

UC Davis

UC Davis Previously Published Works

Title

Undergraduate STEM Students' Perceptions of Grading Practices Reveal that Quiz Retakes Positively Impact Drivers of Self-determination

Permalink

<https://escholarship.org/uc/item/6db6182m>

Journal

CBE—Life Sciences Education, 24(1)

ISSN

1931-7913

Authors

Tripp, Brie

Ravi, Akshaya

Pang, Ethan

et al.

Publication Date

2025-03-01

DOI

10.1187/cbe.24-06-0167

Peer reviewed

Undergraduate STEM Students' Perceptions of Grading Practices Reveal that Quiz Retakes Positively Impact Drivers of Self-determination

Brie Tripp,^{†*} Akshaya Ravi,[†] Ethan Pang,[†] and Robert E. Furrow^{†*}

[†]Department of Neurobiology, Physiology, and Behavior, University of California, Davis, Davis, CA 95616; [†]Department of Wildlife, Fish, and Conservation Biology, University of California, Davis, Davis, CA 95616

ABSTRACT

Grades are a staple of education and a gateway to future career opportunities. Yet, grading practices can (re)produce inequities and cause students to feel inadequate and unmotivated. Alternative grading practices may address these problems, but these strategies are often time intensive and impractical in larger classroom settings. In this study, we explore an easy-to-implement grading practice, in-class quiz retakes, to conceptualize how grades motivate learning and impact well-being for undergraduate students in science, technology, engineering, and mathematics (STEM). Through the lens of self-determination theory, we conducted semistructured interviews with undergraduates who experienced quiz retakes in two STEM courses. Our results revealed that retakes largely improved students' perceptions of their competence in the subject matter, autonomy in grade outcomes, feelings of relatedness to the instructors, and overall motivation to learn. The majority of students also expressed how traditional grading practices negatively impacted their motivation and well-being. In addition, a quantitative analysis revealed that quiz retakes particularly benefitted students who scored lower on their initial quizzes. We aspire for this study to prompt educators to reconsider traditional grading practices by opting for more equitable and just alternative grading approaches that motivate student learning and mitigate systemic barriers in education.

INTRODUCTION

As instructors, we often commiserate on students' fixation regarding grades and their insistence on getting the "right" answer, often to the exclusion of deeper comprehension and learning. It is easy to become frustrated and exhausted by this without questioning what drives students' preoccupation with receiving the maximum points possible. But this grade myopia is cultivated by the structures and incentives that students experience throughout their entire education. The desire to attain exceptional grades is embedded into the fabric of our educational system from K12 to higher education and reinforced by teachers in schools and faculty instructors at universities (Feldman, 2023). And as high grade point averages (GPAs) are required to enter most graduate and professional programs (Allen, 2005), attaining stellar grades can become the defining purpose of education for students. Therefore, is it really students' fault for adhering to the standards we have habitually (and rather inequitably) reinforced?

Some instructors have tried to ameliorate this grade fixation by exploring alternative grading practices that decenter grades and offer students frequent feedback and multiple opportunities to demonstrate proficiency. But these approaches may take lengthy course revisions and require instructors to dedicate substantial time to

Cynthia Brame, *Monitoring Editor*

Submitted Jun 17, 2024; Revised Nov 7, 2024; Accepted Nov 19, 2024

CBE Life Sci Educ March 1, 2025 24:ar4

DOI:10.1187/cbe.24-06-0167

Conflicts of interests: The authors declare no conflicts of interest.

*Address correspondence to: Brie Tripp (btripp@ucdavis.edu); Rob Furrow (refurrow@ucdavis.edu).

© 2024 B. Tripp et al. CBE—Life Sciences Education © 2024 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

"ASCB®" and "The American Society for Cell Biology®" are registered trademarks of The American Society for Cell Biology.

providing extensive feedback. To explore some benefits of alternative grading practices without undertaking large-scale course revisions, we piloted a simple alternative grading practice through the use of quiz retakes integrated into class sessions. These provided students with a second opportunity to exhibit their skills and knowledge associated with learning outcomes, and took place during a part of students' final exam period. Although others have implemented similar retake practices (e.g., Leeming, 2002; Walck-Shannon *et al.*, 2019; Campbell *et al.*, 2020; Harsy, 2020; Mangum, 2020; Harsy *et al.*, 2021; Supriya *et al.*, 2024), few have systematically measured their impact on undergraduate students' perception of motivation and well-being,¹ and even fewer have applied the rich motivational framework of self-determination theory (SDT). This study utilizes SDT to interpret how students view the impact of traditional and alternative grading practices on their academic motivation and success. In the following section, we discuss some ways in which alternative grading methods have been employed, several issues with grades and traditional grading practices, and how we anchored this study in SDT.

Grading in College Courses

Course grading systems, that is, the way that assignment grades are combined to create an overall course grade, fundamentally shape the student experience in a course. Key factors in a course grading system include the level of transparency in learning outcomes, the degree to which students receive actionable feedback, and the quantity and types of opportunities a student has to demonstrate proficiency with each course learning outcome (Sorensen-Unruh, 2024). Table 1 summarizes some features of several commonly referenced grading systems. In this article, we will refer to "traditional grading practices" as systems in which students tend to have a small number of higher-stakes assignments, receive minimal actionable feedback, and have limited transparency both about how grades are assigned and how student learning is assessed. In contrast, we define "alternative grading practices" as having one or more of the following features: a grade primarily composed of many low-stakes assignments, transparent learning outcomes, transparency in the way the final grade is created, and opportunities for feedback and revision or retakes. The degree to which these alternative grading practices are implemented can play a large role in students' course grades and well-being, and may mitigate inequities arising from traditional grading practices.

Inequity Arising from Traditional Grading Practices

Educational inequities between students from dominant and nondominant cultures are replete throughout history and remain prevalent, particularly in STEM education (Whitcomb and Singh, 2021; Feldman, 2023). Traditional grading is typically rooted in male, Eurocentric conceptions of knowledge and competency dating back to the 18th century (Randall *et al.*, 2023). One of the first documentation of grades arose from Cambridge University as a means to rank students in a competitive manner (Schneider and Hutt, 2014). Although this

system was designed for male students from European descent, it has become widespread throughout education with little validation or consideration of its accuracy for women or students of color (Randall *et al.*, 2023). Grading is also vulnerable to instructors' implicit biases, which may cause them to unconsciously reward certain students by giving them "good grades" on assessments, or more detrimentally, unconsciously punish students by assigning poor scores based on factors like skin tone or immigration status (Bonefeld and Dickhäuser, 2018; Feldman, 2023).

Course performance is affected by a large variety of external factors that are unrelated to a student's knowledge, competencies, and learning and that are largely outside students' control. A student's final grade can be influenced by classroom climate (Moos and Moos, 1978; Kwong and Davis, 2015; Zysberg and Schwabsky, 2021), instructor gender (Harris *et al.*, 2020), instructor mindset (Dweck, 2008; Canning *et al.*, 2019; Muenks *et al.*, 2020), and the specific type of summative assessments utilized by faculty (such as multiple choice exams, take home essay exams, final projects, etc.; Cotner and Ballen, 2017). Student grades and course experience are also affected by the point value assigned to each course assessment (Cotner and Ballen, 2017), the way in which the points are combined and potentially curved to yield the final grade (Bowen and Cooper, 2021; Feldman, 2023) and the anxiety induced during tests or other evaluations (Seipp, 1991; Cassady and Johnson 2002). Furthermore, course grades can have profound impacts on students, affecting their self-esteem, confidence, ability to persist within a desired major, and their interest in specific majors (Seymour and Hewitt, 1997; Seymour and Hunter, 2019). These factors interact with different student identities and backgrounds in ways that may benefit or harm the performance of particular students.

Grading systems in which there are a small number of high-stakes assessments can be particularly inequitable. This practice can cause test anxiety that may most strongly impact systematically excluded students (Chapell *et al.*, 2005; Hannon, 2012; Núñez-Peña *et al.*, 2016; Ballen *et al.*, 2017). Such high-stakes assessments are also detrimental to both student welfare and educational attainment (Ecclestone, 1999; Boud and Falchikov, 2006; Jones *et al.*, 2020; Biggs *et al.*, 2022), create inequitable standardized assessment practices (de Brey *et al.*, 2019; Gilman, 2019; Nastal, 2019; Ngo and Melguizo, 2020; Randall *et al.*, 2023), and cause negative affective consequences for students (Ribeiro *et al.*, 2018; Pascoe *et al.*, 2020). According to Au (2008), the adoption of high-stakes assessments has not led to progress but rather perpetuated disparities rooted in race and class.

These equity issues with grades amalgamate into a mountain of barriers for students who hold underserved identities or come from systematically excluded backgrounds. As instructors, we are charged with challenging these deeply ingrained measures of ranking students which reinforce inequities across race, class, gender, and many other social identities and backgrounds.

SDT

Although assigning student grades may feel like only a small part of an instructor's experience, how students receive

¹We define well-being as experiencing minimal distress, encompassing both mental and emotional wellness (American Psychological Association, 2018).

TABLE 1. Key features of several commonly referenced grading systems

Grading system	Learning outcomes transparent?	Additional opportunities to demonstrate proficiency after feedback?	Letter grade components and stakes	Key citation(s)
Traditional	Not necessarily	Not necessarily	Numerical combination of scores on assorted assignments, potentially with some type of grading curve. Often only a few high-stakes assignments	
Standards-based Grading	Yes, key skills/concepts defined as “standards”	Yes, feedback followed by additional opportunities to demonstrate proficiency with a standard	Letter grades not required, but can be created from a combination of proficiency ratings across standards. Typically many low-stakes assignments	Feldman, 2023 ; Vatterot, 2015
Contract grading	Not necessarily, grade defined by completing suite of assignments	Yes, feedback followed by opportunities to revise work until it meets agreed-upon standards	Grade earned based on completing an agreed upon contract between instructor and student. Typically many low-stakes assignments	Shor, 1996
Mastery grading	Yes, key skills clearly defined (individually or in modular units)	Yes, assessments repeated until proficiency is achieved	Letter grades not required, but can be created from a combination of proficiency ratings across units/skills. Typically many low-stakes assignments	Bloom, 1968
Specifications grading	Yes, clear expectations defined for key outcomes or assignments	Not necessarily, though many specific examples of specifications grading schemes incorporate revisions or additional opportunities	Combination of achieving key outcomes (completing different assignment types to a certain standard). Typically many low-stakes assignments	Nilson, 2015

feedback and grades affects their long-term motivation, well-being, future career success, and lifelong approaches to learning and critical thinking. Specific grading practices play a key role, potentially nurturing student creativity and motivating students to engage and pursue new challenges in the classroom and in their future career. Here we outline how SDT can inform instructors' grading practices to promote student motivation in the classroom.

SDT is a theoretical framework to describe the drivers of motivation ([Deci and Ryan, 1985](#); [Ryan and Deci, 2000a, 2000b](#)). When people are intrinsically motivated, they will do something simply because it is interesting to them. SDT posits that intrinsic motivation is always autonomously motivated, that is, under one's control and driven by values and interests within oneself. On the other hand, an extrinsically motivated action is taken in order to achieve a particular outcome ([Deci and Ryan, 1985](#); [Ryan and Deci, 2000b, 2000a](#)). Those outcomes are generally externally defined by others. For example, earning an “A” grade is an outcome that can externally motivate students to complete classwork, while an interesting or meaningful assignment may intrinsically motivate some students.

A sense of autonomous motivation is associated with an array of positive psychological outcomes ([Ryan et al., 1995](#); [Sheldon et al., 1997](#); [Nix et al., 1999](#); [Howard et al., 2021](#)). For most assignments in higher education courses, students will not always be intrinsically motivated, but there are many external motivators built into college course structure. For example, failing a course will delay academic progress, while

earning a high grade may have an effect on future job or scholarship opportunities. If instructors can design course materials, course structure, and the classroom environment to go beyond external motivators to promote autonomous motivation and students' sense of control over their course success, this may positively affect student motivation and well-being, and ultimately improve performance.

SDT characterizes several affective factors that affect the degree to which a behavior may be externally or autonomously motivated: perceived **competence**, level of **autonomy**, and a sense of **relatedness** to others. These factors are often labeled ‘basic psychological needs’ ([Ryan and Deci, 2000b](#)). Perceived competence refers to a person's sense of their expertise in a task ([Rodgers et al., 2014](#)). But even high levels of perceived competence may not inspire action unless an individual feels a sense of autonomy. In this context, autonomy refers to an internal perceived locus of causality, in which an individual's action is originating from within them rather than an individual being forced into this action by external conditions ([Heider, 1958](#); [deCharms, 1976](#)). Because many behaviors happen within a social context, an individual's sense of relatedness, the degree to which an individual feels connected to others, will also influence motivation ([Ryan and Deci, 2000a](#)). Feelings of competence, autonomy, and relatedness interact to support autonomous motivation.

Course grading practices are likely to influence student performance and motivation. Traditional grading practices with high-stakes assessments and no room for revision can reduce students' autonomous motivation ([Deci et al., 2001](#);

Chamberlin *et al.*, 2023). For example, a low midterm score could make it nearly impossible for a student to earn a high final course grade; this may be demotivating and generate a sense that the student no longer has control over their course success. In contrast, grading systems that include a large number of low-weighted quizzes or homework assignments reduce the pressure of mandatory high performance on any single assignment. A course having many low-stakes assessments benefits student learning (Leeming, 2002; Orr and Foster, 2013; Pennebaker *et al.*, 2013; Casselman and Atwood, 2017; Sullivan *et al.*, 2018; Walck-Shannon *et al.*, 2019; Crucho *et al.*, 2020), and this frequent feedback process is likely to promote student autonomy and ultimately motivation (Ryan and Deci, 2000b). More generally, when students perceive that their instructors support autonomy in learning, students experience greater motivation and course engagement, as well as increased feelings of competence, autonomy, and relatedness (e.g., Williams and Deci, 1996; reviewed by Reeve and Cheon, 2021; Guay, 2022).

Grading systems can also promote motivation by providing opportunities to retake assessments. A chance to retake a quiz or exam both reduces the assessment's stakes and provides students with greater control over their learning. Retakes could be based on retaking an identical exam or resubmitting an identical assignment (e.g., Perlmutter *et al.*, 2022; Vyas *et al.*, 2023), or via the completion of novel alternative assignments that assess the same course learning objectives (Campbell *et al.*, 2020; Harsy, 2020; Mangum, 2020; Harsy *et al.*, 2021). Students experiencing multiple chances to complete alternative assignments identify benefits including gains in self-efficacy and self-concept, and reductions in test anxiety (Harsy, 2020; Harsy *et al.*, 2021; Vyas *et al.*, 2023). Retakes may be optional and require students to find time to complete the assignment outside of class. In such cases, students completing the retakes may benefit in their course performance (Cates, 1982; Davidson *et al.*, 1984; Friedman, 1987; Juhler *et al.*, 1998; Walck-Shannon *et al.*, 2019; Supriya *et al.*, 2024) and experience reduced test anxiety (Supriya *et al.*, 2024). However, systems that require optional, outside-of-class work risk lower participation rates by students with lower initial performance (Walck-Shannon *et al.*, 2019) and can perpetuate inequities by preferentially advantaging students with more flexible schedules (Supriya *et al.*, 2024). For this reason, we favor the use of retake opportunities that are integrated into course meeting times and opted to take this approach for this study. We hypothesize that such grading practices promote student motivation via increased student competence, autonomy, and relatedness while avoiding biases in task completion that perpetuate inequities. In this study, we anchor our design, analysis, and interpretation of data in SDT to better understand how in-class retakes impact undergraduate student motivation and well-being.

Problem Statement and Research Questions

Although there have been other studies that elucidate the impact of assessment retakes on students' performance and affect (e.g., Cates, 1982; Davidson *et al.*, 1984; Friedman, 1987; Juhler *et al.*, 1998; Walck-Shannon *et al.*, 2019; Campbell *et al.*, 2020; Harsy, 2020; Harsy *et al.*, 2021; Mangum, 2020; Perlmutter *et al.*, 2022; Vyas *et al.*, 2023; Supriya *et al.*, 2024),

there has yet to be a study that centers the voices of undergraduate students to explore their perceptions of grades and these grading practices through the motivational lens of SDT. Here we describe the implementation of a small, scalable change in course grading practices: the use of in-class quiz retakes that are integrated into class meeting times. These quizzes provided students with an additional opportunity to demonstrate proficiency with course learning outcomes and were integrated into a traditional STEM grading scheme of multiple quizzes and a final exam. We assessed the impact of these quizzes on undergraduate students' course experience and motivation using in-depth student interviews, and quantitatively analyzed student quiz performance. The study was guided by three primary questions:

1. How do grading practices and quiz retakes influence undergraduate students' self-determination in learning?
2. What are undergraduate student perceptions of the purpose of grades?
3. How does initial quiz performance affect student retake participation and improvement?

MATERIALS AND METHODS

This project was approved by the University of California, Davis Internal Review Board under exempt status (1999540-1).

Authors' Positionalities

The authors of this study represent a diverse array of backgrounds, encompassing both shared and distinct identities, and occupy various stages within their academic trajectories. A.R. (able-bodied woman) and E.P. (able-bodied man) are currently undergraduate researchers of color who have experienced various grading practices throughout their academic journey. Their unique perspectives enriched this study in immeasurable ways as undergraduates are best positioned to fully comprehend their peers' reflections. As many participants in the study held marginalized racial and ethnic identities, A.R. and E.P. were critical in unveiling meaning with which the leads of this project, B.T. and R.E.F., had no lived experience. B.T. (able-bodied woman) and R.E.F. (able-bodied man) are White professors who currently teach a variety of science courses in higher education and have varying experience with alternative grading. This diversity among the research team enriched the study with multiple perspectives, while also necessitating a critical awareness of the power dynamics and hierarchies that influenced our collaborative processes and interpretations.

Implementation of In-Class Quiz Retakes

Authors R.E.F. and B.T. similarly implemented in-class quiz retakes within their instructional framework in three different offerings of a statistics for biologists course for R.E.F. and two offerings of one neurobiology course for B.T. Instead of midterms that cover many weeks of material, we embedded several quizzes that were evenly distributed throughout the 10-wk term (R.E.F. implemented 6–7 depending on the year; B.T. implemented 5). These quizzes were delivered at the beginning of the class with a 15-min duration (B.T.) or online

TABLE 2. Course grading schemes for the focal neurobiology and statistics courses.

Course	Assessments	Percentage of course grade
Neurobiology	In-class (retake) quizzes (x5)	30%
	Course assignments (x11)	40%
	Attendance (13 out of 16 for full points)	5%
	Final exam (cumulative)	25%
Statistics	Online (retake) quizzes (x7)	28%
	Small course assignments (x7)	34%
	Projects (x2)	28%
	Final exam (cumulative)	10%

Note: the grading scheme for statistics reflects the second and third iteration of the course; the first year only had six quiz retakes.

with a 60-min duration (R.E.F.). In-class quizzes were followed by the continuation of course content. This approach enabled students to stay current with the material while keeping the relative weight, or “stakes,” of each assessment low in comparison to a structure with only one or two midterms. We chose this approach due to its simplicity and ease of implementation, rather than fully overhauling the course grading practices. The cumulative point value of these quizzes was similar in both courses (28% of grade in statistics, 30% of grade in neurobiology). Refer to [Table 2](#) for further details about course grading schemes.

The instructors integrated retakes into scheduled course time through a two-part final exam; one part contained a cumulative final exam (mandatory) and the second part an opportunity to take new versions of earlier quizzes (optional). The instructors framed these quiz retakes to their students as “second-chance quizzes.” *Importantly, the retake quizzes taken during the final exam assessed the same learning outcomes as the initial versions completed during the term, albeit with novel questions.* Students had two hours to complete both parts of the final in no particular order. R.E.F. allowed students to complete retakes of any and all of the 6–7 course quizzes. B.T. provided students with the opportunity to choose two of the five course quizzes to retake. A student’s highest score for each quiz (initial attempt or retake) was the final score they received for that particular assessment in the course. This encouraged students to revisit material they struggled with during the course while providing them an opportunity to improve their grade.

Interview Recruitment

To explore student perceptions of grading practices, we recruited students from one offering of both courses through a survey designed in Qualtrics. A.R. sent an email to students inquiring about their interest in participating in an interview study regarding the gradings practices they experienced in college. If students indicated “yes,” they were prompted to provide their contact information (name and email address). A.R. contacted potential participants via email and invited them to sign-up for an ~30-min interview time slot. During the interview, participants verbally consented to their responses being used in this study. Upon completion of the interview, students

were sent a \$25 Amazon gift card as a recruitment incentive and token of appreciation.

Interview Protocol Development

R.E.F. and B.T. designed interview questions to probe students’ experiences of grading and assessment practices in college as well as their reflections regarding the second-chance quizzes (see Supplemental Material A for interview protocol). A.R. conducted six practice interviews with STEM undergraduates of color and white students unrelated to this study to obtain face validity ([American Educational Research Association, 2014](#)) on the questions and clarify wording that may cause confusion among participants.

The interviews followed a semistructured format as outlined by [Longhurst \(2003\)](#). A.R. posed a set of uniform questions to all participants, fostering a dialogue between A.R. and the interviewee. This approach provided space for follow-up questions to elucidate participants’ responses or delve into other topics related to the structured interview framework. The interviews took place on Zoom and were recorded in audio format. Subsequently, the audio files were forwarded to Rev.com, a third-party company, for transcription.

Qualitative Data Analysis

All authors of this study engaged in inductive and deductive content analysis ([Patton, 1990](#)). Inductive content analysis uses inductive reasoning to uncover themes from raw data through repeated examination and comparison. Deductive content analysis is applied when the analytical framework is defined in accordance with existing knowledge or theories. A.R. chose 10% of the interviews that had the most diverse responses in order to capture the array of grading experiences. In the initial inductive coding cycle, we independently read and annotated main ideas, or “codes,” that arose based on a preliminary Research Question 1 (influence of grading practices and quiz retakes on student motivation) and Research Question 2 (student perceptions of purpose of grades). We reconvened to compare findings and collaboratively generate a list of preliminary “codes” that captured the essence of participants’ reflections ([Saldaña, 2016](#)). During this process, we discovered themes in our data related to SDT. Thus, we deductively restructured our codes to align with the tenets of SDT (competency, autonomy, and relatedness) and refined Research Question 1 to be “How do grading practices and quiz retakes influence undergraduate students’ self-determination in learning?” Codes that fell outside this framework were categorized based on similarity. Themes were assigned to sets of all similar codes, representing key aspects of how participants experienced and perceived college grading practices. This process resulted in a “codebook” containing emergent themes and associated codes from the initial transcripts.

During the second coding cycle, we used the initial codebook to categorize another 10% of randomly selected interviews. This iterative process continued for two cycles until data saturation was achieved, with no additional codes or themes emerging. Codes with limited traction were either combined or removed from the codebook. The final codebook, along with interview transcripts, was uploaded to MAXQDA 2022 ([VERBI Software, 2021](#)), a qualitative data analysis

software. A.R. and E.P. then recoded all interviews to consensus using the final codebook.

Quantitative Data Analysis

To quantitatively assess the relationship between students' initial quiz scores and their participation in and improvement from in-class quiz retakes (Research Question 3), we analyzed student completion of these quizzes. We assessed student grades and participation in both courses across multiple years, spanning a total of five course offerings ($n = 213$ total students, with 135 from statistics and 78 from neurobiology). This sample is larger than the recruitment pool for interviews, because it includes data from years before and after interview recruitment. Using the programming language R (R Core Team, 2023), we calculated students' initial quiz score grades, the fraction of students who took at least one retake, and the average quiz grade improvement for students who did retake at least one quiz. We could not assess improvement on individual quizzes because of challenges in identifying whether or not students made a substantive attempt at retaking a quiz. During the final exam period, some students started multiple quiz retakes (answering one or two questions), but only completed all the questions for one quiz. Because of this ambiguity in effort, we looked for overall improvement between the initial quiz average (across all initial quizzes) and the finalized quiz average (after incorporating any improvements in any retakes).

We analyzed average improvement, whether the initial quiz average predicted student participation in retakes, and whether it predicted the degree of improvement in a student's grade. To assess the statistical significance of quiz score improvements, we used a Wilcoxon signed rank test, and used a generalized linear model with a logit link function to assess whether initial quiz average significantly predicted a student's binary decision to retake at least one quiz or not, while controlling for course and year. To analyze the relationship between initial quiz average and grade improvement (from retakes), we fit a linear model that included course and year as covariates and used weighted least squares to minimize the impact of heteroskedasticity (Strutz, 2016).

RESULTS

In this section, we share results gathered from interviews conducted with undergraduate students in which we explored their perceptions and experiences of grading practices throughout college as well as their perceptions of quiz retakes in two different courses: statistics and neurobiology. Additionally, we present findings based on students' perceptions of the purpose of grades. The outcomes are structured into subsections based on details about our participant population and research questions. Alongside the qualitative results, we also present the quantitative impact of initial quiz performance on student participation in quiz retakes and on the degree to which students' grades improved.

Participant Population

Of the pool of 112 students initially contacted, 24 students indicated interest in being part of this study (21% response rate). Six students were unable to take part in the inter-

view due to other commitments. Ultimately, 18 students completed the interviews: 13 from the statistics course and 5 from the neurobiology course. Although students in the former were more represented in our sample, every theme was expressed by multiple students from each course. In the recruitment survey, students provided their demographics including race/ethnicity, LGBTQIA+ membership, gender, and first-generation college-going status (see Supplemental Material B for details).

RQ1. How do grading practices and quiz retakes influence students' self-determination in learning?

Interview Results

Competency. Over three-quarters of student participants (78%, $n = 14$) reliably made connections between alternative grading practices and perceived competence (Table 3). According to SDT, perceived competence refers to a person's sense of proficiency or the opportunity to become more proficient in a task (Ryan and Deci, 2000a; Rodgers *et al.*, 2014). Students often reflected on how the quiz retakes provided an opportunity to improve their learning, leading to a decentering of grades, an increase in motivation, and better retention of course concepts.

Adriana reflected on her ability to revisit and become more proficient on certain concepts which decentered grades and promoted learning, while Isabella expressed how improving her learning led to a sense of motivation:

"Because I scored low on quizzes probably meant I wasn't understanding something very clearly... Because I knew there was an opportunity to save my grade in case something went really wrong, I felt like I was focused a lot more on understanding the material rather than trying to get a grade. And honestly even when I got questions wrong on the quiz, instead of being all stressed out that I had done terrible on it, it was more like a minute of clarity. I looked at it as being helpful to get questions wrong because now I could figure out why rather than focusing on, 'Oh my gosh, I'm not gonna do well in this class.'" -Adriana

"[The quiz retakes] definitely directed my attention to quizzes that I had done more poorly on. Because I was like, 'oh, I'm gonna be retaking that. So not only did I get a lower grade, my understanding probably isn't as good, but I'm also more motivated 'cause if I study this more in depth, not only do I understand it better for the final, but I'll be able to do better for my second-chances quiz.'" -Isabella

Many students, including Mira and Derek, reflected on the ways in which revisiting certain concepts led to retention of information both within and outside the classroom:

"[The quiz retakes] really did help and make a difference because I was able to solidify my learning towards the end. It really did give me another opportunity to dive back into the material again and relearn it the right way. I think it really does solidify your learning. To this day I'm still explaining pathways to people and I remember so much from that class compared with other classes that I've taken." -Mira

TABLE 3. Example quotes from students describing the three themes aligning with SDT: Competency, autonomy, and relatedness

Theme % (n) description	Example quotes
Competency 78% (14) <i>Describes students growing in their proficiency to perform well on a task or improve their learning when provided the opportunity to retake assessments</i>	"I think it's a really good learning practice like, 'Okay, here's what I have realized that I did not learn as well based on this quiz, and now I have a chance to go and potentially learn it better than I did before.' And to have my grade reflect that I have learned this stuff later because sometimes not everyone learns at the same pace." —Clara "[With the quiz retakes] I was able to improve past grades, especially if I realized I understood the concepts better than I did at the beginning, and I was able to show that I had a better understanding. The makeup quizzes that were specifically focused on certain concepts, it was nice to go back and relearn everything that I didn't understand before." —Anne "I think that [the quiz retakes] gave more chances for the students to show that they did understand the material. Being more relaxed in the classroom and feeling like you are not going to just instantly fail because you make a singular mistake benefitted my learning and I have retained more of that information." —Derek
Autonomy 72% (13) <i>Describes the increase in student agency when lower stakes assessments are implemented into courses</i>	"The whole idea of making us retake two quizzes in the final I thought was a really great approach because there was no way we were gonna have time to make up every single quiz. And it made you prioritize what you needed to work on. Also in studying a little bit extra for the quiz you did worse on, it was helping you study for the final because clearly that's where I was struggling in the class anyways." —Adriana "I was about to take the final 'cause then I was like, okay, I'm going to retake these two quizzes. So I need to focus more on that material when I am studying. And that worked out for me, so during the quarter, it didn't really make much of an impact because I was trying to ignore it and just focus on like doing well in the moment, but then once I knew like, okay, these are the ones that I have to retake, then when I was studying for the final, I catered more towards that material." —Amara "So I really like that grading practice because I already was aware before going into the final that I could go over what I did wrong and I could go to office hours so I could review the quizzes that were lower before going into the final. The specific likes about the system is I can prepare ahead of time and I can do group work if needed. I can go over it in more advance." —Camila
Relatedness 72% (13) <i>Describes the degree to which students feel connected to their peers and instructors</i>	"I just think there was a lot of support, and the student environment was different. I mean, everybody was motivated, it was very rigorous, but again there were lots of lecture reviews, office hours, and support for students. So despite kind of having the grading schema and being incredibly difficult topics, it offered that support to where it wasn't as maybe challenging as taking other assessments would be" —Mira "Yeah, I think when they, when they do more frequent assignments and ones that are, you know, if there's like weekly homework or weekly online quizzes, I'm more confident that the professor understands what it's like to be a student and how learning actually works." —Clara "Yeah, the starkest difference for me would be when I was taking Spanish classes versus STEM, because those are like two completely different worlds. And to see just how calm and gentle Spanish professors are, they're so easy-going. And if anything ever came up, you say, 'hey, I'm going through something right now, can you give me an extension?' or 'I accidentally clicked on the wrong thing. Can you change my grade for that one question?' —Amara

"I think that [the quiz retakes] gave more chances for the students to show that they did understand the material. Being more relaxed in the classroom and feeling like you are not going to just instantly fail because you make a singular mistake benefitted my learning and I have retained more of that information." Derek

Students like Bianca and Mara also exhibited metacognitive thinking about how retakes refocused their effort on material with which they struggled:

"The second-chance quizzes, knowing that there were opportunities to improve scores on things made it a better environment to actually retain the information... The opportunity to go over the information again. Maybe I didn't

understand it the first time around, but studying off of that first quiz, like what did I get wrong, what was I not understanding? And then being able to go back and take the quiz again and see like, okay, now I'm getting it." —Bianca

"What I would do is just go back to the quizzes and review the material. So, retaking the quizzes just allowed me to look through that material once again and really look into why I missed certain questions and focusing my studying on those questions in particular." —Maria

Autonomy. Almost three-quarters of students in our sample reflected on aspects of autonomy when provided opportunities to engage in lower stakes assessments ($n = 13$, 72%; Table 3). SDT describes autonomy as relating to an internal sense of

self-determination, where an individual's actions originate from within themselves rather than being compelled by external circumstances (deCharms, 1976; Ryan and Deci, 2000a). In this study, students' autonomy was expressed through their ability to make decisions on how to navigate a course, such as prioritizing material, choosing where to focus studying efforts, and feeling a sense of control. Consider Mira's reflection about allocating study efforts:

"I really enjoyed [being able to retake quizzes]. And especially 'cause I was taking anatomy last quarter and it took up the majority of my study time. Except for kinda towards the end of class, I left the [neurobiology of mental illness] class a little bit to kinda prioritize the anatomy class, but I took those second-chance quizzes in the neuro class again during the final and it shot my grade up like 10%." -Mira

Many students expressed how quiz retakes provided control to better manage their time, such as Isabella and Bao:

"The second-chance quizzes allowed me to arrange my studying for that class around what I was busy with, like other classes or extracurriculars. So because I had that backup in my mind where like, oh, I can do the second-chance, if I had like a bunch of midterms for another class or had a deadline due for an extracurricular, I can be like, 'okay, I can put this off a little bit and then catch back up toward the final. So, it allowed me to reorganize my time throughout the quarter and helped to direct my studying towards that material I didn't understand well.'" -Isabella

"One thing about grades is they shouldn't be a fixed structure. They represent my performance. And sometimes my performance changes and it's unpredictable and I cannot control that. So the fact that I have a chance to redo and retry, which allows me to prove that my performance is still good and it's not really dependent on what happened just that one time." -Bao

Some students, like Elena, described having power to change their grade regarding the quiz retakes, thus leading to a sense of relaxation and a decrease in stress, while Morgan expressed having options to improve their grade which were motivating:

"As for the grades on the makeup quizzes, it's just relaxing. I knew that I had the power to change what I did and what I got wrong. Nothing was set in stone, so I wasn't freaking out about the class. I was mentally in a better spot about it." -Elena

"[The quiz retakes] made it more of a positive impact because once you get a bad grade it's hard to bring your grade back up to an A or B or something. But I feel like doing this grading system provides a better playing field and options to raise your grade. So it is easier and motivating." -Morgan

A few students reflected on the ways in which low-stakes assessments allow time for them to apply the material to real-life experiences:

"If it's more of a chill exam or assessment that covers less material, I feel like I can learn the material on my own time

and also learn about it in ways that apply to me. I can think about applying it to my life and the experiences. And I can cater what I take from that class, like what I learned from that class to both like what I want to do and what I currently do right now." -Clara

Relatedness. A majority of students also identified ways in which grading practices affected feelings of relatedness ($n = 13$, 72%, Table 3). SDT describes relatedness as a feeling of connection to others (Ryan and Deci, 2000a); in a classroom this could describe connection with peers and/or instructors. Students discussed relatedness mostly in terms of their comfort and connection to their course instructors, explaining both positive impacts from alternative grading approaches and negative feelings arising from courses with traditional grading practices. The positive impacts of greater instructor relatedness connected to both learning and student well-being more generally. For instance, Morgan and Clara reflect on the caring nature of the instructor in relation to quiz retakes:

"[This] class was very refreshing 'cause [the] grading system really made it feel like [the instructor] cared about the students. It's like [they] really just wanted us to learn and retain the information and not so much focus on getting the good grade." -Morgan

"With my grade and that flexibility to fix something for a week where I just wasn't doing well, it was really, really nice and made me feel like [the instructor] really cared about us doing well and cared about our well-being and wanted to do what he could to let us get a good grade in that class." -Clara

On the contrary, many students also reflected upon experiences in courses with rigid grading structures that allowed little to no room for accommodation, expressing how this could harm their relationship with the instructor:

"And then when you have STEM professors, and they're like, 'No, I'm not going to be flexible for you. That's not fair towards everyone else.' And you get responses that are kind of robotic and they don't really care about what you're going through" -Amara

"Whereas if it were exam based and all the memorization, there's not really anywhere you can get help. And I found that in those classes it was harder for me to reach out to resources such as the TA instructor, tutoring to go and get help because I felt ashamed that I didn't understand that topic yet." -Elena

"Or are they going to have a panic attack on the exam? Which I've seen, I've seen that happen in Organic Chemistry, Lord knows how many times and I was on Zoom, I would see someone crying and they couldn't turn off their camera and I was, 'oh my god, it's horrible.'" -Mia

Only a single student mentioned positive peer-peer interactions. Clara discussed the way that particular low-stakes course assignments helped create a sense of community with peers.

TABLE 4. Salient themes, codes, and example quotes by students who expressed traditional grading practices as harmful to their well-being and learning

Theme % (n)	Codes % (n)	Example quotes from students
Grades and traditional grading practices are harmful for students' well-being and learning 89% (16)	Traditional grading practices, such as high-stakes assessments and curving, cause stress and anxiety 78% (14)	"I've had classes where your grade is pretty direct, so you know your midterm and your final and that's it. And that's a lot of pressure. And I know that's kind of a quarter system thing, but it really does, especially for students that have testing anxiety that struggle with testing and it can be a huge stress factor over a quarter. It's kind of brutal when you have just three exams and you either get an A in all of them or you don't get an A in the class, you know?" —Mira "Even though it is more stressful to have a final exam that is 50% of your grade versus in those other classes it would be more broken up." —Isabella "Curves were not my favorite, just because a lot of the time it was subjective, so the professors would say, 'The curve can be applied at the end of the quarter only if the students didn't do well.' So there was always a, 'What if other students do well, though.'" So it was always a struggle to have to wait until the final to see the curve" —Maria
	Exam content coverage inhibits student learning 55% (10)	I struggle with really big tests that have a lot of cumulative information without quizzes to prior backup that knowledge. It is too much information. —Alice If I take a midterm that covers five weeks of stuff, it's going to be really hard for me to know all of that material and be able to fit it all into my brain when I do take that midterm. —Clara
	Grades are inequitable 28% (n = 5)	I actually struggle with test taking, especially because exams in the sciences require a lot of memorization. —Camila It would be good to start changing the grading system. I think it's, very, especially, in a quarter system where everything goes so fast, being labeled with a grade it can be daunting for students, and I think there should be more studies and research on how to still obviously grade your work. But in a less, "oh, you're an A student or you're an F student," you know, that type of deal. —Adithi
		I have issues with the current grading system. Especially when some courses or professors maybe aren't super flexible for students and their needs, then the grades don't really reflect that student's character. And so as a result, there can be a lot or maybe even the teacher just doesn't really like to hold office hours or like to provide the necessary support for their students. That's also a really big issue that can affect a student who has no background in STEM whatsoever versus another student who's older sister or brother went to [that university] and took the exact same course. Or someone has a private tutor. —Mia

"I feel like discussion posts are nice because they create a sense of community with your class and when they're not graded overly harshly, it's nice to hear opinions from other people in your class. I think there's a lot of learning that can be done. Learning from your peers and from perspectives you don't know." — Clara

Grades and Traditional Grading Practices Harm Motivation.

The vast majority of students (89%, $n = 16$) expressed dissatisfaction with traditional gradings practices such as high-stakes assessments and opaque grade distributions due to curving. Over three-fourths of students (78%, $n = 14$) expressed how such practices were detrimental to their overall well-being and learning, causing stress and anxiety. Over half of the students (55%, $n = 10$) also identified excessive content

coverage on exams as a factor contributing to memorization and regurgitation and led to some students feeling underprepared for the workforce (Table 4). Last, 28% ($n = 5$) described grades and traditional grading practices as inequitable to certain groups of students.

Many students like Alice related high-stakes assessments to lack of motivation and procrastination which contributed to stress, while other students like Olivia spoke to heavy exam weighting as a source of stress:

"Classes with just three tests and nothing else due makes me unmotivated because of the amount of stress. And so it was like I should have been studying, but I was putting it off, because of how stressed it was making me." —Alice

"I think classes where the grade is 100% based on tests are a lot more stressful, especially when there's only two midterms is the most stressful, three midterms is a little less stressful, but when the grade is based completely off of just three tests that's a little stressful." -Olivia

Some students like Adriana reflected on the opacity of grades they experienced in some of their courses as a source of stress:

"I struggled in classes where your grade is not clear to you. The class I currently have right now, the professor is keeping all of the grades locked on campus until the end of the course and it's just very stressful because seeing the grade wouldn't make me try any less, it would just give me a sense of peace." -Adriana

Students also described how curving reduces performance transparency, causing demotivation and damaging students' self-worth and competence, despite their experiences that curving had generally raised their grades:

"I think curving adds a level of stress because in classes where I do have a curve, I usually feel kind of like a failure the whole time 'cause I'm doing so bad. I've had a class where I had a D, and then all of a sudden the curve was so big it became a B, and that makes it discouraging." -Morgan

"For curves, those actually make me feel so much less competent. If I don't know the grade I have because of a potential curve, to open Canvas and to see an F in your grade is not a great feeling. Even if, by the end of the course it ends up turning into an A plus. Sometimes the curve was so large that I never knew what grade I had in that class, which was very stressful." -Mia

In addition to the above reflections on how traditional grading practices hamper students' well-being and learning, students discussed being assessed on extensive amounts of content and details, as detrimental to their learning and future careers. Elena described how exams that cover too much material left them feeling unprepared for the workplace:

"Exams that make up the majority of your grade, I personally have to cram for those tests and after I'm done with it, half of that information is just gone. It's just memorization because you are being tested on such a large amount of material. I haven't heard of any workplace where you just take tests and just memorize and regurgitate information onto a piece of paper. I, honestly, personally feel very unprepared for my field." -Elena

Students such as Christine and Anne expressed how details and facts required on exams decreased deep thinking and are designed for certain types of students who consistently do well on exams:

"A lot of exams in college are just mostly regurgitation of facts and not really interacting with the material on a deeper level and there's just so much of those tiny facts.

And I feel like a lot of people have been able to get by without actually knowing the material on a deeper level... but only the ones that are good test takers are able to do that." -Anne

"I would say cumulative finals stress me out where all of the questions are based on minor facts that were only covered once in lecture. That's where I would say I struggle 'cause I can understand how everything integrates as a whole, big picture, and a few relevant minor facts. But it's definitely hard to memorize all of them." -Christine

Along with the above expressions, we invite readers to consider Maria and Derek's critical awareness of the inequitable nature of grades in academia:

"Based on what I've learned in my education, it's just very subjective. Grades are just there to measure people based on the white education model. It's unfair. A lot of the time they don't [consider] people's different backgrounds. And people learn differently and sometimes these courses that offer exams as the only way to grade people, it's biased toward white students." -Maria

"[The purpose of grades] is to favor certain people, and there's a reason why a lot of other students that make it in academia have money and time. They don't have to work a job during college and they can spend all this time on studying. And a lot of the students that do have to work a job or do have to do something on top of learning, you see that they don't do as well in grades because even though they're trying their best, grades are representative of your privilege." -Derek

A few students, such as Clara, highlighted the dehumanizing effect of grades:

"Grades really make learning inaccessible and can be really discouraging because they focus on everything you get wrong. And that is so incredibly unhelpful because you're going to make mistakes. That's just how learning works. And then recovering and learning from those mistakes is how you learn things. So, when courses are graded with grades, it completely devalues the humanity of learning and just assigns these numbers to sometimes very arbitrary measures of how much of the textbook we fit inside our brain in a matter of time." -Clara

RQ2: What are student perceptions of the purpose of grades?

Students' Perspective on the Purpose of Grades as de/motivating. Sixty-six percent ($n = 12$) of students described the intent of grades as motivators or demotivators for learning. Some students described the purpose of grades as a metric that quantifies and communicates student performance, such as Camila who stated, "*I think courses have grades so that you can see how well or poorly you are doing,*" and Christine who said, "*Grades are just a gauge of how well you know the material and how well you reproduce it on a test.*" Other

TABLE 5. Student completion of and grade improvement from in-class quiz retakes

Course	% students re taking at least one quiz (n)	% retakers who improved at least one quiz score (n)	Average quiz grade improvement % (mean/median)
Statistics 2021	84% (27)	100% (27)	4.4 / 5.9
Statistics 2022	85% (34)	94% (32)	6.1 / 7.4
Statistics 2023	76% (48)	81% (39)	3.6 / 5.2
Neurobiology 2022	87% (34)	91% (31)	4.5 / 6.2
Neurobiology 2023	95% (37)	86% (32)	3.5 / 6.1

students took this one step further by applying performance learned in school to real-life goal setting, such as Mia:

“I feel like the overall hope here is that once you go out into the, “real world” and get a job and start working that you’ll be able to kind of tell how you’re doing on your work and then like what needs improvement to kind of achieve this level of whatever goal you want to set for yourself.” -Mia

Several students identified the purpose of grades as informing professional schools on how successful students might be, like Bianca, “*Grades are for things like grad school. We need to have grades, so they know that you can move on to grad school or that you might be successful in grad school.*” Isabella similarly explained how grades are important when moving beyond college and getting a job, “*We have grades so that GPAs can be used by people fresh out of college, to show their employer.*”

Only a few students identified grades as a personal “*motivation for student learning*” (Olivia). Alice described grades as demotivating, “*I feel like [the purpose of grades] are a way to motivate students but if your grade gets too low, that negatively affects your motivation.*” Mira reflected on the importance of grades as a way to “*separate a person that is motivated and does wanna learn the concept and cares about the material from those who don’t.*”

RQ3: How does initial quiz performance affect student re-take participation and improvement?

Across all offerings of both courses, a large majority of the 213 total students participated in at least one in-class quiz retake (range among courses: 76–95% of students within course; Table 5), and almost all the students who completed a retake (“retakers”) improved their score on at least one quiz (range 86–100%; Table 5; Figure 1, A and B). The median improvement to retakers’ overall quiz average was large enough to substantially influence overall course grades, and did not vary greatly by course offering (range 5.2–7.4% improvement; Table 5). Scores on the retakes were consistently higher than the initial quiz scores ($p < 0.05$ for all five course offerings), despite the fact that many students (particularly in statistics) appeared to start several of the retake quizzes but ultimately only complete one or two of the retakes.

Initial quiz average (across all quizzes within the course) was a significant predictor of whether or not students completed a quiz retake ($p < 10^{-5}$, Figure 1C); students with a lower initial average were more likely to retake. For example, students with an initial quiz average of less than 90% were

1.45 times more likely to retake at least one quiz than students with an initial quiz average of 90% or higher. We found no difference in participation by course type (statistics versus neurobiology) when controlling for initial quiz average ($p = 0.12$).

Among retakers, students with the lowest initial quiz average generally improved their grades the most ($p < 10^{-15}$, Figure 1D). We found no difference in improvement by course type (statistics vs. neurobiology) when controlling for initial quiz average ($p = 0.74$).

DISCUSSION

In this study, we implemented a form of alternative grading using in-class quiz retakes to decenter grades and improve students’ motivation to learn and measured its impact by conducting interviews with undergraduate students from two upper-division STEM courses. We anchored our findings in SDT, revealing that students had an increased sense of competence with the course material, greater autonomy and feelings of self-determination, and improved relatedness to the course instructors. Additionally, students reflected on the harmful effects of traditional grading practices and described grades as negative motivators, including how the traditional grading model (re)produces inequity in education. Last, students with lower initial performance on the quizzes were more likely to retake quizzes and most improved their course grades. Below we explore the implications of these findings for students and instructors and the need for systemic change in grading practices.

Second-Chance Quizzes Promote Student Competence, Autonomy, and Relatedness

Although we distinguished student quotes that particularly emphasized competence, autonomy, or relatedness, these dimensions are fundamentally intertwined. In the same breath, students often mentioned how second-chances gave them an opportunity to improve (competence), offered more control over their final grade (autonomy), and made them feel seen and cared for by their instructor (relatedness). SDT posits that autonomous motivation is not achieved by addressing one single construct but, instead, it is the interaction between them that yields the greatest gains in self-determination (Ryan and Deci, 2000a). In the classroom, these constructs interweave to shape a student’s perceived learning environment, and effective changes can shift classwork from a series of chores into a motivating set of tasks that are aligned with a student’s values, long-term goals, and sense of community.

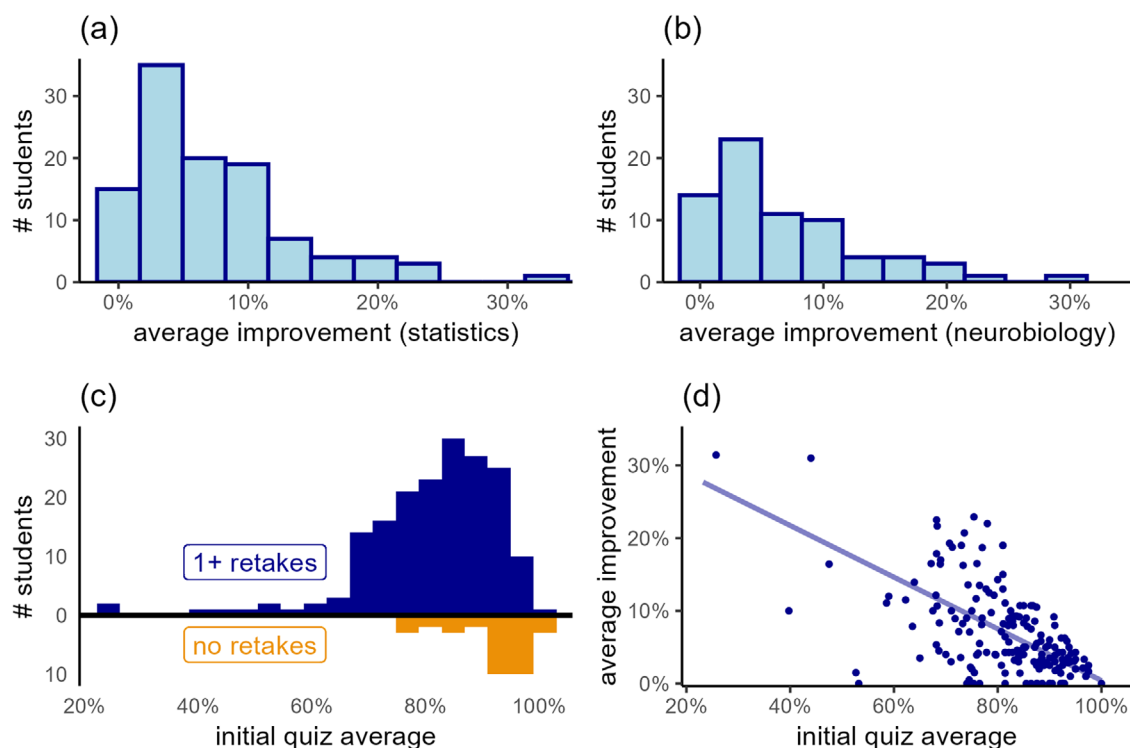


FIGURE 1. (a) Histogram of average quiz score improvement for retakers in statistics, including the 2021, 2022, and 2023 course offerings ($n = 109$). (b) Histogram of average quiz score improvement for retakers in neurobiology, including the 2022 and 2023 course offerings ($n = 71$). (c) Mirror histogram showing initial quiz average distribution for all students in both courses ($n = 213$); blue bars represent students who completed quiz retakes and orange bars represent students who did not. (d) Average quiz score improvement for all retakers in both courses ($n = 180$), as a function of the initial quiz average.

With the goal of promoting student motivation, increasing student autonomy may be the most crucial target for revisions to a course grading scheme (Ryan and Deci, 2000a). Not only did the opportunities for reassessment improve students' confidence and provide students with control over their course learning, it allowed them to shift workload priorities week-by-week to strategically tackle the most salient challenges across all of their commitments. Students experiencing a grading system with in-class retakes felt that their instructors cared about them. When instructors support autonomy, they show their empathy for students' individual needs, promoting relatedness and creating new opportunities for students to build competence.

Students Perceive Grades and Traditional Grading Practices as Harmful to Their Learning, Well-Being, and Motivation

In our investigations, students overwhelmingly described the harmful effects of traditional grading practices—particularly high-stakes assessments—on their mental well-being (e.g., higher stress and anxiety) which ultimately curtailed their motivation. Many lines of evidence highlight the ways in which high-stakes assessments negatively impact student well-being and learning (Ecclestone, 1999; Boud and Falchikov, 2006; Jones *et al.*, 2020; Biggs *et al.*, 2022), and receiving poor grades can directly harm student motivation in multiple ways (Butler and Nisan, 1986; Ryan and Weinstein, 2009). Literature suggests when students receive damaging grades, they

experience less competence, autonomy, and relatedness, and are more inclined to interpret stressful situations as a threat, instead of a positive challenge, which decreases their intrinsic motivation (Weinstein and Ryan, 2011). Unsurprisingly, Kasser and Ryan (2001) found that greater intrinsic motivation is related to lower anxiety and stress for college students. Thus, increasing students' agency over extrinsic outcomes such as grades may ultimately lead to greater well-being, through satisfaction of basic psychological needs for autonomy, competence, and relatedness (Ryan and Deci, 2000b).

Notably, students in this study also indirectly spoke about thoughts indicative of mental disorders such as lack of self-worth and feeling inept due to poor course grades often determined by one or two exams. More grave consequences of high-stakes assessments include student depression, disordered eating, self-harm, and suicide (Ribeiro *et al.*, 2018; Pascoe *et al.*, 2020). Thus, our data suggest that high-stakes assessments not only threaten students' motivation and positive self-perception but also may have more dire consequences on mental and emotional health.

Even when an instructor may feel strongly constrained to use a small number of high-stakes assessments, reassessment opportunities like in-class quiz retakes that promote autonomy and emphasize end-of-term performance can have compounding positive impacts on student motivation, performance, and well-being. The fact that students perceived these benefits from an intervention as simple as a single suite of in-class quiz retakes is heartening and reveals that small instructional

changes may be able to mitigate a substantial chunk of the harmful effects of traditional grading.

Students Identify Grades as Inequitable and Demotivating

In this study, almost a third of students critically reflected on the inequities in grades. One student poignantly described how grades are biased toward White students, while others reflected how grades privilege certain groups of students either through financial means (having to work or not being able to hire a tutor) or social capital (having family members who have taken the same courses). The American grading system was built on 18th century European assignment of scores for White students from affluent families (Schneider and Hutt, 2014). Being that the grading model was first intended for ranking these students, and evolved to ranking all students according to this population (which perpetuated the idea of lower intelligence among immigrants, people in poverty, and students of color), the application of this model in American education continues to disproportionately impacts students from systematically excluded backgrounds in detrimental ways (Tyack, 1974; Feldman, 2023). If we aspire to challenge these long-standing racist and inequitable grading systems, adoption of grading alternatives that are culturally responsive need to be widespread across departments, campuses, and education policy makers to enact real change in grading systems that lead to liberation for people from underserved backgrounds. Although instructors are required to report end of term grades, they can make small yet meaningful assessment changes in their classrooms to subvert these dominant, oppressive grading systems.

Through a more traditional lens, students identified a much more benign purpose of grades, at least at face value, as a way to quantify and communicate student performance and learning. Many scholars would agree with this perspective (Bailey and McTighe, 1996; Snowman and Biehler, 2002; Allen, 2005) as communication of grades is often used to relay information about a student's learning aptitude to graduate and professional schools (Allen, 2005), and also to students themselves. In line with the latter, several students in this study proposed the purpose of grades as an external motivator which directly aligns with the extrinsic regulation of SDT (Ryan and Deci, 2000a). According to Ryan and Deci (2000a), however, extrinsic regulation represents the lowest level of self-determination among extrinsic motivators by involving participation in an activity to secure rewards or social validation, evade penalties, or adhere to external standards. When considering students' future career opportunities and long-term well-being, it seems that grades offer little to no benefit as an external motivator.

Second Chance Quizzes Benefit Students with the Lowest Initial Performance

In addition to promoting student motivation and mitigating stress, our in-class implementation of quiz retakes specifically benefited students who initially performed poorly. Students with lower quiz averages were more likely to complete the quiz retakes, and they earned the greatest overall improvement in their quiz scores, thus reducing performance disparities within these courses. This result contrasts with recent findings for optional quiz or exam retakes. In a course with ungraded voluntary requizing, Walck-Shannon *et al.* (2019)

found that students who were less successful in a previous STEM course were also less likely to retake quizzes. Supriya *et al.* (2024) studied the impact of optional graded exam retakes in large introductory biology courses, finding that the positive impacts of retakes were not evenly distributed among students. Students who completed exam retakes generally improved their scores, but Black/African American students and students working more than 20 h/wk were less likely to complete these retakes, potentially increasing grade disparities among student identities and experiences. Although we could not include demographic data in our quantitative analysis, we were encouraged by our finding that quiz retakes preferentially helped the students with the lowest average quiz scores, which suggests the potential for retakes to reduce rather than amplify grade disparities. Other college exam retakes implemented during class time have similarly found greater participation for students with lower grades (Cates, 1982; Davidson *et al.*, 1984; Friedman 1987). We hypothesize that using class time for retakes is crucial both because this signals that retakes are an expected, integral part of the course and because the retakes can be completed without requiring extra outside-of-class time.

A central goal in equitable teaching is to identify and address the external factors that produce performance disparities. Our results support the notion that grading practices are a powerful tool to combat education inequity when they emphasize learning and provide multiple opportunities to demonstrate proficiency (Nilson, 2015; Blum and Kohn, 2020; Clark and Talbert, 2023; Feldman, 2023). Moreover, the pattern of improvement we observed is not likely to be caused solely by regression to the mean (Galton, 1886; Marsden and Torgerson, 2012), because these initial scores and improvements are averaged across multiple quizzes rather than based on the improvement in individual quiz scores. Reducing classroom opportunity gaps should be a central goal when implementing new alternative grading schemes, and in-class quiz retakes fundamentally address this goal by providing the greatest opportunity for improvement to students with low initial performance.

Many lines of evidence highlight the ways in which high-stakes assessments negatively impact student well-being and learning (Ecclestone, 1999; Boud and Falchikov, 2006; Jones *et al.*, 2020; Biggs *et al.*, 2022). Receiving grades can directly harm student autonomy (Butler and Nisan, 1986; Ryan and Weinstein, 2009), and tests early in the quarter may strongly influence a student's final grade regardless of their end-of-term proficiency (Feldman, 2023). Simple changes to allow second chances shift the nature of early test grades, turning them into sources of direct feedback on learning, which may be less likely to harm student autonomy (Cameron *et al.*, 2001).

LIMITATIONS

There are several noteworthy limitations in this study. Although our study design and approach was rigorous and data saturation criteria were met, a qualitative investigation that included a larger sample size across a wider range of student majors may have revealed other patterns of thought not captured in our interviews. We were also limited in our ability to capture student perceptions of grades who completed

second-chance assessments from other locations outside of a R1 university on the West Coast, potentially hampering our ability to make generalized claims about students' perceptions of grades and alternative grading.

Additionally, students may have exhibited social desirability bias (Paulhus, 1984) in which they may not have been transparent about their perceptions during their interviews based on the authors' dual role as instructors of record for the classes examined and their involvement in the study. We attempted to mitigate this by having an undergraduate researcher unrelated to the courses (A.R.) conduct interviews as well as ensuring anonymity of student responses to the instructors/authors B.T. and R.E.F.

Finally, we note that reassessment opportunities do not fully address the system's overemphasis on grades. Although we observed greater motivation to learn and perform well in these courses, the ultimate goal of learning was fundamentally conflated with the goal of getting a good grade. Second-chance quizzes address some of the inequities that arise within our university grading systems, but do not address the underlying white supremacy that undergirds education.

IMPLICATIONS AND FUTURE DIRECTIONS

Measuring student perceptions of second-chances on assessments reveals the impact that small grading modifications have on student motivation and overall well-being. This gentle grading shift can have long lasting implications for student-instructor relationships by potentially decreasing instructors' stress of accommodating student emergencies and reducing students' hyperfixation on points (Feldman, 2023). We posit that by shifting from a "no-excuses approach" to a "no-excuses needed" mindset, both instructors and students can focus their attention on the learning process in a less stressful and more motivating environment.

We encourage readers to consider the purpose of exams and ask whether they ever effectively measure student learning and proficiency. As a starting point, alternative grading approaches such as in-class retakes and low-stakes assessments are an improvement. They allow for growth in learning and guide students to refine and then re-exhibit their proficiency, centering learning and hopefully cementing longer-term learning and motivation. After all, is this not how the real world works? As one of our participants dutifully noted, when are we ever given only one shot to be tested on minor details and facts in our work? We also echo the calls of many scholarly predecessors to encourage instructors to use formative feedback as an antidote to high-stakes assessments, allowing for pedagogical and curricular adjustments that flexibly serve students from all backgrounds. These first steps may start the process of increasing students' sense of autonomy over their learning and can support their growth as deep thinkers and creators regardless of identity or background.

We recognize that second-chances on assessments are not enough to change the tide of oppressive systems in education. Much more work beyond the redesign of our grading system needs to be done to combat the perpetual exclusion of students who hold minoritized identities. But we contend that second-chance grading and other alternative grading approaches move us one step closer to abolishing an education

model steeped in white supremacy. This begs the question, if students in this study recognized the racist nature of grades, are instructors cognizant of this, too? And if instructors are, why are we still using this system? Out of convenience, habit, or the old adage, "Because we've always done it this way"? We hope this study motivates scholars to systematically explore instructors' beliefs on the purpose of grades and their reasoning for implementing specific grading practices over others.

ACKNOWLEDGMENTS

We would like to thank the gracious student participants who volunteered to contribute to this study. Without their participation, this project would not have been possible. We also appreciate our colleagues, Dr. Dax Ovid and Dr. Mallory Rice, for providing constructive feedback to a draft version of this paper. Thank you to the monitoring editor and reviewers of this work for improving this manuscript by providing invaluable feedback.

REFERENCES

- Allen, J. D. (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 78(5), 218–223. <https://doi.org/10.3200/TCHS.78.5.218-223>
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing. (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- American Psychological Association. (2018). *Well-being*. *APA Dictionary of Psychology*. Washington, DC. Retrieved from <https://dictionary.apa.org/well-being>
- Arjoon, J. A., Xu, X., & Lewis, J. E. (2013). Understanding the state of the art for measurement in chemistry education research: Examining the psychometric evidence. *Journal of Chemical Education*, 90(5), 536–545. <https://doi.org/10.1021/ed3002013>
- Au, W. W. (2008). Devising inequality: A Bernsteinian analysis of high-stakes testing and social reproduction in education. *British Journal of Sociology of Education*, 29(6), 639–651. <https://doi.org/10.1080/01425690802423312>
- Bailey, J., & McTighe, J. (1996). Reporting achievement at the secondary level: What and how. In T. R. Guskey (Ed.), *Communicating student learning*. Arlington, VA: ASCD.
- Ballen, C. J., Salehi, S., & Cotner, S. (2017). Exams disadvantage women in introductory biology. *PLoS One*, 12(10), e0186419. Retrieved from <https://plos.org/10.1371/journal.pone.0186419>
- Baron, P. A. B. (2000). Consequential validity for high school grades: What is the meaning of grades for senders and receivers? Retrieved from <https://eric.ed.gov/?id=ED445051>
- Biggs, J. B., Tang, C. S., & Kennedy, G. (2022). *Teaching for quality learning at university* (5th ed.). Maidenhead, Berkshire, United Kingdom: Open University Press.
- Bloom, B. S. (1968). *Learning for Mastery. Instruction and Curriculum. Regional Education Laboratory for the Carolinas and Virginia, Topical Papers and Reprints, Number 1*. Washington, DC. Evaluation Comment. Retrieved from <https://eric.ed.gov/?id=eD053419>
- Blum, S. D., & Kohn, A. (eds.). (2020). *Ungrading: Why rating students undermines learning (and what to do instead)* (1st ed.). West Virginia University Press.
- Bonefeld, M., & Dickhäuser, O. (2018). (Biased) Grading of students' performance: Students' names, performance level, and implicit attitudes. *Frontiers in Psychology*, 9, 481. <https://doi.org/10.3389/fpsyg.2018.00481>
- Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education*, 31(4), 399–413. <https://doi.org/10.1080/02602930600679050>
- Bowen, R. S., & Cooper, M. M. (2021). Grading on a curve as a systemic issue of equity in chemistry education. *Journal of Chemical Education*, 99(1), 185–194. <https://doi.org/10.1021/acs.jchemed.1c00369>

- Brookhart, S. (2004). *Grading*. London, United Kingdom: Pearson Education.
- Butler, R., & Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal of Educational Psychology*, 78(3), 210–216. <https://doi.org/10.1037/0022-0663.78.3.210>
- Cameron, J., Banko, K. M., & Pierce, W. D. (2001). Pervasive negative effects of rewards on intrinsic motivation: The myth continues. *The Behavior Analyst*, 24(1), 1–44. <https://doi.org/10.1007/BF03392017>
- Campbell, R., Clark, D., & OShaughnessy, J. (2020). Introduction to the special issue on implementing mastery grading in the undergraduate mathematics classroom. *PRIMUS*, 30(8–10), 837–848. <https://doi.org/10.1080/10511970.2020.1778824>
- Canning, E. A., Muenks, K., Green, D. J., & Murphy, M. C. (2019). STEM faculty who believe ability is fixed have larger racial achievement gaps and inspire less student motivation in their classes. *Science Advances*, 5(2), eaau4734. <https://doi.org/10.1126/sciadv.aau4734>
- Cassady, J. C., & Johnson, R. E. (2002). Cognitive test anxiety and academic performance. *Contemporary Educational Psychology*, 27(2), 270–295. <https://doi.org/10.1006/ceps.2001.1094>
- Casselmann, B. L., & Atwood, C. H. (2017). Improving general chemistry course performance through online homework-based metacognitive training. *Journal of Chemical Education*, 94(12), 1811–1821. <https://doi.org/10.1021/acs.jchemed.7b00298>
- Cates, W. M. (1982). The efficacy of retesting in relation to improved test performance of college undergraduates. *The Journal of Educational Research*, 75(4), 230–236. <https://doi.org/10.1080/00220671.1982.10885386>
- Chamberlin, K., Yasué, M., & Chiang, I.-C. A. (2023). The impact of grades on student motivation. *Active Learning in Higher Education*, 24(2), 109–124. <https://doi.org/10.1177/1469787418819728>
- Chapell, M. S., Blanding, Z. B., Silverstein, M. E., Takahashi, M., Newman, B., Gubi, A., & McCann, N. (2005). Test anxiety and academic performance in undergraduate and graduate students. *Journal of Educational Psychology*, 97(2), 268–274. <https://doi.org/10.1037/0022-0663.97.2.268>
- Clark, D., & Talbert, R. (2023). *Grading for growth*, 1st ed. Routledge.
- Cotner, S., & Ballen, C. J. (2017). Can mixed assessment methods make biology classes more equitable? *PLoS One*, 12(12), e0189610. <https://doi.org/10.1371/journal.pone.0189610>
- Crucho, C. I. C., Avó, J., Diniz, A. M., & Gomes, M. J. S. (2020). Challenges in teaching organic chemistry remotely. *Journal of Chemical Education*, 97(9), 3211–3216. <https://doi.org/10.1021/acs.jchemed.0c00693>
- Davidson, W. B., House, W. J., & Boyd, T. L. (1984). A test-retest policy for introductory psychology courses. *Teaching of Psychology*, 11(3), 182–184. <https://doi.org/10.1177/009862838401100320>
- de Brey, C., Musu, L., McFarland, J., Wilkinson-Flicker, S., Diliberti, M., Zhang, A., ... & Wang, X. (2019). *Status and trends in the education of racial and ethnic groups 2018*. Washington, DC: National Center for Education Statistics.
- deCharms, R. (1976). *Enhancing Motivation: Change in the Classroom*, ix, 279. New York, NY: Irvington.
- Deci, E., Koestner, R., & Ryan, R. (2001). Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of Educational Research*, 71(1), 1–27.
- Deci, E. L., & Ryan, R. M. (1985). Conceptualizations of intrinsic motivation and self-determination. In E. L. Deci & R. M. Ryan (Eds.), *Intrinsic motivation and self-determination in human behavior* (pp. 11–40). New York, NY: Springer. https://doi.org/10.1007/978-1-4899-2271-7_2
- Dweck, C. S. (2008). *Mindset: The New Psychology of Success*. New York, NY: Random House Digital, Inc.
- Ecclestone, K. (1999). Empowering or ensnaring?: The implications of outcome-based assessment in higher education. *Higher Education Quarterly*, 53(1), 29–48. <https://doi.org/10.1111/1468-2273.00111>
- Feldman, J. (2023). *Grading for equity: What it is, why it matters, and how it can transform schools and classrooms* (2nd ed.). Thousand Oaks, CA: Corwin Press.
- French, S., Dickerson, A., & Mulder, R. A. (2023). A review of the benefits and drawbacks of high-stakes final examinations in higher education. *Higher Education*, 88, 893–918. <https://doi.org/10.1007/s10734-023-01148-z>
- Friedman, H. (1987). Repeat examinations in introductory statistics courses. *Teaching of Psychology*, 14(1), 20–23. https://doi.org/10.1207/s15328023top1401_4
- Fromme, K., Corbin, W. R., & Kruse, M. I. (2008). Behavioral risks during the transition from high school to college. *Developmental Psychology*, 44(5), 1497–1504. <https://doi.org/10.1037/a0012614>
- Galton, F. (1886). Regression Towards Mediocrity in Hereditary Stature. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 15, 246–263. <https://doi.org/10.2307/2841583>
- Gibbs, G. (2005). *Improving the quality of student learning*. Newport, Wales: University of Glamorgan.
- Gilman, H. (2019). Are we whom we claim to be? A case study of language policy in community college writing placement practices. *Journal of Writing Assessment*, 12(1). Retrieved from <https://escholarship.org/uc/item/3kh925t2>
- Goldman, L. (1985). The betrayal of the gatekeepers: Grade inflation. *The Journal of General Education*, 37(2).
- Guay, F. (2022). Applying self-determination theory to education: Regulations types, psychological needs, and autonomy supporting behaviors. *Canadian Journal of School Psychology*, 37(1), 75–92. <https://doi.org/10.1177/08295735211055355>
- Handelsman, J. (2007). *Scientific teaching*. New York, NY: Macmillan Learning. Retrieved from <https://store.macmillanlearning.com/us/product/Scientific-Teaching/p/1429201886>
- Hannon, B. (2012). Test anxiety and performance-avoidance goals explain gender differences in SAT-V, SAT-M, and overall SAT scores. *Personality and Individual Differences*, 53(7), 816–820. <https://doi.org/10.1016/j.paid.2012.06.003>
- Harris, R. B., Mack, M. R., Bryant, J., Theobald, E. J., & Freeman, S. (2020). Reducing achievement gaps in undergraduate general chemistry could lift underrepresented students into a “hyperpersistent zone”. *Science Advances*, 6(24), eaaz5687. <https://doi.org/10.1126/sciadv.aaz5687>
- Harsy, A. (2020). Variations in mastery-based testing. *PRIMUS*, 30(8–10), 849–868. <https://doi.org/10.1080/10511970.2019.1709588>
- Harsy, A., Carlson, C., & Klamerus, L. (2021). An analysis of the impact of mastery-based testing in mathematics courses. *PRIMUS*, 31(10), 1071–1088. <https://doi.org/10.1080/10511970.2020.1809041>
- Heider, F. (1958). *The psychology of interpersonal relations* (pp. ix, 326). Hoboken, NJ: John Wiley & Sons Inc. <https://doi.org/10.1037/10628-000>
- Hicks, T., & Heastie, S. (2008). High school to college transition: A profile of the stressors, physical and psychological health issues that affect the first-year on-campus college student. *Faculty Working Papers from the School of Education*, 14, Fayetteville, NC. Retrieved from https://digitalcommons.unctsu.edu/soe_faculty_wp/14
- Howard, J. L., Bureau, J. S., Guay, F., Chong, J. X. Y., & Ryan, R. M. (2021). Student motivation and associated outcomes: A meta-analysis from self-determination theory. *Perspectives on Psychological Science*, 16(6), 1300–1323. <https://doi.org/10.1177/1745691620966789>
- Jones, E., Priestley, M., Brewster, L., Wilbraham, S. J., Hughes, G., & Spanner, L. (2020). Student wellbeing and assessment in higher education: The balancing act. *Assessment & Evaluation in Higher Education*, 46(3), 438–450. <https://doi.org/10.1080/02602938.2020.1782344>
- Juhler, S. M., Rech, J. F., From, S. G., & Brogan, M. M. (1998). The effect of optional retesting on college students' achievement in an individualized algebra course. *The Journal of Experimental Education*, 66(2), 125–137. <https://doi.org/10.1080/00220979809601399>
- Kain, D. L. (1996). Looking beneath the surface: Teacher collaboration through the lens of grading practices. *Teachers College Record*, 97(4), 569–587.
- Kasser, T., & Ryan, R. M. (2001). Be careful what you wish for: Optimal function and the relative attainment of intrinsic and extrinsic goals. In P. Schmuck & K. Sheldon (Eds.), *Life goals and well-being: Towards a positive psychology of human striving*. Seattle: Hogrefe & Huber.
- Kellaghan, T., & Greaney, V. (2019). *Public Examinations Examined*. Washington, DC: World Bank Publications.
- Klausman, J., & Lynch, S. (2022). Chapter 2. From ACCUPLACER to informed self-placement at Whatcom Community College: Equitable placement as an evolving practice. In J. Nastal, M. Poe, & C. Toth (Eds.), *Writing placement in two-year colleges: The pursuit of equity in postsecondary education* (pp. 59–83). Louisville, Colorado: The WAC Clearinghouse; University Press of Colorado. <https://doi.org/10.37514/PRA-B.2022.1565.2.02>
- Kwong, D., & Davis, J. R. (2015). School climate for academic success: A multilevel analysis of school climate and student outcomes. *Journal of Research in Education*, 25(2), 68–81.

- Leeming, F. C. (2002). The exam-a-day procedure improves performance in psychology classes. *Teaching of Psychology*, 29(3), 210–212. https://doi.org/10.1207/S15328023TOP2903_06
- Longhurst, R. (2003). Key methods in geography. *Semi-structured interviews and focus groups* (4th ed., pp. 168–183). Thousand Oaks, CA: Sage.
- Mangum, A. (2020). Implementation of mastery-based testing in Calculus I. *PRIMUS*, 30(8–10), 869–884. <https://doi.org/10.1080/10511970.2019.1709587>
- Marsden, E., & Torgerson, C. J. (2012). Single group, pre- and post-test research designs: Some methodological concerns. *Oxford Review of Education*, 38(5), 583–616. <https://doi.org/10.1080/03054985.2012.731208>
- Marzano, R. (2011). *Formative assessment & standards-based grading*. Bloomington, IN: Solution Tree Press.
- Moos, R. H., & Moos, B. S. (1978). Classroom social climate and student absences and grades. *Journal of Educational Psychology*, 70(2), 263–269. <https://doi.org/10.1037/0022-0663.70.2.263>
- Muenks, K., Canning, E. A., LaCrosse, J., Green, D. J., Zirkel, S., Garcia, J. A., & Murphy, M. C. (2020). Does my professor think my ability can change? Students' perceptions of their STEM professors' mindset beliefs predict their psychological vulnerability, engagement, and performance in class. *Journal of Experimental Psychology: General*, 149(11), 2119. <https://doi.org/10.1037/xge0000763>
- Nastal, J. (2019). Beyond tradition: Writing placement, fairness, and success at a two-year college. *Journal of Writing Assessment*, 12(1). Retrieved from <https://escholarship.org/uc/item/4wg8w0ng>
- Ngo, F., & Melguizo, T. (2020). The equity cost of inter-sector math misalignment: Racial and gender disparities in community college student outcomes. *The Journal of Higher Education*, 92(3), 410–434. <https://doi.org/10.1080/00221546.2020.1811570>
- Nilson, L. B. (2015). *Specifications grading: Restoring rigor, motivating students, and saving faculty time*. Sterling, VA: Stylus Publishing, LLC.
- Nix, G. A., Ryan, R. M., Manly, J. B., & Deci, E. L. (1999). Revitalization through self-regulation: The effects of autonomous and controlled motivation on happiness and vitality. *Journal of Experimental Social Psychology*, 35(3), 266–284. <https://doi.org/10.1006/jesp.1999.1382>
- Núñez-Peña, M. I., Suárez-Pellicioni, M., & Bono, R. (2016). Gender differences in test anxiety and their impact on higher education students' academic achievement. *Procedia - Social and Behavioral Sciences*, 228, 154–160. <https://doi.org/10.1016/j.sbspro.2016.07.023>
- Orr, R., & Foster, S. (2013). Increasing student success using online quizzing in introductory (majors) biology. *CBE—Life Sciences Education*, 12(3), 509–514. <https://doi.org/10.1187/cbe.12-10-0183>
- Pascoe, M. C., Hetrick, S. E., & Parker, A. G. (2020). The impact of stress on students in secondary school and higher education. *International Journal of Adolescence and Youth*, 25(1), 104–112. <https://doi.org/10.1080/02673843.2019.1596823>
- Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Thousand Oaks, CA: Sage.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46(3), 598–609. <https://doi.org/10.1037/0022-3514.46.3.598>
- Pennebaker, J. W., Gosling, S. D., & Ferrrell, J. D. (2013). Daily online testing in large classes: Boosting college performance while reducing achievement gaps. *PLoS One*, 8(11), e79774. <https://doi.org/10.1371/journal.pone.0079774>
- Perlmutter, L., Everson, J., Yasuhara, K., Wortzman, B., & Lin, K. (2022). Reading between the lines: Student experiences of resubmission in an introductory CS course. *Proceedings of the 53rd ACM Technical Symposium on Computer Science Education*, 2, 1137. <https://doi.org/10.1145/3478432.3499112>
- Perna, L. W. (2010). Understanding the working college student. *Academe*, 96(4). Retrieved from <https://www.aaup.org/article/understanding-working-college-student>
- R Core Team. (2023). *R: A language and environment for statistical computing* (4.3.1) [Computer software]. R Foundation for Statistical Computing. Retrieved from September 1, 2023 <https://www.R-project.org/>
- Ramsden, P. (2003). *Learning to teach in higher education*, 2nd ed. Milton Park, United Kingdom: RoutledgeFalmer. Retrieved from <https://www.routledge.com/Learning-to-Teach-in-Higher-Education/Ramsden-Ramsden/p/book/9780415303453>
- Randall, J., Poe, M., Oliveri, M. E., & Slomp, D. (2023). Justice-oriented, antiracist validation: Continuing to disrupt white supremacy in assessment practices. *Educational Assessment*, 29(1), 1–20. <https://doi.org/10.1080/10627197.2023.2285047>
- Reeve, J. (2016). Autonomy-supportive teaching: What it is, how to do it. In W. C. Liu, J. C. K. Wang, & R. M. Ryan (Eds.). *Building autonomous learners: Perspectives from research and practice using self-determination theory* (pp. 129–152). New York, NY: Springer. https://doi.org/10.1007/978-981-287-630-0_7
- Reeve, J., & Cheon, S. H. (2021). Autonomy-supportive teaching: Its malleability, benefits, and potential to improve educational practice. *Educational Psychologist*, 56(1), 54–77. <https://doi.org/10.1080/00461520.2020.1862657>
- Ribeiro, Í. J. S., Pereira, R., Freire, I. V., de Oliveira, B. G., Casotti, C. A., & Boery, E. N. (2018). Stress and quality of life among university students: A systematic literature review. *Health Professions Education*, 4(2), 70–77. <https://doi.org/10.1016/j.hpe.2017.03.002>
- Rodgers, W. M., Markland, D., Selzer, A.-M., Murray, T. C., & Wilson, P. M. (2014). Distinguishing perceived competence and self-efficacy: An example from exercise. *Research Quarterly for Exercise and Sport*, 85(4), 527–539. <https://doi.org/10.1080/02701367.2014.961050>
- Ryan, R. M., & Deci, E. L. (2000a). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 54–67. <https://doi.org/10.1006/ceps.1999.1020>
- Ryan, R. M., & Deci, E. L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Ryan, R. M., Deci, E. L., & Grolnick, W. S. (1995). Autonomy, relatedness, and the self: Their relation to development and psychopathology. *Developmental psychopathology, Theory and methods* (Vol. 1, pp. 618–655). Hoboken, NJ: John Wiley & Sons.
- Ryan, R. M., & Weinstein, N. (2009). Undermining quality teaching and learning: A self-determination theory perspective on high-stakes testing. *Theory and Research in Education*, 7(2), 224–233. <https://doi.org/10.1177/1477878509104327>
- Saldaña, J. (2016). *The coding manual for qualitative researchers*. Thousand Oaks, CA: Sage. Retrieved from <https://www.torrossa.com/en/resources/an/5018667>
- Schneider, J., & Hutt, E. (2014). Making the grade: A history of the A–F marking scheme. *Journal of Curriculum Studies*, 46(2), 201–224. <https://doi.org/10.1080/00220272.2013.790480>
- Scott-Clayton, J., Crosta, P. M., & Belfield, C. R. (2014). Improving the targeting of treatment: Evidence from college remediation. *Educational Evaluation and Policy Analysis*, 36(3), 371–393. <https://doi.org/10.3102/0162373713517935>
- Seipp, B. (1991). Anxiety and academic performance: A meta-analysis of findings. *Anxiety Research*, 4(1), 27–41. <https://doi.org/10.1080/0891779108248762>
- Seymour, E., & Hewitt, N. (1997). *Talking about Leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview Press.
- Seymour, E., & Hunter, A.-B. (eds.). (2019). *Talking about leaving revisited: Persistence, Relocation, and loss in undergraduate STEM education*. Cham, Switzerland: Springer International Publishing. <https://doi.org/10.1007/978-3-030-25304-2>
- Sheldon, K. M., Ryan, R. M., Rawsthorne, L. J., & Ilardi, B. (1997). Trait self and true self: Cross-role variation in the Big-Five personality traits and its relations with psychological authenticity and subjective well-being. *Journal of Personality and Social Psychology*, 73(6), 1380–1393. <https://doi.org/10.1037/0022-3514.73.6.1380>
- Shor, I. (1996). *When students have power: Negotiating authority in a critical pedagogy*. Chicago, IL: The University of Chicago Press.
- Snowman, J., & Biehler, R. (2002). *Psychology applied to teaching*, 10th ed. Boston, MA: Houghton Mifflin College Div.
- Sorensen-Unruh, C. (2024). The ungrading learning theory we have is not the ungrading learning theory we need. *CBE—Life Sciences Education*, 23(3), es6. <https://doi.org/10.1187/cbe.24-01-0031>
- Strutz, T. (2016). *Data fitting and uncertainty*. Weisbaden, Germany: Springer Vieweg Wiesbaden. Retrieved from <https://link.springer.com/book/9783658114558>
- Sullivan, S. G. B., Hoiriis, K. T., & Paolucci, L. (2018). Description of a change in teaching methods and comparison of quizzes versus midterms scores in a research methods course. *Journal of Chiropractic Education*, 32(2), 84–89. <https://doi.org/10.7899/JCE-17-20>
- Supriya, K., Bang, C., Ebbe, J., Pagliarulo, C., Tucker, D., Villegas, K., ... & Brownell, S. (2024). Optional exam retakes reduce anxiety but may

- exacerbate score disparities between students with different social identities. *CBE—Life Sciences Education*, 23(3), ar30. <https://doi.org/10.1187/cbe.21-11-0320>
- Tripp, B., & Shortlidge, E. E. (2020). From theory to practice: Gathering evidence for the validity of data collected with the Interdisciplinary Science Rubric (IDSR). *CBE—Life Sciences Education*, 19(3), ar33. <https://doi.org/10.1187/cbe.20-02-0035>
- Tyack, D. B. (1974). *The one best system: A history of American urban education*. Cambridge, MA: Harvard University Press.
- van der Meer, J., Jansen, E., & Torenbeek, M. (2010). 'It's almost a mindset that teachers need to change': First-year students' need to be inducted into time management. *Studies in Higher Education*, 35(7), 777–791. <https://doi.org/10.1080/03075070903383211>
- Vatterot, C. (2015). *Rethinking grading: Meaningful assessment for standards-based learning* (1st ed.). Alexandria, VA: ASCD.
- VERBI Software. (2021). *MAXQDA 2022 [computer software] (Version 2022) [Computer software]*. Retrieved from July 1, 2023 maxqda.com
- Vyas, V. S., Nobile, L., Gardinier, J. R., & Reid, S. A. (2023). Impact of a low-stakes assessments model with retake in general chemistry: Connecting to student attitudes and self-concept. *Education Sciences*, 13(12), 1235. <https://doi.org/10.3390/educsci13121235>
- Walck-Shannon, E. M., Cahill, M. J., McDaniel, M. A., & Frey, R. F. (2019). Participation in voluntary re-quizzing is predictive of increased performance on cumulative assessments in introductory biology. *CBE—Life Sciences Education*, 18(2), ar15. <https://doi.org/10.1187/cbe.18-08-0163>
- Weinstein, N., & Ryan, R. M. (2011). A self-determination theory approach to understanding stress incursion and responses. *Stress and Health*, 27(1), 4–17. <https://doi.org/10.1002/smi.1368>
- Whitcomb, K. M., & Singh, C. (2021). Underrepresented minority students receive lower grades and have higher rates of attrition across STEM disciplines: A sign of inequity? *International Journal of Science Education*, 43(7), 1054–1089. <https://doi.org/10.1080/09500693.2021.1900623>
- Williams, G. C., & Deci, E. L. (1996). Internalization of biopsychosocial values by medical students: A test of self-determination theory. *Journal of Personality and Social Psychology*, 70(4), 767–779. <https://doi.org/10.1037/0022-3514.70.4.767>
- Wladis, C., Hachey, A. C., & Conway, K. (2018). No time for college? An investigation of time poverty and parenthood. *The Journal of Higher Education*, 89(6), 807–831. <https://doi.org/10.1080/00221546.2018.1442983>
- Zysberg, L., & Schwabsky, N. (2021). School climate, academic self-efficacy and student achievement. *Educational Psychology*, 41(4), 467–482. <https://doi.org/10.1080/01443410.2020.1813690>