Daniel Guberman, PURDUE UNIVERSITY, dguberma@purdue.edu

Student Perceptions of an Online Ungraded Course

ABSTRACT

What do grades mean? What purpose do they serve? What role do they play in the learning process? Teachers and scholars have recently begun to re-examine these questions central to our current grading system. As a result, many have started to re-assess how grades are assigned in their classes. In this case study, I examine the effectiveness of ungrading, an approach centered around students assigning their own grades through reflecting on the learning process. After contextualizing and describing the approach developed for this fully online, asynchronous history class, I share quantitative and qualitative data regarding student perceptions, motivation, and information usage to argue that systems such as ungrading have potential for contributing to the construction of highly effective and meaningful learning environments.

KEYWORDS

ungrading, assessment, partnership, self-determination theory

INTRODUCTION

In most higher education systems, multi-interval grading schemes function as a primary representation of success and/or failure, summarizing learning from a whole course, or even an entire curriculum into a single letter or number. Yet, explorations of what grades mean, and critical looks at how they help or hinder student accomplishment remain largely absent. In recent years, critiques of grades and grading systems have become more prominent, yet there are few studies examining alternative approaches. This study seeks to explore some of this territory through examining student experiences in a 50-student, online, asynchronous history class that follows an "ungraded" model in which students identify their own grade through reflection on their learning. First, I provide background on the development of grades and grading, suggesting that the now-prominent systems are both not as long-standing as many assume and more opaque than should be acceptable for such common practice. Second, I describe the structure of this ungraded course. Third, I document evidence of student perceptions relating to their experience, motivation, and engagement through three survey tools and student voices. I conclude with an invitation for others to expand upon the potential revealed through these data.

GRADES: HOW DID WE GET HERE?

Depending on how we view grades and higher education, we can trace either a very long or relatively short history. Some trace grading back to the Chinese Civil Service Exam seeking to compare and measure candidates for potential positions, with a history well over a thousand years (Franke 1960). Cambridge University began sorting students into three categories (25% at the top, 50% in the middle, and 25% at the bottom) in the middle of the 16th century (Winter 1993). In the United States, grades

became fairly common by the start of the 20th century, and as described by Smith and Smith (2019) these final grades represented standing in the course based on a combination of "progress, conduct, attentiveness, effort, class attendance, chapel attendance, and academic achievement." In short, grades had little consistency in practice and meaning. In the second decade of the 20th century, grades began to see reform in American K-12 and higher education institutions alongside the emerging science of statistics (Brookhart and Guskey 2019). With grades already in place primarily as a determinant of relative standing, statisticians felt that these rankings should align with a normal distribution. Scholars assumed that ability and achievement should follow this distribution so grading could only be accurate if these were the results (3% As, 22% Bs, 50% Cs, etc.).

In their examination of recent studies of grades in higher education, Smith and Smith (2019) explore the question of "grade inflation." Research suggests that the Vietnam War and associated draft pushed many faculty members to raise student grades through recognizing the potentially extreme consequences for failing (being sent to war). While this correlation seems reasonable, this era also saw the emergence of criterion-based grading, described by Glaser (1963). This represents the beginning of a large-scale shift from grades based on relative standing to grades based on absolute levels of achievement, aligning with Bloom's description of "learning for mastery" (1968) and arguments that all or nearly all students could succeed at very high levels. Outside of the United States, student movements in the late 1960s also challenged grading practices and structures. In their documentation of student uprisings and resistance in Paris, France, at this time, Schnapp and Vidal-Naquet (1971) reprinted numerous documents depicting ways to challenge and/or undermine the validity of examination- and grade-based structures. Entering the 1970s, this "inflation" of grades generally stopped and student protest movements in the United States and Europe slowed.

Many have documented a new rise in average grades, starting in the 1990s, still often termed "inflation," particularly in media (Rampell 2010), but few studies have really examined what grades mean. Scholars, for example, have tried to argue that a rise in grades might result from institutional policies and incentives. Love and Kotchen (2010) proposed a model in which an increased emphasis on teaching evaluations in the promotion process would lead instructors to inflate grades in an effort to raise their evaluations. Ginexi (2003) also argued for a correlation between higher grades and course evaluations, positing that either students may use evaluations as a means to exact revenge or, more likely, that students would associate their lack of success with an instructor's ineffective teaching. There should be some connection between an instructor's teaching approach and a student's lack of learning, but it does not seem to influence course evaluations directly as much as many assume. Two meta-analyses (Centra and Creech 1976; Feldman 1997) demonstrated little support connecting grades directly with evaluations. More recent studies have further challenged assumptions that evaluations tie directly to grades. For example, Wright and Palmer (2006) found that grades could not independently predict course ratings, but instead ratings reflected a combination of motivation, ability, and learning in conjunction with grades. Johnson, Jones, Weidner, and Manwell (2019) go a step further in contradicting theories of retributional bias (students using course ratings as retribution for poor grades) through their findings about course rigor's influence on course and instructor ratings.

During the decades since Glaser and Bloom, higher education in the United States experienced a dramatic overhaul in grading philosophies at the university level, moving from norm-based systems of comparison to systems rooted directly in criterion and measurement. Unfortunately, we do not have any longitudinal studies exploring how and why these changes were made, nor what they resulted in. In eras

of grade inflation fears, we may merely be missing the fact that we are more appropriately supporting and accurately measuring achievement. We may also be partly looking in the wrong place. While some suggest that universities have "inflated" grades to make them more appealing to potential students, the reality is that this coincides with significant investments in professional and peer support for students (tutoring centers, pro-active advising, mentoring programs, high-impact practices, etc.), as well as a boom in research and dissemination of research on effective teaching practices across most disciplines (Chick, Nowell, and Lenart 2019 describe their attempt to track the range of development of the field). These ideas are interrelated. If our grading systems reveal more clearly where students struggle and the degree to which they struggle, then we are more likely to develop our teaching in a way that addresses these struggles directly and enhances overall learning and achievement.

While my proposal that rising grades reflect a shift to achievement-based systems sounds nice for those of us studying teaching and learning, the change from norm-based grading as a comparison with peers to criterion-based systems remains incomplete at best. I suspect that criterion-based approaches expanded significantly starting in the 1990s, with the emergence and promotion of the Scholarship of Teaching and Learning as a discipline (Boyer 1990; Hutchings, Huber, and Ciccone 2011), as well as the professionalization and expansion of teaching and learning centers (Sorcinelli et al. 2006). Around the same time, faculty developers began helping instructors apply backward design principles (Fink 2002; Wiggins and McTighe 2005) focused on aligning assessments with outcomes and activities. While none of these developments deal explicitly with translating these assessments into grades, the presumption always seems to be that grades will reflect achievement of learning outcomes. More recently, tools and advice for creating more effective rubrics, as well as broader promotion of rubrics, also suggests broadening adoption of transparent achievement-focused grading. This has been particularly prominent in disciplines traditionally resistant to rubric construction (Haugnes, Holmgren, and Springborg 2018). Nilson (2015) has pushed these boundaries further, developing an explicit grading system from Bloom's ideas about mastery, including multiple options for how faculty may translate mastery levels into currently common grading schema.

Writings throughout the past few years suggest the emergence of a new trend focused on aligning grades with learning. Franke's (2018) exploration of weighting final exams focuses on how different weights for the exam might reward students for growth throughout a semester. Similarly, increasing discussions around growth mindsets in higher education have shifted the instructors' role from sorting students into groups based on ability (normative systems) to agents of change who play a significant role in students' ability to succeed (Dweck 2006). Growth mindsets have aligned with the promotion of universal design for learning, which pushes instructors to focus on providing access, emphasizing thinking about different ways of viewing, and measuring accomplishment (CAST 2018; Tobin and Behling 2018).

Looking beyond the United States, norm-based grades and examination processes reflect one of the many forms in which colonialist forces and traditions maintain a lasting influence. For example, Divala describes how grading and ranking students in South African universities reflect the forces of neoliberal-driven globalization, which derail the possibility of a civic-role of the African university and by extension "a new kind of being-in the-world, where the individual and community do not exist as the antithesis of each other, but rather as complementary aspects of human existence" (2017, 157). More broadly, Ndlovu-Gatsheni identifies the "imposition of Western epistemology and shaping the formative process of development of black subjectivity" (2013, 8) as one of "four constitutive elements of Western

domination and exploitation of the non-Western world" (7). Later, Ndlovu-Gatsheni expands: "It is within African institutions of learning that the African agenda continues to be lost… African values and aspirations have remained outside the school, college and university curriculum. These institutions continue to produce 'mimic' men and women" (62). Obviously, these issues run much more deeply than grading systems, but the imposition of Western grading structures, whether it is the direct imposition of French examination systems in Cameroon in the 1960s (Ramoupi and Ntongwe 2017) or the continued practice of sending doctoral theses to readers in the United States and Europe for assessment (Ndofirepi and Cross 2017), have been part of the practices that impede efforts toward decolonization of higher education.

THE EMERGENCE OF UNGRADING

Stommel has been the strongest proponent of what might be termed a learning-focused grading system, which he calls ungrading, through his insistence on not grading his students' work. This has appeared throughout his writings, particularly many of the essays collected in his co-authored collection *An Urgency of Teachers* (Morris and Stommel 2018). He has also discussed this approach on his website (Stommel 2020) and extensively on Twitter. Following his lead, many others have adopted ungrading in various ways, and it has appeared in formats such as Twitter and publications such as the *Chronicle of Higher Education* (e.g., Supiano 2019).

In the class studied here, I attempted to merge Stommel's approach with principles of student partnerships (Cook-Sather, Bahti, and Ntem 2019; Cook-Sather, Bovill, and Felten 2014; Mercer-Mapstone and Abbot 2020). These concepts pair naturally as they may both be viewed as outgrowths of critical pedagogy movements, drawing heavily on ideas from Freire (2005) and hooks (1994), while they look at different elements of the course. The addition of methods from student partnership enabled me to successfully translate ideas from ungrading to the field of history, which traditionally has had a stronger emphasis on content than in the composition classes that form the basis of many of Stommel's writings.

This paper seeks to address one of the core challenges facing the expansion of ungrading practices: the lack of traditional evidence of effectiveness. In many studies examining the effectiveness of new teaching practices we rely on semi-controlled experiments (Freeman et. al. 2014; Theobald et. al. 2020). These experiments often rely on one of two measures of effectiveness, either a comparison of overall grades or a comparison of examination scores. I would not consider either of these valid measures in a switch to an ungraded course. By its nature, ungrading does not align well with the practice of testing, or at least relying on a single attempt at a test as indicative of learning. Similarly, ungrading fundamentally changes what a final grade measures (learning as opposed to achievement or comparative standing), so any attempt to compare based on these standard measures seems disingenuous. Significant work could still be done in comparing student written work before and after changes to an ungraded system. However, I would be wary of such a comparison because the final product in my approach to ungrading is not the essay assignment, but the reflection a student writes about that assignment. Following Eyler (2018), ungrading allows me to promote a culture that embraces and learns from struggle and failure, rather than punishing it.

COURSE STRUCTURE

The course examined here, titled "Society, Culture, and Rock and Roll" was taught as a fully online, asynchronous course during an eight-week summer session. It is listed as an upper level elective in the history department and it draws diverse majors from the large, U.S., public, STEM-focused, doctoral-granting institution. While open to everyone on campus, it reaches the maximum enrollment of 50 students before first- and second-year students have the opportunity to register because it fulfills a core-curriculum requirement and addresses a popular topic. Therefore, students tend to be upper-division (third or fourth years). The instructor has taught the course in a variety of formats and structures across three institutions over the past decade. The eight-week version maintains the same content and structure (with some adjustments enabled by the omission of exams from the new version of the course). In reverting to a full, 16-week version of the course the primary change would involve adding an additional reflection.

Course design did not follow a strictly linear backward model (outcomes; then assessments; then activities and content). The structure of assessment here led to me to rethink the types of outcomes possible without tests and with student-driven content. The resulting outcomes focused more on communication and research skills (defined broadly), as well as metacognitive learning skills. These fit with my broader teaching aims and directions. Additionally, because students would assess their level of accomplishment for each outcome, the outcomes had to be framed in non-absolute terms. Often contemporary outcomes state a capability to do something specific, but for me this grading structure lacked such absolutes (nor do I view learning through absolutes). The outcomes I developed reflected ways of engagement that students could assess through different levels of complexity and perceptions of learning. For example, instead of mastering one specific type of presentation, I described high achievement through experimentation with multiple modes and reflection on success, as well as benefits and drawbacks of using these modes.

The assignments also changed significantly. An ungraded system does not align well with tests and quizzes other than in the context of memory or mastery practice. With this in mind, I transitioned to placing greater emphasis on students researching and sharing information. In moving away from individual graded assignments, I had to determine what students would do with their learning. I decided that, with the exception of reflections, all student work would be shared as presentations. Students completed five such presentations during the course. The first and last presentations asked specific tasks (in the first students analyzed a song of their choice and in the last they make a case for including a recent song in a future iteration of the course). The three middle presentations were left open such that students could present on any topic that they viewed as connected to the content (generally two decades of music). I provided a variety of prompts students could choose to use or ignore, as well as guides for accessing and navigating archival resources of music magazines and historical newspaper databases.

Aligning with the concept of "non-disposable assignments," student presentations were designed to be shared asynchronously in consistent groups of 12 students, rather than for my consumption or limited to the traditional student-teacher dyad (Seraphin et al. 2019). This allowed me to emphasize skills in finding and analyzing information, as well as sharing and responding to others. Furthermore, because the assessment methods no longer assigned strict points to discussion, participation changed in quality. Students were expected to explain why they thought their participation and engagement demonstrated meaningful responses, rather than meeting word or post counts. This approach to discussion drew on Gillis (2019). The only number I provided for students with regard to

length was that their colleagues should be able to engage meaningfully with their presentation in three to seven minutes. This does not mean there was a seven-minute limit; I encouraged the inclusion of options for peers who wanted to dive more deeply.

The core assignments in the class fit into three categories. First, each student created and shared a series of five presentations in which each person decided what they wanted to explore and share. I included possible topics and guidance for anyone who wanted. Second, each student engaged with others through responding to their presentations. Students were divided into four groups of 12 to make sharing more reasonable and facilitate the construction of community. Finally, students each completed three learning reflections, which were guided by a series of questions. The last question in each reflection asked students what grade they felt they were currently earning and why, which allowed us to have a clear and precise discussion regarding their grades and what they might do to enhance their grade. The last series of assignments I described as "optional." This involved a series of three paper/assignment prompts that I have found useful in previous iterations of the course. If a student sought to pass the course, they could ignore these assignments. However, the assignments provided students with opportunities to enhance their learning and by extension their grade, through exploring the content and ideas in new ways and sharing/discussing with others.

While I did not "grade" the presentations or discussions in terms of numbers or letters, I did provide feedback on every presentation. This included asking follow-up questions, highlighting other possible resources or interpretations, or just sharing an element that I found particularly interesting. I engaged in the discussion forums lightly, often either asking further questions if I felt like a group could use a boost, or by sharing my own experiences as part of broader discussions. I gave much more detailed feedback with each of the individual reflections, particularly in response to the question about students' desired feedback. This meant that I could give detailed and precise feedback that students were likely to engage with. I also gave detailed feedback about the grade they assigned themselves, particularly in the rare cases where we disagreed. In terms of time commitment, this process was comparable to time spent with earlier iterations of the class, which included three exams (multiple choice, short answer and essay questions), two papers (~1,200 words each), weekly journal entries (~250 words), and weekly discussion forums (marked for completion). While both versions involved significantly more time spent providing feedback than other potential course structures, in the new version I got the impression the feedback offered more value to students in the class (in earlier iterations, I suspect less than half of the students engaged with the feedback in a meaningful way).

Because of the nature of my institution, ultimately I needed to give students a specific grade at the end of the semester, summarizing their work, so all of the reflections ended with students assigning themselves a grade based on qualities described in the syllabus. If a student suggested a significantly higher grade than I thought their work represented and did not sufficiently explain the disparity in their reflection, I wrote an email to them directly. More often, students suggested grades slightly lower than I had in mind, or sometimes their reflections highlighted elements of learning that I had not noticed, justifying the grade they proposed. At the end of the semester, fewer than 10% of the 50 students suggested final grades that differed from my own expectations.

EXAMINING STUDENT PERCEPTIONS

In changing to ungrading structures, it would not be appropriate to use grades as a point of comparison for student learning, and the new presentation assignment format prevented an attempt to

compare student work directly. While I found a great deal of meaningful discussion about students' learning in the reflections, here I focus on the data collected through an optional, anonymous survey distributed two weeks before the semester ended and completed by 20 students, focusing on their perceptions. I explained to students at the start and throughout the semester that this structure was new to me and that I wanted to understand their experiences as part of my scholarship on this approach to teaching. I combined a set of survey tools that provide a sense of the environment, motivation, and use of information. Two of the three survey tools draw on self-determination theory (Ryan and Deci 2017), a theory of human motivation. Self-determination theory focuses on the degree to which an environment meets people's three basic psychological needs: competence, autonomy, and relatedness. This theory drove many of the ideas underlying my approach to the course, focusing on autonomy in how students approach the content and work, developing competence through the reflective learning process, and building community by having all work shared in small groups. Below I describe and share results from the three survey tools I used, as well as responses to the open-ended comments, which reveal what these results look like in the words of the students in the class.

The Learning Climate Questionnaire (table 1) used a seven-point scale (1=strongly disagree, 7=strongly agree, Williams and Deci 1996), across six questions, where students indicated the extent to which they found the course autonomy supportive. This scale has been used in a range of studies focusing on higher education, for example, Levesque-Bristol, Knapp, and Fisher's (2010) study of service learning and Bonem, Fedesco, and Zissimopoulos' (2020) comparison of course design models. Each individual question had a mean of 6.4 or higher, with a total mean of 6.53. Of the students who responded, only one question included a score of 4 (neutral) and no questions had responses lower than 4, meaning that students never indicated disagreement with the prompts.

Table 1. Learning climate questionnaire

These items were rated on a 7-point Likert scale (1=strongly disagree; 7=strongly agree) n=20				
Learning climate prompt	Min.	Max.	Mean	Std. dev.
I feel that my instructor provides me choices and options.	5	7	6.70	.571
I feel understood by my instructor.	5	7	6.45	.686
My instructor conveyed confidence in my ability to do well in the course.	6	7	6.60	.503
My instructor encouraged me to ask questions.	5	7	6.55	.605
My instructor listens to how I would like to do things.	5	7	6.50	.607
My instructor tries to understand how I see things before suggesting a new way to do things.	4	7	6.40	.821
Total LC	**5.5**	**7**	**6.53**	**.497**

Student comments reinforced the value of autonomy in the class, particularly around the freedom to select their own assignments. For example, one student wrote: "I didn't think I would like the open format of the assignments. I actually almost dropped the class. But now I feel it was the best part of the class. I was able to study and learn about what I wanted to. This made it easier, and I feel I was able to put more of my personality into the presentations because of it." This comment also reflects the broader philosophical and political aims of this course structure, including an effort to respect the knowledge and experiences that students bring to the classroom through the flexibility in their presentations and

discussions. For me, this responds to a call to bring the interrelated concepts of culturally responsive teaching (Ladson-Billings 1995) and Funds of Knowledge (Moll et al. 1992) together in higher education pedagogy (Kiyama, Rios-Aguilar, and Deil-Amen 2018). Other student commentary highlighted the value of learning from the knowledge and experiences of others, even when they did not follow my urging to explore beyond their comfort zone in their own presentations: "Class discussion helped expose me to different kinds of music and artists even if I chose a topic that I already knew much about."

The Situational Motivation Scale (table 2) (Guay, Vallerand, and Blanchard 2000), relies on 18 questions, using the same seven-point scale, to assess student perceptions of their motivation for participating in the course. This breaks down into six subscales aligning with the six forms of motivation proposed by Ryan and Deci (2017): intrinsic motivation, integration, identification, introjection, extrinsic motivation, and amotivation (listed here in order from most internalized to most controlled types). This index is often used in conjunction with the Learning Climate Questionnaire (including the two examples cited above). These subscales revealed high levels of internalized or autonomous motivation and low levels of externalized or controlled motivation. It is important to remember that each of the six motivational types are measured independently so these patterns do not always appear. For example, because many students enroll in courses like this one expecting them to be a fun and easy way to fulfill a requirement, when they enter the class we might find high levels of both intrinsic motivation, which asks if it is fun, and external regulation, which reflects a sense of being forced to do something (fulfilling a university requirement).

Table 2. Situational motivation scale

These items were rated on a 7-point Likert scale (1=strongly disagree; 7=strongly agree) n=20				
Motivation type	Min.	Max.	Mean	Std. dev.
Intrinsic regulation	3.67	7.00	6.0784	.88608
Integration	1.67	7.00	5.5098	1.60346
Identification	3.00	7.00	5.8039	1.20219
Introjection	1.00	6.00	2.4706	1.30734
Extrinsic regulation	1.00	5.00	3.0196	1.35129
Amotivation	1.00	2.67	1.3922	.48926

Because my course topic has a reputation for being "fun," the ratings for identification and integration strike me as particularly valuable. I consider integration and identification the ideal forms of motivation for any class. Identified motivation occurs when a student cognitively understands what they are being asked to do and how it will help them reach their aims or goals. Integration reflects a further step of internalization in which students move beyond that cognitive comprehension to an internalized belief that the course goals, objectives, and activities align with their personal values and goals. These are particularly important constructs when considering elective classes like the one examined here. As many discussions have pushed higher education both in practice and in students' minds toward a job preparation model, students often perceive only classes within a major as having a clear contribution to their future aims and goals. This neoliberal shift in university structures has aligned with claims that the humanities, particularly humanities electives such as this class, are irrelevant. I do not suggest that these results indicate students found direct connections between this class and their future careers, but rather

they found something that relates to who they want to be, what they value, or what they want to accomplish.

Student comments, while not dealing explicitly with motivational types, reveal hints of how these constructs influenced their experience. For example, some students wrote about how the feedback, through focusing on responding to reflections rather than grades on individual assignments supported learning in meaningful, and presumably relevant ways: "I think that the learning reflections were really beneficial. I don't believe I've had any classes that have done this in the past. I think it is a great way to hear honest, in-depth feedback about your performance in the course as a whole, instead of just hearing feedback on individual assignments only." Other students wrote about the structure supporting enjoyment and fun, but still in the context of learning (specifically by supporting autonomy): "I truly enjoyed this class. Much like the rock and roll we discussed throughout the semester, this course allows each student to be their own person, and research/present on topics that they find special to themselves, or just interesting. There were general guidelines, but you allowed for creative freedom… And this just isn't something I've encountered very often in my time in college. This made the course refreshing, and actually enjoyable to take (even though it's over the summer!)."

The Informed Learning Scale (table 3) (Flierl, Bonem, and Maybee, n.d.) derives from Informed Learning, a theoretical framework developed by Christine Bruce (2008), seeking to look holistically at student engagement with information within disciplinary contexts as opposed to traditional views of information literacy as a series of distinct skills. While providing students with so much freedom and encouraging them to draw on their existing knowledge and interests, I was curious to see how this might influence their perception regarding information use. This also aligned with another element of course design in which I replaced the textbook with institutional access to Rocksbackpages, a searchable online collection of magazines. While I encouraged students to engage with these materials or find their own for their presentations, I did not require any particular level of engagement with these sources. Instead, descriptions of high grades, as well as feedback and suggestions, highlighted the importance of engaging with information in meaningful ways that reflect the discipline.

Table 3. Informed learning scale

These items were rated on a 7-point Likert scale (1=strongly disagree; 7=strongly agree) n=20				
Informed learning prompt	Min.	Max.	Mean	Std. dev.
I believe it is important for me to carefully evaluate information I use in this course.	3	7	5.87	1.09
I believe I can learn in this course by using information.	5	7	6.27	0.57
For this course, my instructor encourages me to use my prior experiences of using information.	6	7	6.40	0.49
I build upon my previous experiences of using information to learn subject content in this course.	5	7	6.27	0.68
My instructor encourages me to use information for specific purposes.	4	7	6.13	0.88
When I consider my life after college, I feel confident in my ability to learn with information sources.	6	7	6.20	0.40
I feel confident in my ability to synthesize information from different sources.	5	7	6.13	0.50

I feel confident in my ability to use information to learn subject content in this course	5	7	6.13	0.50

Students agreed consistently with the constructs in this scale suggesting that the lack of rigid structures for information usage did not interfere with its importance. In fact, one student suggested that the grading structure motivated them to do research: "I felt that the freedom in grading criteria helped my learning and made me more inclined to research music history for personal interest." Another saw this flexibility as enabling them to develop their information usage. "I really liked the grading in this course. It set me up to try an[d] exceed my own expectations each time we did a review. I think my work effort and ability to use different sources of information continued to grow throughout the course."

While I did not ask or survey about it specifically, students wrote regularly about the value they found in the reflection process, which I found pleasantly surprising compared to my perception that students did not take learning reflections particularly seriously when I have previously incorporated them in traditionally graded classes. For example, above we saw a student comment explaining how this was more beneficial than the traditional practice of feedback limited to the scope of individual assignments. Another student saw the reflections as a way to develop a more meaningful relationship with the course: "The learning reflections were also helpful to see how you were doing and what could be improved. They made the class seem more personal."

FUTURE DIRECTIONS/REFLECTION

In this paper I have explored student perceptions of ungrading in a single online elective course. The results are certainly promising from numerous perspectives, but this should only be the start of more varied explorations about ungrading in general and this course in particular. Looking at future iterations of this course, it would be valuable to delve more deeply into an analysis of student learning through a detailed qualitative analysis of student reflections. This will provide greater insight into what learning looks like for individual students over the course of the semester. It might also reveal how perceptions change. Initially, many students felt confused or concerned by the structure so tracking individuals and how they deal with the unfamiliar and how that contributes to or detracts from their learning would offer rich material for analysis. A deeper analysis of reflections could also lead to ways of demonstrating learning effectiveness for skeptics on the campus and/or accreditors who are generally accustomed to measurement based on standardized tests or clear rubrics. Ultimately, those of us who adopt these methods will be held accountable for providing evidence of student learning and achievement of outcomes and detailed analysis of reflections represents one way to satisfy this need. More broadly, I hope that more approaches to examining student learning, perceptions, and experiences will shed light further on alternate grading structures in a variety of disciplines and contexts. Will students perceive the same levels of learning in courses treated as prerequisites for others, or is the concept limited in potential applications? What about other approaches to non-graded work, and how does the experience of courses that challenge traditional grading structures influence a student's broader education. Is this course perceived as an outlier by students, or do they take some of the reflective skills developed here and apply them elsewhere? Are they more critical of standard grading structures after taking this course? Future projects would benefit from more direct student involvement in the analysis of course materials as well as following their experiences after the course.

ACKNOWLEDGMENTS

Daniel Guberman is a senior instructional developer with the Center for Instructional Excellence and a provost fellow with the Division of Diversity and Inclusion at Purdue University (USA). https://orcid.org/0000-0002-2669-3197.

REFERENCES

Bloom, Benjamin S. 1968. "Learning for Mastery." *Evaluation Comment* 1, no. 2.

Bonem, Emily M., Heather N. Fedesco, and Angelika N. Zissimopoulos. 2020. "What You Do Is Less Important than How You Do It: The Effects of Learning Environment on Student Outcomes." *Learning Environment Research* 23: 27–44. https://doi.org/10.1007/s10984-019-09289-8.

Boyer, Ernest L. 1990. *Scholarship Reconsidered: Priorities of the Professoriate*. New York: The Carnegie Foundation for the Advancement of Teaching.

Brookhart, Susan M., and Thomas R. Guskey. 2019. "Reliability in Grading and Grading Scales." In *What We Know about Grading: What Works, What Doesn't and What's Next*, edited by Thomas R. Guskey and Susan M. Brookhart, 13–31. Alexandria, VA: Ascd.

Bruce, Christine S. 2008. *Informed Learning*. Chicago: Association of College and Research Libraries.

CAST. 2018. "Universal Design for Learning Guidelines Version 2.2." Accessed November 10, 2020. http://udlguidelines.cast.org.

Centra, John. A., and F. Reid Creech. 1976. "The Relationship between Student, Teacher, and Course Characteristics and Student Ratings of Teacher Effectiveness." *ETS Program Report*, 76-1. Princeton: Educational Testing Service.

Chick, Nancy, Lorelli Nowell, and Bartlomiej A. Lenart. 2019. "The Scholarship of Teaching and Learning: A Scoping Review Protocol." *Teaching & Learning Inquiry* 7, no. 2: 186–97. https://doi.org/10.20343/teachlearninqu.7.2.12.

Cook-Sather, Alison, Melanie Bahti, and Anita Ntem. 2019. *Pedagogical Partnerships: A How-To Guide for Faculty, Students, and Academic Developers in Higher Education*. Elon, NC: Elon University Center for Engaged Learning. https://doi.org/10.36284/celelon.oa1.

Cook-Sather, Alison, Catherine Bovill, and Peter Felten. 2014. *Engaging Students as Partners in Learning and Teaching: A Guide for Faculty*. San Francisco: Jossey-Bass.

Divala, Joseph J. 2017. "Interrogating the Civic Role of South African Universities." In *Knowledge and Change in African Universities*, edited by Michael Cross and Amasa Ndofirepi, 145–60. Rotterdam: Sense Publishers. https://doi.org/10.1007/978-94-6300-842-6_9.

Dweck, Carol S. 2006. *Mindset: The New Psychology of Success*. New York: Random House.

Eyler, Joshua R. 2018. *How Humans Learn: The Science and Stories behind Effective College Teaching*. Morgantown, WV: West Virginia University Press.

Feldman, Kenneth A. 1997. "Identifying Exemplary Teachers and Teaching: Evidence from Student Ratings." In *Effective Teaching in Higher Education: Research and Practice*, edited by Raymond P. Perry and John C. Smart, 368–95. New York: Agathon.

Fink, L. Dee. 2002. *Creating Significant Learning Experiences: An Integrated Approach to Designing College Courses*. San Francisco: Jossey-Bass.

Flierl, Michael, Emily Bonem, and Clarence Maybee. n.d. "Developing the Informed Learning Scale: Measuring Information Literacy in Higher Education."

Franke, Matthew. 2018. "Final Exam Weighting as Part of Course Design." *Teaching & Learning Inquiry* 6, no. 1: 91–103. https://doi.org/10.20343/teachlearninqu.6.1.9.

Franke, Wolfgang. 1960. *The Reform and Abolition of the Traditional Chinese Examination System*. Cambridge, MA: Center for East Asian Studies, Harvard University. https://doi.org/10.2307/j.ctt1tg5m5p.

Freeman, Scott, Sarah L. Eddy, Miles McDonough, Michelle K. Smith, Nnadozie Okoroafor, Hannah Jordt, and Mary P. Wenderoth. 2014. "Active Learning Increases Student Performance in Science, Engineering, and

Mathematics." *Proceedings of the National Academy of Sciences* 111, no. 23: 8410–415. https://doi.org/10.1073/pnas.1319030111.

Freire, Paulo. 2005. *Pedagogy of the Oppressed*, translated by M. Bergman Ramos. New York: Continuum. (Original work published 1970).

Gillis, Alanna. 2019. "Reconceptualizing Participation Grading as Skill Building." *Teaching Sociology* 47, no. 1: 10–21. https://doi.org/10.1177/0092055X18798006.

Ginexi, Elizabeth M. 2003. "General Psychology Course Evaluations: Differential Survey Response by Expected Grade." *Teaching of Psychology* 30, no. 3: 248–51.

Glaser, Robert. 1963. "Instructional Technology and the Measurement of Learning Outcomes: Some Questions." *American Psychologist 18*, no. 8: 519.

Guay, Frédéric, Robert J. Vallerand, and Céline Blanchard. 2000. "On the Assessment of Situational Intrinsic and Extrinsic Motivation: The Situational Motivation Scale (SIMS)." *Motivation and Emotion* 24, no. 3: 175–213. https://doi.org/10.1023/A:1005614228250.

Haugnes, Natasha, Hoag Holmgren, and Martin Springborg. 2018. *Meaningful Grading: A Guide for Faculty in the Arts*. Morgantown, WV: West Virginia University Press.

hooks, bell. 1994. *Teaching to Transgress: Education as the Practice of Freedom.* New York: Routledge.

Hutchings, Pat., Mary T. Huber, and Anthony Ciccone. 2011. *The Scholarship of Teaching and Learning Reconsidered: Institutional Integration and Impact.* San Francisco: Jossey-Bass.

Johnson, James E., James A. Jones, Thomas G. Weidner, and Allison K. Manwell. 2019. "Evaluating Academic Rigor, Part II: An Investigation of Student Ratings, Course Grades, and Course Level." *Journal of Assessment and Institutional Effectiveness* 9, no. 1/2: 49–78. https://doi.org/10.5325/jasseinsteffe.9.1-2.0049.

Kiyama, Judy M., Cecilia Rios-Aguilar, and Regina Deil-Amen. 2018. "Funds of Knowledge as a Culturally Responsive Pedagogy in Higher Education." In *Funds of Knowledge in Higher Education: Honoring Students' Cultural Experiences and Resources as Strengths*, edited by Judy Marquez Kiyama and Cecilia Rios-Aguilar, 176–88. New York: Routledge. https://doi.org/10.4324/9781315447322-11.

Ladson-Billings, Gloria. 1995. "Toward a Theory of Culturally Relevant Pedagogy." *American Educational Research Journal* 32, no. 3: 465–91. https://doi.org/10.3102/00028312032003465.

Levesque-Bristol, Chantal, Timothy D. Knapp, and Bradley J. Fisher. 2010. "The Effectiveness of Service-Learning: It's Not Always What You Think." *Journal of Experiential Education* 33, no. 2: 208–24. https://doi.org/10.1177/105382590113300302.

Love, David A., and Matthew J. Kotchen. 2010. "Grades, Course Evaluations, and Academic Incentives." *Eastern Economic Journal* 36, no. 2: 151–63. https://doi.org/10.1057/eej.2009.6

Mercer-Mapstone, Lucy, and Sophia Abbot. 2020. *The Power of Partnership: Students, Staff, and Faculty Revolutionizing Higher Education*. Elon, NC: Elon University Center for Engaged Learning. https://doi.org/10.36284/celelon.oa2.

Moll, Luis C., Cathy Amanti, Deborah Neff, and Norma Gonzalez. 1992. "Funds of Knowledge for Teaching: Using a Qualitative Approach to Connect Homes and Classrooms." *Theory Into Practice* 31, no. 2: 132–41. https://doi.org/10.1080/00405849209543534.

Morris, Sean Michael, and Jesse Stommel. 2018. *An Urgency of Teachers: The Work of Critical Digital Pedagogy*. Lexington, KY: Hybrid Pedagogy Incorporated.

Nilson, Linda B. 2015. *Specifications Grading: Restoring Rigor, Motivating Students, and Saving Faculty Time.* Sterling, VA: Stylus.

Ndlovu-Gatsheni, Sabelo J. 2013. *Coloniality of Power in Postcolonial Africa: Myths of Decolonization*. Dakar: CODESRIA.

Ndofirepi, Amasa, and Michael Cross. 2017. "University Knowledge for Societal Change in Africa: Unpacking Critical Debates." In *Knowledge and Change in African Universities*, edited by Michael Cross and Amasa Ndofirepi, 1–13. Rotterdam: Sense Publishers. https://doi.org/10.1007/978-94-6300-842-6_1.

Ramoupi, Neo L. L., and Roland N. Ntongwe. 2017. "Africanisation of Humanities Knowledge in Universities in Africa: A Critique of the Cameroon and South African Experiences." In *Knowledge and Change in African Universities*, edited by Michael Cross and Amasa Ndofirepi, 195–214. Rotterdam: Sense Publishers. https://doi.org/10.1007/978-94-6300-842-6_12.

Rampell, Catherine. 2010. "Grade Inflation: Your Questions Answered" *The New York Times*. May 13, 2010. https://economix.blogs.nytimes.com/2010/05/13/grade-inflation-your-questions-answered/

Ryan, Richard M., and Edward L. Deci. 2017. *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. New York: The Guilford Press. https://doi.org/10.1521/978.14625/28806.

Schnapp, Alain, and Pierre Vidal-Naquet. 1971. *The French Student Uprising, November 1967-June 1968: An Analytical* Record, translated by Maria Jolas. Boston: Beacon Press. (Original work published 1969).

Seraphin, Sally B., J. Alex Grizzell, Anastasia Kerr-German, Marjorie A. Perkins, Patrick R. Grzanka, and Erin E. Hardin. 2019. "A Conceptual Framework for Non-Disposable Assignments: Inspiring Implementation, Innovation, and Research." *Psychology Learning and Teaching* 18, no. 1: 84–97. https://doi.org/10.1177/1475725718811711.

Smith, Jeffrey K., and Lisa F. Smith. 2019. "Grading in Higher Education." In *What We Know about Grading: What Works, What Doesn't and What's Next*, edited by Thomas R. Guskey and Susan M. Brookhart, 195–213. Alexandria, VA: Ascd.

Sorcinelli, Mary Deane, Ann E. Austin, Pamela L. Eddy, and Andrea L. Beach. 2006. *Creating the Future of Faculty Development: Learning from the Past, Understanding the Present*. Bolton, MA: Anker Press.

Stommel, Jesse. 2020. "How to Ungrade: A FAQ." *Jesse Stommel*. Published, February 6, 2020. https://www.jessestommel.com/ungrading-an-faq/.

Supiano, Beckie. 2019. "Grades Can Hinder Learning. What Should Professors Use Instead?" *Chronicle of Higher Education*, July 19, 2019. https://www.chronicle.com/interactives/20190719_ungrading.

Theobald, Elli J., Mariah J. Hill, Elisa Tran, Sweta Agrawal, E. Nicole Arroyo, Shawn Behling, Nyasha Chambwe, Dianne L. Cintrón, Jacob D. Cooper, Gideon Dunster, Jared A. Grummer, Kelly Hennessey, Jennifer Hsiao, Nicole Iranon, Leonard Jones II, Hannah Jordt, Marlowe Keller, Melissa E. Lacey, Caitlin E. Littlefield, Alexander Lowe, Shannon Newman, Vera Okolo, Savannah Olroyd, Brandon R. Peecook, Sarah B. Pickett, David L. Slager, Itzue W. Caviedes-Solis, Kathryn E. Stanchak, Vasudha Sundaravardan, Camila Valdebenito, Claire R. Williams, Kaitlin Zinsli, and Scott Freeman. 2020. "Active Learning Narrows Achievement Gaps for Underrepresented Students in Undergraduate Science, Technology, Engineering, and Math." *Proceedings of the National Academy of Sciences 117*, no. 12: 6476–483. https://doi.org/10.1073/pnas.1916903117.

Tobin, Thomas J., and Kirsten T. Behling. 2018. *Reach Everyone, Teach Everyone: Universal Design for Learning in Higher Education*. Morgantown, WV: West Virginia University Press.

Wiggins, Grant, and Jay McTighe. 2005. *Understanding by Design*. Alexandria, VA: Ascd.

Williams, Geoffrey C., and Edward L. Deci. 1996. "Internalization of Biopsychosocial Values by Medical Students: A Test of Self-Determination Theory." *Journal of Personality and Social Psychology 70*, 767–79. https://doi.org/10.1037/0022-3514.70.4.767.

Winter, Richard. 1993. "Education or Grading? Arguments for a Non-Subdivided Honours Degree." *Studies in Higher Education* 18, no. 3: 363–77. https://doi.org/10.1080/03075079311331382271.

Wright, Robert E., and John C. Palmer. 2006. "Comparative Analysis of Different Models Explaining the Relationship between Instructor Ratings and Expected Student Grades." *Educational Research Quarterly* 30, no. 2: 3–19.