, and effort were required to earn greater than a “B.” Examples for evidence of advanced learning were discussed on the first day. Examples included: outstanding class engagement as demonstrated by engagement scores of “3,” thorough and thoughtful self-reflections, excellent completion and revisions of all reading assignments, meaningful synthesis of research in problem set questions, complete and thoughtful revision of the problem sets and lab reports, and demonstration of group leadership in lab. A final meeting between the student and instructor was used to co-decide the grade for each student based on their completion of work, evidence of understanding, and self-reflections of learning. More detail about the UG course structure can be found in the appendix accompanying this manuscript.

Data analysis

The researcher examined data from Likert scale questions on the institutional student opinion surveys and student self-reflection surveys with descriptive statistics and frequencies. Likert responses were converted to numerical scores from 0 to 5 (Excellent=5). A one-way Analysis of Variance (ANOVA) was used to examine differences between UG1, UG2, TG1, and TG2 for two survey questions. The researcher also used a one-way ANOVA to examine differences in final grades between the four sections. Effect sizes and least significant difference post hoc tests were also calculated. Open-ended response questions were analyzed using the thematic and content analysis method (Terry et al. 2017). The researcher de-identified comments from the final self-reflection survey before qualitative analysis was performed. Responses were read and coded independently by the researcher and an external reviewer. Independent themes were compared and finalized, which established inter-rater reliability. Open-ended responses supported interpretation of quantitative data and provide perspectives on student perceptions of the grading method.

RESULTS

Students in all course sections completed the anonymous institutionally required student opinion surveys. The response rates were 55.5% for TG1 (n=10/18), 100% for TG2 (n=13/13), 79% for UG1 (n=15/19), and 92% for UG2 (n=11/12). Figure one displays students’ ratings of the “amount I learned in this course” and the “evaluation and grading techniques.” A significant difference was found between sections for perceived learning ($F_{3,45}=4.18$; $p=0.01$; $\eta^2=0.22$). Perceived learning was significantly lower in TG2 (3.69 ± 1.03) than UG1 (4.53 ± 0.74 ; $p=0.007$), UG2 (4.64 ± 0.50 ; $p=0.005$), and TG1 (4.6 ± 0.70 ; $p=0.008$). A significant difference was found between courses for rating of the evaluation and grading techniques ($F_{3,45}=10.66$; $p<0.001$; $\eta^2=0.42$). Evaluation and grading techniques rating was significantly lower in TG2 (2.69 ± 1.49) than UG1 (4.67 ± 0.62 ; $p<0.001$), UG2 (4.45 ± 0.52 ; $p<0.001$), and TG1 (3.9 ± 0.99 ; $p=0.006$).

Figure 1. Responses to student opinion surveys for two Likert scale questions

*Indicates the mean score is significantly different from other sections.

On the self-reflection survey given only in UG1 and UG2, a combined 87.1% ($n=27/31$) of the students indicated they agreed or strongly agreed that ungrading enhanced their learning. All students in both sections either strongly agreed ($n=23$) or agreed ($n=8$) with the statement “I learned a lot in this course” (Figure 2).

Most importantly, students perceived that both they and the instructor valued their learning in the course (Figure 2). Overall, 90% ($n=28$) of students strongly agreed that they valued their own learning, and 90% ($n=28$) of students strongly agreed that the instructor valued their learning.

Student responses to the additional questions added to the self-reflection survey for UG2 are found in Figure 3. Notably, nine of 12 students perceived ungrading as motivating to complete the work of the course, and 100% of students agreed or strongly agreed that ungrading did not detract from their learning. All students agreed or strongly agreed that ungrading enhanced their understanding of the course content and their ability to remember course content compared with other traditionally graded courses. All students either agreed or strongly agreed that they liked receiving feedback better than a letter or number grade on their work.

Figure 2. Responses to student self-reflection questions on three Likert-scale questions

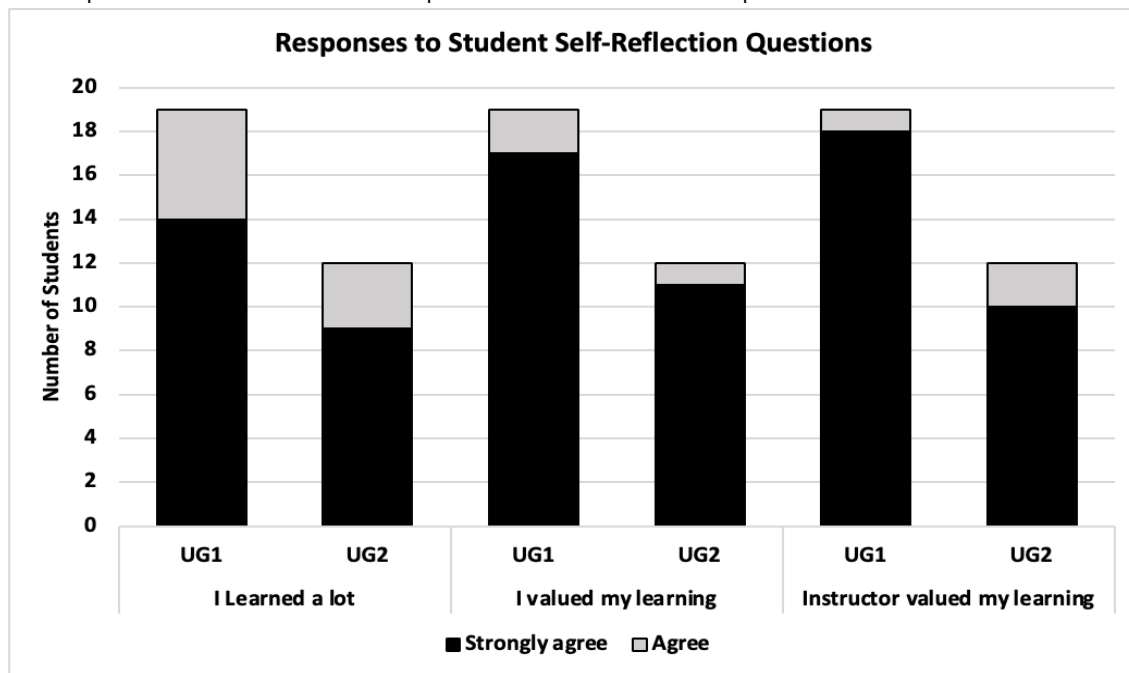
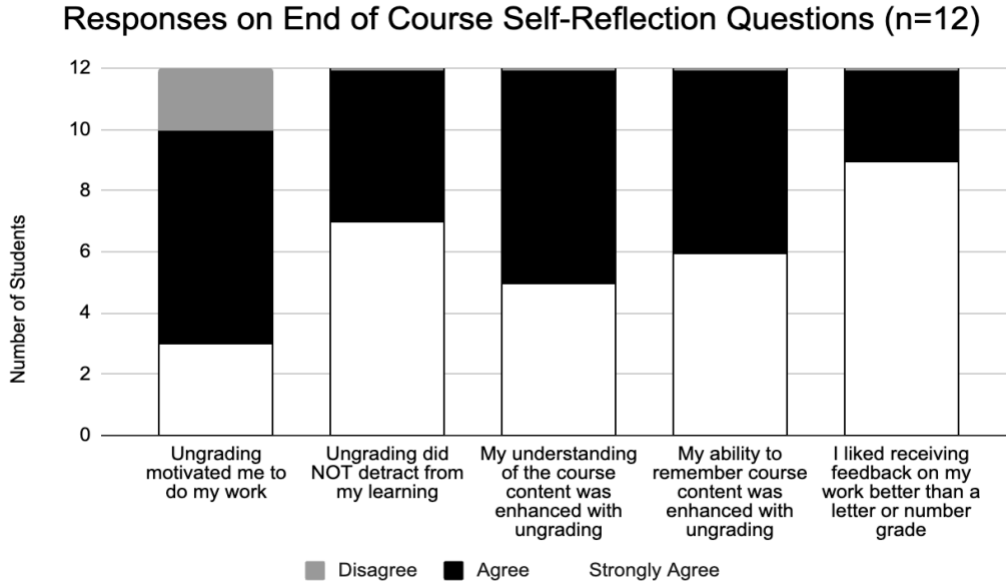
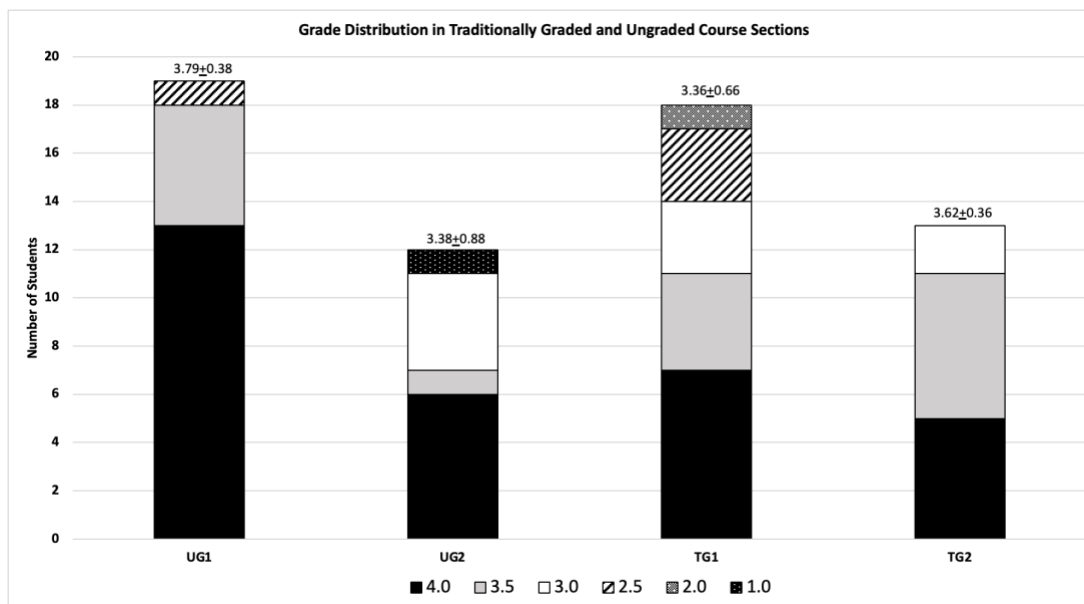


Figure 3. Responses on end-of-course self-reflection questions for five Likert-scale questions added to the self-reflection during UG2



Grade distributions and mean final grades for the UG and TG sections of the course are found in Figure 4. There was no significant difference in final grades between the four sections ($F_{3,58}=2.06$; $p=0.11$; $\eta^2=0.10$).

Figure 4. Final grade distributions for each section of the course



Values above each bar are mean \pm SD grades for each section.

In UG1 and UG2 students provided input on their final course grades which aligned with their work in the course and their articulation of learning from the self-reflection survey. Upon self-reflection, two students in each section indicated they should earn grades of B or higher without meeting the minimum criteria for a B. In these cases, the final instructor meeting to co-determine final grades was crucial. The instructor and student reviewed the student's coursework and determined that since the student did not meet the requirements outlined to earn a B in the course, a different grade should be assigned. The instructor meeting gave students a chance to further articulate their learning and to provide more context for their self-reflection responses.

Qualitative analysis of open-ended comments

Section TG1

Two main themes emerged from analysis of the open-ended comments from the institutionally required student opinion surveys in TG1 (Table 1). First, students indicated that the course was challenging and intellectually stimulating. They noted that in-class lectures contributed to their learning. Students said: "This course was intellectually stimulating. It expanded my understanding of exercise physiology and forced me to reevaluate previous beliefs and thoughts on many of the subjects covered in the course," and "the lectures and explanations contributed the most to my understanding of the course material."

Second, the students noted that the course was too much work and that the work was anxiety-producing. One student wrote, "Lab reports were especially challenging for me and I think it would've been helpful to look at a previous one in order to make less mistakes starting off." Another student noted, "I think the lab reports were hard for me and caused me a lot of anxiety. . . I thought they were really useful, and I learned a lot from them, I just felt really overwhelmed by them."

Section TG2

Two main themes emerged from the TG2 student open-ended comments on the institutionally required student opinion surveys (Table 1). First, the course was challenging. In some comments, this theme was positive and reflected the intellectual stimulation students perceived. One student commented:

This class challenged me, but it's the main reason why I loved it. Every topic we covered I always found myself using other resources besides our books, such as the internet or scholarly articles to become more aware of the topic. Overall it stretched out my thinking and made me take what I've learned and placed it into my everyday life.

In other comments, this theme was negative such that the course was too challenging and required too much time and effort:

I think the grades I earned on our assignments did not represent the amount I actually knew which was very frustrating. I put substantial amounts of time into each assignment and searched a lot of literature and I felt my grades did not reflect that.

The second theme was related to learning from assignments. One student wrote, "There were so many occasions where we needed research. I did feel like I learned things when reading the research, but it did, at times, feel like I wasn't learning anything." Another student noted, "I did not like the problem sets. We were asked questions on things that we did not go over in class. This made it very difficult for me to feel confident that what I was putting for my answer was the right one."

Section UG1

Three main themes developed from the qualitative analysis of open-ended questions on the final student self-reflection survey (Table 2) and the institutional student opinion survey (Table 1). The most prominent theme was that students learned by making and fixing mistakes in their work. One student said, "I learned to be okay with making mistakes in my reading assignments because I learned from those mistakes." Other students reflected:

I noticed that I obtained and retained things that I have learned when I make a mistake in applying or explaining it to someone else. When I made a mistake in the class discussion or when answering a question about it in the smartbook assignments or class, I had a mental note there to remember the mistake I made and to not make the same mistake again.

Editing our problem sets was a very helpful process, especially when combined with other research articles. They served as a way to a) check your initial work and b) see how NEW information connects to your initial responses.

The second theme was that students understood the course concepts well. One student noted, "I also enjoyed that I got to demonstrate my learning by doing the problem sets and revising them instead of having exams. They took the emphasis off memorizing content for a test and allowed me to work to apply the content." Another student said, "The ungrading helped me actually process the content and learn it instead of memorizing facts and spitting them out on a test."

The third theme was that ungrading reduced grade-based anxiety. One student said, “I also loved the ungrading system because it allowed me to focus on my knowledge instead of focusing on my grade.” Another student commented, “[Ungrading] relieves the stress of worrying about what grade I will end up with and allowed me to feel like I was more focused on what I was learning instead of what will be on the test and if that will accurately reflect my knowledge.”

Section UG2

Three main themes emerged from the qualitative analysis of the open-ended questions on the final student self-reflection survey (Table 2) and the institutional student opinion survey (Table 1). The first and most prominent theme was that ungrading reduced grade-based anxiety. A student noted, “I think ungrading reduced the grade-based anxiety. I didn’t have to worry about getting a certain score, so I feel like it enhanced my learning as I was able to deepen my knowledge about the topic.” Another student said, “It definitely reduced my anxiety without completely eliminating it. I realized that although a specific number won’t hurt me, I still have to do quality work.” A third student expanded on why they felt less anxiety:

It reduced grade base anxiety and at some point I found myself caring a lot more about this course than my other courses in which I could actually fail assignments. I think this is because it was more rewarding to receive a comment assuring that I have learned something, or a comment explaining something I was misunderstanding than a letter grade.

The second most prominent theme was that students understood the concepts well, demonstrating that students perceived understanding as more important than short term memorization. One student said, “Ungrading helped my learning because I didn’t just have to memorize something and then forget about it. I was able to understand the topics on a less superficial level and felt as though I could remember a lot more information.” Another student reflected, “It decreased my grade-based anxiety and it enhanced my learning by allowing me to actually learn the material in contrast to memorizing a study guide/answers.”

The third theme was that students learned from feedback. One student commented, “I genuinely feel like I learned way more with feedback than a letter grade because I was able to learn from my mistakes without fear of making a mistake and costing my grade.” Another student noted:

I think that it made me more willing to do my work because I knew that I wouldn’t be getting a 4/5 for something I would just be getting feedback because learning is a process where you aren’t always right so it was nice to get acknowledged for completing the work and then knowing where I didn’t understand a concept with the feedback.

A third student reported on their increased understanding due to feedback:

Feedback allows me to consider the quality of the answer and throughout each PS it got easier to answer the questions. My responses lacked step by step explanation of how the body is responding or functioning in regards to the question. Receiving feedback made me consider The Who, What, Where and Why components of the body functioning which would be a reflection of whether I comprehended the concept or not.

A fourth, less prominent theme present in both ungraded courses was that students felt they had freedom to take risks in their learning. This theme manifested in two ways. First, students were motivated to investigate topics of interest outside of class. The student indicated they had more freedom to be curious and grow as learners. One student noted, “I got more feedback on how to improve or that I was on the right track. This really allowed me to grow as a student and learn more in the course and about my work ethic.” Second, students indicated a freedom to make mistakes which was accompanied by a decreased fear of failure. Normalizing errors and allowing revisions helped reduce grade-based anxiety and helped students focus on learning. One student explained, “It gave me more freedom to take risks in my work and not be afraid to get something wrong.”

Across both sections, only one student was critical about ungrading. This student wrote:

I personally learn better when motivated with traditional grading styles. While I appreciated using the ungrading policy, which also allowed me to learn what I am passionate about and not merely memorize material, I do well being pushed to learn for a grade. However, learning by this method also allowed me to stay curious about topics I would not have asked myself if we focused on traditional grades.

DISCUSSION

Ungrading can successfully replace traditional forms of assessment in causal reasoning-based disciplines, such as exercise physiology, in order to improve perceived learning, reduce students’ grade-based anxiety, and increase students’ reflection on their learning. The primary finding of this study was that students reported greater satisfaction and greater perceived learning with ungrading than traditional grading, while performing similarly on final course grades, dispelling the notion that ungrading causes grade inflation. Two conclusions can be drawn from the qualitative analysis of student comments. First, students perceived their learning as enhanced by making and fixing mistakes. Second, students perceived a greater understanding of the course content without feeling a need to memorize material. Students noted they could focus on understanding and not on grades, which reduced grade-based anxiety. Less grade-based anxiety allowed students to feel as though they could take risks, make mistakes, and feed their curiosity through enhanced engagement with course content.

Because exercise physiology courses are structured with scaffolded content to demonstrate the interconnectedness of body systems, the potential effect of correcting mistakes on future work is noteworthy. Revising work also incorporates repetition. Repeatedly asking students to refer back to and explain concepts fosters learning and enhances causal reasoning skills which are important to understanding course content within scaffolded learning structures (Dobson, Linderholm, and Perez 2018; Michael 2007). In the current study, students felt that normalizing mistakes and correcting them was helpful in furthering their understanding of course content. When students are required to respond to the instructor feedback by resubmitting assignments, they may improve performance on subsequent tasks, such as application, and are afforded another opportunity to demonstrate their understanding (Butler and Nisan 1986; Schinske and Tanner 2014). Additionally, when students received descriptive feedback without grades, they performed significantly better on follow-up assignments than students who received grades or no feedback (Butler and Nisan 1986). These findings are applicable to other disciplines, especially those based in causal reasoning and those with scaffolded learning.

Students reported reduced grade anxiety in the ungraded sections because of the opportunity to revise their work. Students tend to associate grades with learning or understanding and students often tend to place more emphasis on grades over understanding (Marcus and Tomasi 2020; Pulfrey, Buchs, and Butera 2011). Grades provide extrinsic motivation for students, which in turn limits intrinsic motivation for learning to understand instead of striving for a grade (Chamberlin, Yasué, and Chiang 2023; Kohn 2011). Reducing grade anxiety and fostering intrinsic motivation for learning allows students in any discipline to focus on understanding and applying course content. Students noted in open-ended responses that they felt freer to investigate topics of interest outside of the class reading or discussion, further evidence of intrinsic motivation for learning.

In both ungraded courses, students perceived that they understood the course material well without memorization. While grades may be rewarding for tasks such as memorization, they do not promote deep processing or long-term retention (Chamberlin, Yasué, and Chiang 2023). Students without a complete understanding of the material may feel they need to memorize facts in order to demonstrate learning but lack the ability to explain how they have arrived at their answer when applying the information. Students rely on memorizing likely because it was a study strategy they successfully used previously in their education (Momsen et al. 2013).

In both ungraded sections, the instructor repeatedly emphasized the importance of being able to explain and demonstrate understanding of the processes and deemphasized memorizing. Students' perceptions that they were able to demonstrate their understanding reflects how much students thought about their own learning in UG sections. This contrasts with the limited and superficial comments about learning from students in the TG sections. TG students lacked reflective prompting from the instructor and equivalent transparency in course design to the UG sections. Students in TG2, for example, commented that they did not learn much while researching answers to questions or did not understand that the course content was intended to be applied within the problem sets. While students in the TG courses did perceive they learned a lot, they could not articulate why, how, or what contributed to their learning other than lecture and discussion. Students did not perceive significant learning from the major assignments in the course. It is possible the students struggled to understand the purpose of the assignments and did not engage in the same degree of metacognition as students in UG sections who were more frequently prompted to reflect on their learning. Ungrading and the emphasis on metacognition in this course made students more aware of how and what they were learning in the course. Asking students to reflect on their learning can occur without ungrading; however, metacognition is a hallmark of ungrading which did not receive as much prominence in the TG courses as the UG courses.

Dissociating grades from assessments in ungraded sections meant the students could demonstrate what they understood without the pressure of missing points they could not recover. Revising mistakes integrates well with ungrading because students can earn full credit on all assignments if they choose to revise their work and do so satisfactorily. Some may argue that allowing revisions can contribute to grade inflation. For example, if students could continually revise, they could all earn A's in the course. In the present study, final grades were not different between grading methods, indicating that grade inflation did not occur with ungrading. The opportunity to revise work likely helped the students demonstrate understanding of content and more accurately gauge their learning. From the instructor's perspective, students in the ungraded section were more engaged with course content both directly (in-class, verbally) and indirectly (outside of class, revising assignments). Not only did all students take the opportunities to revise their work, but they were also more vocal in class asking and answering questions. The instructor's perception of engagement could have influenced the students' perceptions. Instructor comments such as "great discussion today" or

increased instructor enthusiasm due to more direct, verbal engagement may have reinforced and increased frequency of those behaviors in the classroom.

Limitations

This paper's conclusions are limited to a single instructor and an upper division course with small section enrollments. While some aspects of the course can be scaled up to larger class sizes, others may present challenges to instructors, with feedback workload being the largest challenge. From the instructor's perspective, the workload associated with providing feedback for this course was similar between UG and TG sections. Qualitative feedback was not different between the two grading methods. The instructor, however experienced a reduced mental workload in the UG section as they didn't have to assign points. Being able to provide feedback without justifying a specific number or letter grade was less stressful than the constant decision making for allotted points that accompanied the feedback. In larger class sections, providing detailed feedback on some main assignments as opposed to all assignments or providing more general feedback to the entire class may help reduce the feedback workload.

The course content and evaluation measures were consistent between the two UG sections; however, the section structure was very different. Differences in student comments and student learning between the sections could be due to the intensity of the four-week course during which students did not take any other courses and were able to focus their attention on exercise physiology. The four-week course structure used in UG1 was an institutional change implemented during the COVID-19 pandemic. This structure, however, is similar to a typical summer section delivery structure, whereas UG2 followed a typical 16-week semester schedule.

Since the reflective questions posed to students differed between sections, the perceptions are constricted to individual sessions. Institutionally mandated student opinion survey questions were the same for all sections. Students were asked to perform significant self-reflection in the ungraded sections, especially in UG2 when specific questions about the effects of ungrading on learning were included in student reflections. Student perceptions may have also been biased by instructor enthusiasm and the initial information presented about ungrading and issues with traditional grading before the courses began. An explicit emphasis on demonstrating understanding and increased transparency about course design and assessment in ungraded sections also pose major limitations to the comparison between grading methods.

The relative emotional and intellectual development of the students also poses a limitation. Students in these course sections were upper-division (third and fourth year) students who were likely in the "contextual relativism" stage of Perry's Scheme of Intellectual and Ethical development (Perry 1981). In this stage, students recognize that solutions are supported by evidence in context and because some solutions are better than others, the student's task is to evaluate solutions. In the present study, students were asked to perform self-evaluation on major assignments and provide evidence of their learning in the context of the course. Questions that had no clear correct answer prompted this self-evaluation, and students were tasked with providing evidence for their response.

While ungrading might be better suited for students with more advanced metacognitive skills and intellectual development, starting ungrading with students earlier in their intellectual development, perhaps with a partially ungraded course, several low stakes ungraded assignments, or integrating assignment self-evaluation or self-reflection of learning into a course may help them build metacognitive and critical thinking skills as they move through their college career. Instructors using ungrading should investigate student and faculty perceptions of different methods of ungrading and the use of ungrading with larger class sizes. Not only should students reflect on their learning, but also

on the approaches used to assess their learning as this is truly putting the students at the center of the learning experience. Instructors should also consider measuring content knowledge acquisition and understanding in ungrading and traditional grading, perhaps using a standardized assessment.

CONCLUSION

Overall, ungrading increased student satisfaction with grading and perceived learning in exercise physiology without changes in perceived rigor or final grades when compared to sections with traditional grading. The combination of reducing the focus on grades, encouraging understanding over memorization, and providing opportunities to learn from mistakes created a student-centered course environment where students perceived reduced grade anxiety.

ACKNOWLEDGMENTS

I would like to thank L. Taylor and J. Schaefer for their assistance in designing the ungraded courses described in this paper. I am also grateful for the editing and feedback provided by L. Taylor.

AUTHOR BIOGRAPHY

Mary Stenson, PhD (USA) is associate professor and program coordinator for the Exercise and Rehabilitation Sciences Program at the University of Minnesota Duluth. Her research interests are in physiological recovery from exercise and exercise science pedagogy.

DISCLOSURES

The author has no conflicts of interest to disclose. No funds were used from any source for the production of this research. AI was not used in any way during the course of this research or the preparation of this manuscript.

ETHICS

Research was reviewed by the College of Saint Benedict/Saint John's University Institutional Review Board.

REFERENCES

- Bas, Gokhan. 2021. "Relation Between Student Mental Health and Academic Achievement Revisited: A Meta-Analysis." *Health and Academic Achievement - New Findings*. <https://doi.org/10.5772/intechopen.95766>.
- Blum, Susan Debra, editor. 2020. *Ungrading: Why Rating Students Undermines Learning (and What to Do Instead)*, first edition. Morgantown: West Virginia University Press.
- Brookhart, Susan M. 2017. *How to Give Effective Feedback to Your Students*, second edition. Alexandria, Virginia: ASCD.
- Butler, Ruth. 1988. "Enhancing and Undermining Intrinsic Motivation: The Effects of Task-Involving and Ego-Involving Evaluation on Interest and Performance." *British Journal of Educational Psychology* 58 (1): 1–14. <https://doi.org/10.1111/j.2044-8279.1988.tb00874.x>.
- Butler, Ruth, and Mordecai Nisan. 1986. "Effects of No Feedback, Task-Related Comments, and Grades on Intrinsic Motivation and Performance." *Journal of Educational Psychology* 78 (3): 210–16. <https://doi.org/10.1037/0022-0663.78.3.210>.
- Chamberlin, Kelsey, Maï Yasué, and I-Chant A. Chiang. 2023. "The Impact of Grades on Student Motivation." *Active Learning in Higher Education* 24 (2): 109–24. <https://doi.org/10.1177/1469787418819728>.
- Crisp, Beth R. 2007. "Is It Worth the Effort? How Feedback Influences Students' Subsequent Submission of Assessable Work." *Assessment & Evaluation in Higher Education* 32 (5): 571–81. <https://doi.org/10.1080/02602930601116912>.

- Deci, Edward L., and Richard M. Ryan. 2002. "The Paradox of Achievement." In *Improving Academic Achievement*. 61–87. Elsevier. <https://doi.org/10.1016/B978-012064455-1/50007-5>.
- Dobson, John, Tracy Linderholm, and Jose Perez. 2018. "Retrieval Practice Enhances the Ability to Evaluate Complex Physiology Information." *Medical Education* 52 (5): 513–25. <https://doi.org/10.1111/medu.13503>.
- Farias, Gerard, Christine M. Farias, and Kent D. Fairfield. 2010. "Teacher as Judge or Partner: The Dilemma of Grades Versus Learning." *Journal of Education for Business* 85 (6): 336–42. <https://doi.org/10.1080/08832321003604961>.
- Frank, Nicholas, and James Sutherland-Smith. 2021. "Evidence-Based Approach to Switching to a Pass-Fail System for Clinical Year Veterinary Student Grading." *Journal of Veterinary Medical Education* 48 (4): 503–10. <https://doi.org/10.3138/jvme.2019-0103>.
- Gibbs, Laura. 2019. "OU Digital Teaching: Getting Rid of Grades (Book Chapter)." *OU Digital Teaching* (blog). <http://oudigitools.blogspot.com/2019/03/getting-rid-of-grades-book-chapter.html>.
- Gorichanaz, Tim. 2022. "'It Made Me Feel Like It Was Okay to Be Wrong': Student Experiences with Ungrading." *Active Learning in Higher Education*, May. <https://doi.org/10.1177/14697874221093640>.
- Harackiewicz, Judith M., Steven Abrahams, and Ruth Wageman. 1987. "Performance Evaluation and Intrinsic Motivation: The Effects of Evaluative Focus, Rewards, and Achievement Orientation." *Journal of Personality and Social Psychology* 53: 1015–23. <https://doi.org/10.1037/0022-3514.53.6.1015>.
- Hayek, Anne-Sophie, Claudia Toma, Sofia Guidotti, Dominique Oberlé, and Fabrizio Butera. 2017. "Grades Degrade Group Coordination: Deteriorated Interactions and Performance in a Cooperative Motor Task." *European Journal of Psychology of Education* 32 (1): 97–112. <https://doi.org/10.1007/s10212-016-0286-9>.
- Kohn, Alfie. 1999. "From Degrading to De-Grading." *High School Magazine*. <https://www.alfiekohn.org/article/degrading-de-grading/>.
- Kohn, Alfie. 2011. "The Case Against Grades." *Alfie Kohn* (blog). November 2, 2011. <https://www.alfiekohn.org/article/case-grades/>.
- Malouff, John M., Ashley J. Emmerton, and Nicola S. Schutte. 2013. "The Risk of a Halo Bias as a Reason to Keep Students Anonymous During Grading." *Teaching of Psychology* 40 (3): 233–37. <https://doi.org/10.1177/0098628313487425>.
- Marcus, Mackenzie, and David Tomasi. 2020. "Emotional and Cognitive Responses to Academic Performance and Grade Anxiety." *Journal of Medical Research and Health Sciences* 3 (4): 919–23. <https://doi.org/10.15520/jmrhs.v3i4.172>.
- Melrose, Sherri. 2017. "Pass/Fail and Discretionary Grading: A Snapshot of Their Influences on Learning." *Open Journal of Nursing* 7 (2): 185–92. <https://doi.org/10.4236/ojn.2017.72016>.
- Michael, Joel. 2007. "What Makes Physiology Hard for Students to Learn? Results of a Faculty Survey." *Advances in Physiology Education* 31 (1): 34–40. <https://doi.org/10.1152/advan.00057.2006>.
- Milton, Ohmer, Howard R. Pollio, and James A. Eison. 2010. *Making Sense of College Grades*. San Francisco: Jossey-Bass.
- Momsen, Jennifer, Erika Offerdahl, Mila Kryjevskaja, Lisa Montplaisir, Elizabeth Anderson, and Nate Grosz. 2013. "Using Assessments to Investigate and Compare the Nature of Learning in Undergraduate Science Courses." *CBE—Life Sciences Education* 12 (2): 239–49. <https://doi.org/10.1187/cbe.12-08-0130>.
- Newell-Caito, Jennifer. 2024. "Exploring Ungrading in a Biochemistry Laboratory Course." In *Inclusive STEM: Transforming Disciplinary Writing Instruction for a Socially Just Future*, edited by: Heather M. Falconer and LaKeisha McClary. 223–46. Denver: University Press of Colorado. <https://doi.org/10.37514/ATD-B.2024.2364.2.14>.
- Patton, Kevin, and Staci Johnson. 2021. "Ungrading w/ Standards-Based Grading | Chat w/ Staci Johnson." Accessed April 19, 2023. <https://theapprofessor.org/podcast-episode-106.html>.
- Perry, William G., Jr. 1981. "Cognitive and Ethical Growth: The Making of Meaning." In *The Modern American College*, edited by: Arthur W. Chickering. 76–116. San Francisco: Jossey-Bass.
- Price, Michael J, and Christopher Simons. 2010. "The Effect of Sodium Bicarbonate Ingestion on High-Intensity Intermittent Running and Subsequent Performance." *Journal of Strength and Conditioning Research* 24 (7):

- 1834–42. <https://doi.org/10.1519/JSC.0b013e3181e06e4a>.
- Pulfrey, Caroline, Céline Buchs, and Fabrizio Butera. 2011. “Why Grades Engender Performance-Avoidance Goals: The Mediating Role of Autonomous Motivation.” *Journal of Educational Psychology* 103 (3): 683–700. <https://doi.org/10.1037/a0023911>.
- Schinske, Jeffrey, and Kimberly Tanner. 2014. “Teaching More by Grading Less (or Differently).” *CBE—Life Sciences Education* 13 (2): 159–66. <https://doi.org/10.1187/cbe.14-03-0054>.
- Soric, Mate M., Jennifer D. Robinson, and Timothy R. Ulbrich. 2019. “Grade Point Average Is an Inappropriate Measure of a Residency Candidate’s Knowledge and Skills.” *Journal of the American College of Clinical Pharmacy* 2 (2): 177–80. <https://doi.org/10.1002/jac5.1076>.
- Spring, Laura, Diana Robillard, Lorrie Gehlbach, and Tiffany A Moore Simas. 2011. “Impact of Pass/Fail Grading on Medical Students’ Well-Being and Academic Outcomes.” *Medical Education* 45 (9): 867–77. <https://doi.org/10.1111/j.1365-2923.2011.03989.x>.
- Stommel, Jesse. 2017. “Why I Don’t Grade.” Jesse Stommel. <https://www.jessestommel.com/why-i-dont-grade/>.
- Stommel, Jesse. 2023. *Undoing the Grade: Why We Grade, and How to Stop*. Madison: Hybrid Pedagogy Inc.
- Terry, Gareth, Nikki Hayfield, Victoria Clarke, and Virginia Braun. 2017. *The SAGE Handbook of Qualitative Research in Psychology*. London: SAGE Publications Ltd. <https://doi.org/10.4135/9781526405555>.
- Walker, Christopher O., Barbara A. Greene, and Robert A. Mansell. 2006. “Identification with Academics, Intrinsic/Extrinsic Motivation, and Self-Efficacy as Predictors of Cognitive Engagement.” *Learning and Individual Differences* 16 (1): 1–12. <https://doi.org/10.1016/j.lindif.2005.06.004>.
- West, Colin P., Steven J. Durning, Bridget C. O’Brien, John H. Coverdale, and Laura Weiss Roberts. 2020. “The USMLE Step 1 Examination: Can Pass/Fail Make the Grade?” *Academic Medicine* 95 (9): 1287. <https://doi.org/10.1097/ACM.0000000000003537>.

APPENDIX

IMPLEMENTATION OF UNGRADING

Learning goals

The course included six main content areas: bioenergetics of exercise, neuromuscular control and function, cardiorespiratory physiology, acid-base and fluid balance, exercise in extreme environments (heat, cold, altitude), and fatigue. The acute responses and chronic adaptations to different types of exercise were discussed in each content area. In addition to content knowledge related learning goals associated with each of the content areas above, the course and lab were uniquely designed around several departmental academic skills-based learning goals including: evidence based decision making, professionalism and teamwork, skilled communication, and ethically guided decisions. The learning goals and course content in ungraded sections of exercise physiology were the same as previous semesters when a traditional weighted grading system was used.

Student learning assessments

The student learning assessments were designed to meet the learning goals and allow students to demonstrate an understanding of the concepts with an emphasis on creativity of thought, application of concepts, and causal reasoning. The assignments were the same in both types of grading; however, there were some differences in course set up in the ungraded sections of the course.

For most class periods (approximately 85%), students were asked to complete a pre-class reading assignment which served as a check for reading comprehension. In each ungraded section of the course, students were also assigned pre-class video lectures with the intent to preserve in class time for discussion and application activities. In-class and post-class application assignments included case studies, concept maps, graph interpretation, and application problems. These assignments lend themselves well to ungrading because the assignments capture varying degrees of a student's skills and understanding which go beyond forced choice questions with a single correct answer. Not only do student responses provide opportunities for specific feedback, but the instructor can more easily identify misconceptions and allow opportunities for reassessment without a grade penalty. Additionally, assigning a number grade equitably across an assignment, such as a concept map or data prediction, would require an extensive rubric which may not fully capture the interconnected relationships of physiological processes, or the diversity of thought and creativity in a student response.

Ungrading allows the exercise physiology instructor to recognize degrees of understanding which go beyond correct or incorrect work. For example, students created a concept map to demonstrate understanding of how glucose, fatty acid, and amino acid breakdown processes interact with one another to fuel activity. In this case, students may make different connections between bioenergetic processes depending on feeding status or the type or intensity of exercise. Another concept map was created around changes in blood pressure during exercise. Students then applied the connections and changes to different situations such as prolonged exercise in a hot environment or exercise in the cold. On any concept map, a student may have correctly labeled a connection on the map, but the context or detail of the connections might differ between students.

Interpreting data on graphs also helped students practice reading graphs and applying or anticipating changes with new information. For example, students interpreted a graph of changes in mean skin and core temperature during exercise in the heat. Students explained the graph using their knowledge of thermoregulation and made predictions for temperature changes as the exercise

progressed beyond the graph. Additional hypothetical questions about fluid consumption, humidity, wind speed, and intensity required students to demonstrate their understanding of thermoregulation.

Problem solving and application problems were used in both lab report discussion questions and problem sets. For each lab students wrote a results section and answered discussion questions which were based on interpreting and applying the data collected in the lab. Discussion questions began with data interpretation and ended with application questions for the lab topic (Figure A1).

Figure A1. Lab report and problem set sample questions

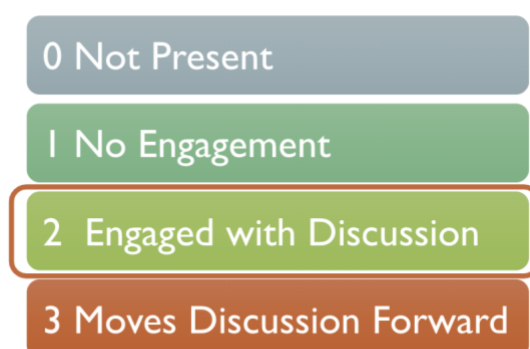
Lab report discussion question examples
<p>Metabolic Responses to Upper and Lower Body Resistance Exercise</p> <ol style="list-style-type: none"> 1. Discuss the impact of resistance exercise on plasma volume changes during exercise. Were our results expected, why or why not? 2. Describe the differences in metabolic variables between the upper body exercise and the lower body exercise. Explain any differences that occurred, whether or not those changes were expected, and why or why not. 3. Describe the causes of delayed onset muscle soreness (DOMS). What types of exercise commonly result in DOMS and why? Explain the advantages of the repeated-bout effect.
Problem set question examples
<ol style="list-style-type: none"> 1. Is the thirst response during and/or after exercise different (greater or less than) in a cold environment compared to a thermoneutral environment? Does the thirst response in a cold environment change if sweating is profuse? 2. From an exercise biochemistry and metabolic perspective, why can women compete with (or out-compete!) men at ultra-endurance distances, but not at shorter endurance distances (half-marathon, marathon)? 3. Sodium bicarbonate (NaHCO_3) has been investigated as a potential ergogenic aid. Carefully examine the following graphs and answer the questions below. (Figures 1 and 4 provided from Price and Simmons, 2010) (Price and Simons 2010) <ol style="list-style-type: none"> a. What type of exercise would potentially benefit most from using sodium bicarb as an ergogenic aid? Why? b. After examining Figure 1, explain the physiological theory behind using NaHCO_3 as an ergogenic aid (in other words, what kind of physiological responses are expected from using NaHCO_3?). What is the effect demonstrated in the figure? c. After examining Figure 4, is NaHCO_3 an effective ergogenic aid? Explain how you arrived at this conclusion and provide a physiological explanation for your conclusion.

While lab report discussion questions also required students to use data collected in the lab to support their responses, the problem sets allowed more advanced thinking about exercise physiology topics. Problem sets were assigned as take-home exams in the course and required students to make connections between the class content and new ideas. Students used course information, textbooks, and primary research to provide a detailed response to the questions on these assignments. Students found primary research articles which sometimes led to different answers to the questions; however, using ungrading the instructor could provide feedback on the quality and use of the information in the response, the completeness of the argument, and how the responses demonstrated an understanding of course content. Because each student responded slightly differently, the ungrading approach allowed the instructor flexibility in assessing understanding of the content as opposed to forced choice exam questions which often do not fully capture a student's thought process.

Classroom engagement

In the ungraded sections, class discussion and regular class engagement was emphasized more than the traditionally graded sections. The intent of this emphasis was to improve student engagement and provide different opportunities for students to demonstrate understanding. Students were expected to make regular and thoughtful contributions during class. Daily class engagement for each student was recorded by the instructor on a four-point scale (Figure A2). During the first class, students brainstormed a list of good engagement techniques for the class which included defining the requirements for earning a “3” and options for how students could contribute in class if they felt nervous about engaging in the discussion. While the criteria for each level of engagement were fluid, the classes decided together that a “1” was appropriate for a student who was present in class, but did not engage at all by speaking in class, volunteering, or exhibited disengagement such as unauthorized computer use, sleeping, or not taking notes. A score of “2” (the standard) was characterized by some vocal participation, attentiveness, and no disengagement. A score of “3” was characterized by moving the discussion forward, asking questions or making comments which demonstrated an advanced understanding of the material or applied the content in a new way. During each class day, students were encouraged to ask questions about the material that they did not understand, to answer questions posed by the instructor or other students, to apply the material to a new concept, or to make connections to a previous class or a real-world situation. Each student had ample opportunity for outward class engagement daily. While the daily engagement scores did not count toward earning a grade of B in the course, students could use their engagement scores as evidence of understanding during their individual meetings with the instructor to co-determine the final grade.

Figure A2. Four-point daily class engagement scale used during class meetings. A score of “2” was considered the standard for daily engagement



Feedback on assessments

During the traditionally graded sections qualitative feedback was provided and a number grade was assigned with partial credit allowed for all assignments. In the ungraded sections, all assignments were marked as complete or incomplete in the online learning management system. If the student’s response required additional work or contained incorrect answers, feedback was provided, and students were given the option to revise and resubmit the assignment to earn a

“complete.” The intent of resubmission was to allow students to correct and learn from their mistakes after receiving feedback. Students could choose if they wanted to resubmit the work or not. The time commitment for instructor feedback on assignments varied from a few seconds per student for reading assignments (e.g., “nice job explaining the process”) to approximately 10-12 minutes per student for lab reports or problem sets. Feedback on longer assignments was more detailed and, when necessary, identified student misconceptions and where the students might improve their work (e.g., “this explanation of substrate use during exercise does not consider the intensity of the exercise, what is the exercise intensity and how would that change your response here?”)

Instructor meetings and student self-reflection

At the midpoint of the course students completed a mid-term self-evaluation and attended a five-minute meeting with the instructor about their responses. In areas where students needed improvement, the instructor guided them to think about actionable steps to improve in the second half. For example, for students who did not regularly participate in class discussions, the student and instructor developed actionable goals to be sure the student had opportunities to contribute over more vocal classmates. Examples included making eye contact with the student to see if they had something to contribute, specifically calling on the student, or the student contributing comments such as “I was going to say that as well” or “I also noticed that about the article.” The one-on-one conversations proved to be the most successful part of the ungrading method. Students were excited to discuss what they learned and the successes and challenges they faced in the learning process. The mid-term meeting was also used as a learning progress check. The instructor answered questions about ungrading and asked students about their perception of their learning in the course.

At the end of the course, each student completed a final learning self-reflection and met with the instructor again. The final self-reflection was used to encourage students to honestly assess their learning in and engagement with the course, as well as how their learning, engagement, and work quality changed over the course. Within the self-reflection students submitted a letter grade and rationale for themselves. Students then met with the instructor for approximately five minutes to discuss their self-reflection, work in the courses, and alignment with the B grade anchor. The student presented their grade rationale and co-decided with the instructor what final grade would be submitted.



Copyright for the content of articles published in *Teaching & Learning Inquiry* resides with the authors, and copyright for the publication layout resides with the journal. These copyright holders have agreed that this article should be available on open access under a Creative Commons Attribution License 4.0 International (<https://creativecommons.org/licenses/by-nc/4.0/>). The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited, and to cite *Teaching & Learning Inquiry* as the original place of publication. Readers are free to share these materials—as long as appropriate credit is given, a link to the license is provided, and any changes are indicated.