

Ungrading in the Music Theory Classroom: A One-Year Case Study

BY DAVID ORVEK

In this paper, I present a summary of and critical reflection upon an alternative assessment practice known as “ungrading” that I used during the 2023–24 academic year. My version of the practice involved providing students with feedback without grades and letting students determine their own final grades through a process of guided self-assessment. Many students’ reactions to ungrading were positive. Some reported being able to focus on learning rather than grades, and even skeptical students admitted some benefits.

Ungrading benefited me in various ways as a teacher, but there were challenges as well. Some students struggled with the lack of external accountability that grades provide. I also had issues with attendance and students not submitting work. I do believe that aspects of ungrading can improve students’ attitudes toward school, their learning, and their relationships with their instructors, but there are challenges with the implementation of this idea within education as it is currently practiced in the United States.



Introduction

This paper describes and evaluates an alternative assessment model I used in all my courses during the 2023–24 academic year. Before discussing the model, though, I want to explain the motivation behind it. And toward this end, I begin with a story that illustrates some of the ways in which traditional grades-based assessment fails.

I took a required computer course in college where most of the course grade was determined by a series of quizzes and exams. The quizzes were offered online, used only multiple-choice and true/false questions, and could be taken as many times as one wanted. The questions did not change at all from attempt to attempt, and after each attempt one was told which answers were incorrect. The in-person exams consisted entirely of questions from these online quizzes.

With this knowledge, I developed a strategy to ensure I would get an A in this course. The first time I took each quiz, I would give my best guess on each question. As I went along, I would record the question number and my answer in a notebook. Once I received my score on my first attempt, I would make a note of which questions

I presented an earlier version of this paper at Pedagogy into Practice 2024 at the University of Oklahoma. Special thanks to the two anonymous reviewers for their suggestions that substantially improved this article.

I had gotten wrong. Then I would take the quiz again, making sure to give the same answers for each question except those I had gotten wrong. After a few iterations of this, I could without fail achieve 100% on each quiz. And then all I had to do to prepare for the exams was memorize these quiz questions. I did not even have to actually understand the concepts covered in these assessments; all I needed to know was that the answer to Question X was C, and so on. My method worked, and I earned an A in the course, but I learned pretty much nothing in the process. Several years later I actually did develop an interest in computers, but nothing I “learned” in that course had stuck.

This story illustrates two problems with traditional grading models. First, grades determined by points-based systems do not assess learning. I got an A in that computer course, but my grade had nothing to do with my knowledge of the course content. Second, grades cause students to focus on the wrong things. If I had cared about learning anything in that computer course, I would have tried to understand concepts rather than just memorizing answers. The issue was partly that I did not care about the course content at the time, but more importantly, I cared more about my grade than learning. Another contributing factor was the pressure of school. I seem to remember not being particularly happy about how I was getting through the course, but I did not feel that studying to actually learn was something I had the bandwidth for that semester.

The failure of grades to measure learning is something I am sure all teachers have experienced. If you are a teacher, just think of the students you have had in your courses that have ostensibly passed the prerequisites but seem to be completely unprepared. Or think of students you have had in your courses that you know learned a lot but who were not able to pass because of some reason not related to their learning. Part of the reason this happens is because, as mathematicians David Clark and Robert Talbert point out, the way final grades are usually calculated in traditional points-based models is not statistically meaningful. Specifically, they explain that data can either be numerical (the result of a measurement or count) or categorical (names, labels, etc.). Grades are categorical data. And while averages of numerical data are meaningful, averages of categorical data are not. As they put it, “Computing averages of the categorical data called *grades* makes no more sense than averaging ZIP codes.”¹

My story about how grades negatively impacted my motivation to learn is not unique either. Surely many instructors can relate to my anecdote, either from experience as students or interactions with students. This is not to say that students

1 Clark and Talbert (2023, 20).

do not care about learning. Some probably do not, but I would bet that most are just spread so thin that they feel they cannot care about learning as much as they have to care about their grades. And who can blame them with all the pressure that our society puts on grades? I remember hearing a biology professor at my undergraduate institution tell the pre-med majors in the class something like: “If you don’t get As in all your science classes, you can forget about going to medical school.”

There is also an enormous body of research regarding the effects of grades on learning that supports the conclusions from my story. Here is Alfie Kohn’s concise summary of this literature:

[W]hen students from elementary school to college who are led to focus on grades are compared with those who aren’t, the results support three robust conclusions: Grades tend to diminish students’ interest in whatever they’re learning. . . . Grades create a preference for the easiest possible task. . . . Grades tend to reduce the quality of students’ thinking.²

All of this was exactly true for me in that computer course.

I hope it is clear at this point that assessment is not a neutral practice. If we as teachers are concerned with helping students learn, it is time to stop taking the way we assess students for granted.

I began my journey into alternative assessment in the summer of 2023. By August, I decided that I was going to move as far away from traditional grading practices as I could. In the remainder of this paper, I will explain the alternative assessment model I used during the 2023–24 academic year, share some results about how it went, and discuss my thoughts about what I will do in the future.

The Alternative Assessment Model

Teaching Context

Before explaining the alternative assessment model I used, let me share some information about my teaching context. I teach at Butler University in Indianapolis, Indiana. Butler is a small, private, liberal arts university with about 5,500 students, of which about 150 are music majors. Most music students are undergraduates majoring in music education or performance. We also have a small graduate program that offers master’s degrees. I am one of two full-time music theory faculty, and I have total control over the content and policies of my courses. My courses generally range in size from about 10 to 25 students, and I do not have any teaching assistants to help with grading.

² Kohn (2013, 144).

Example 1 lists the courses I taught during the 2023–24 academic year. Note that these are all exclusively written-theory courses; none had any aural skills or keyboard components.³ For simplicity, I used essentially the same assessment model in each course, though I did make some small changes between the Fall and Spring semesters discussed below.

Term	Title	Level	# of Students
Fall 2023	Music Theory 3	Undergraduate core theory	34 total (2 sections)
Fall 2023	Form and Analysis	Upper-level undergraduate	14
Fall 2023	Graduate Theory Review	Remedial graduate	5
Spring 2024	Music Theory 2	Undergraduate core theory	44 total (2 sections)
Spring 2024	Music Theory 4	Undergraduate core theory	13
Spring 2024	Analysis of Tonal Music	Graduate	14

Example 1

List of courses I taught during the 2023–24 academic year.

The Core Components of the Assessment Model

I had two goals for my new assessment model: (1) for students to be able to assess their learning, and (2) for students to forget about grades and focus on learning for its own sake. That is, I wanted students and myself to be able to know whether learning was in fact taking place, and, if not, to understand what sort of steps we might take to see that it did. And I wanted students to be doing the work of learning out of a true desire to learn something that they wanted to learn. I *did not* want students to come to class and submit assignments because they were being forced to or in order to earn a grade. But, because my institution requires final grades for all courses, I still needed a way of determining those grades.

³ I did not teach any aural skills courses during the academic year discussed in this paper, nor have I used anything like the assessment model described in this paper in an aural skills course. I do believe that feedback without grades and self-assessment can be beneficial in any context, but I imagine some careful thought would be needed to integrate these practices into an aural skills or other skills-based course. One way that one could implement these ideas into an aural skills course would be to limit the grade categories on assessments to “pass” and “retry” and ask students to write short self-assessments on how they think they did on dictation exams and/or hearings. See Alegant (2018) for an example of minimal grade categories. I am currently (Fall 2024) experimenting with an alternative assessment model in an aural skills course based on the principles of standards-based grading. But since this is an entirely different assessment model, it is beyond the scope of this article to report any information on this course here. I hope to publish an article in the future that reports on my experience with this course.

Based on my research into alternative assessment practices, I decided to build my assessment model around gradeless feedback and self-assessment. Models like this have been called “ungrading.”⁴ What I came up with is heavily indebted to practices outside the field of music theory described in Sackstein (2015), Blum (2020), and Stommel (2023). While there has been discussion of non-traditional grading practices like standards-based grading and mastery learning in music theory pedagogy, to my knowledge the work of Brian Alegant and Barbara Sawhill is the only published discussion of an assessment model using minimal grading and self-assessment in the context of music theory, and a very brief one at that.⁵

As Stommel and others note, ungrading is not a single unified practice. In Stommel’s words, ungrading is the act of “raising an eyebrow at grades as a systemic practice, distinct from simply ‘not grading.’”⁶ Sean Michael Morris adds:

- Ungrading is a humanizing approach to education.
- Ungrading is not about grades, but about people.
- Ungrading is more than not-grading, it’s an approach that informs and is informed by our whole pedagogical practice. It is an “if this, then what else” *modus operandi*, and something we must inhabit fully when we teach or prepare to teach.
- Ungrading does not play at being objective, but rather acknowledges the utter subjectivity of learning.
- Ungrading asks us to question the notions of teaching content, of expertise, of the location of knowledge in the room, which means that ungrading is inherently more equitable than grading.⁷

Day to day, my courses ran similarly to courses I have taught with traditional grading models. I still asked students to complete assignments, projects, and occasional quizzes. But instead of grading these assessments with numbers or letters, I provided detailed written feedback only. I set deadlines for assignments and projects, but there was no penalty for submitting them late. I also encouraged students to revise and resubmit assignments. Finally, I did not grade attendance or participation. There were

4 Terminology for alternative assessment models is not consistent and is constantly evolving. In some cases, people use the same term to describe different models and different terms to describe essentially the same models. Recently, some have begun arguing that “ungrading” is not a useful term. See discussion in Clark (2023).

5 See Alegant and Sawhill (2013), Duker et al. (2015); Johnson (2015); and Michaelsen (2019).

6 Stommel (2023).

7 Morris (2021).

thus no points or letter grades attached to anything in my courses. The gradebook in Canvas, my institution's learning management system, had assignments and projects in it, but the only markings were complete and incomplete. To determine final grades at the end of the semester, students chose the grades that they felt best represented their learning through a process of self-assessment.

Feedback

The kind and amount of feedback I gave on assignments and projects depended on the nature of the submission. The way that I gave feedback was not fundamentally different from how I gave feedback in traditional assessment models, but I did find myself giving much more feedback since there were no grades to communicate my sense of the quality of the work. Generally, I would make note of mistakes or issues to address in annotations on assignments and then provide a broader summative comment in which I would make detailed suggestions as to how the submission could be improved in a subsequent revision. Sometimes these comments would be as simple as "look at this chord again." Sometimes I might give a hint about a different way of looking at something. At other times, I might leave a more detailed comment explaining why something was not right or why something different would work better. Then in longer comments, I would provide a broader summary of the kinds of issues I was seeing in the submission and what the student ought to address in a revision of the assignment.

I emphasize the importance of feedback in my assessment model because I hope to dispel two impressions about ungrading: first, that ungrading means the teacher does not provide students with feedback (i.e., "ungrading" means "no grading"); and second, that students in an ungraded course have no idea whether they are learning. As I mentioned above, I spent much more time reviewing students' work in my ungraded courses than I did in previous courses.⁸ With larger projects—especially projects students designed themselves—I could easily spend half an hour or more writing up

8 I do not intend to suggest that most instructors do not give detailed feedback (although I have certainly received my fair share of grades without feedback on papers I have submitted as final projects, especially in graduate seminars). However, my experience as a student and a graduate assistant suggests that most instructors do not give undergraduate students the opportunity to revise and resubmit most assessments. In the absence of such an opportunity, students rarely internalize the feedback they receive, even when later work builds directly on the current assignment. I am sure most instructors have had the experience of leaving detailed comments on students' work, only to find the same issues in subsequent assessments. Or, even worse, passing back assessments and watching students quickly glance at their grade and then toss the assessment in the trash on their way out the door. For a discussion of the complex relationship between grades and feedback, see Guskey (2019).

my thoughts on a submission. While this work is hard and incredibly time consuming, it is work that feels worthwhile to me because it is directed at meaningful learning. And students' work almost always improves when they take the time to complete the recommended revisions.

Syllabus Statements

I explained my alternative assessment model to students in the syllabus and on the first day of class. The statement on assessment and grades I used in my syllabi in Fall 2023 appears in Example 2 and the statement for Spring 2024 appears in Example 3.

Research shows that students who focus on grades are less interested in their work, tend to choose whatever task is easiest, and overall do not learn as well as students who are intrinsically motivated to learn (see Alfie Kohn, "The Case against Grades"). In order to minimize the negative impact of grades on your learning in this course, I will not be assigning grades to any of your work, and you will choose your own final grade through a process of self-assessment. However, I do reserve the right to change your final grade if your self-assessment obviously wasn't done in good faith. Your colleagues and I will be providing you with feedback on your work, but this feedback will be qualitative rather than quantitative. My goal is for this grading policy to take grades and all stress associated with them off your mind. If this proves not to be the case for you, please let me know so that we can discuss alternative assessment options for you.

Example 2

Syllabus statement on assessment and grades for Fall 2023.⁹

There is a substantial body of research that shows that grades harm learning. Specifically, grades "tend to diminish students' interest in whatever they're learning," "create a preference for the easiest possible task," and "tend to reduce the quality of students' thinking." Additionally, grades do not actually do a good job of measuring learning, nor are they particularly objective. While it is necessary for you and me to be able to assess your learning in this course, my goal is to do so without negatively impacting your learning in the process.

Here is the system I propose we use this semester: rather than assigning grades to your work, I'll provide you with qualitative feedback. Then, at the end of the semester, you'll compile a portfolio of your best work that shows you've achieved the course learning outcomes and choose the final grade (see below) that best represents your learning through a process of guided self-assessment. My goal is for this assessment practice to take grades off your mind and to allow you to focus on *learning* rather than achievement. If this proves not to be the case for you, please let me know so that we can discuss alternative assessment options for you.

Example 3

Syllabus statement on assessment and grades for Spring 2024.¹⁰

⁹ The source cited in the example is Kohn (2013).

¹⁰ The quote in the example comes from Kohn (2013).

Example 4 reproduces the self-assessment I had students complete at the end of the Spring 2024 semester.¹¹

1. Student Learning Outcome X for this course states that, by the end of the course, students will be able to: "... **Explain in detail how you have achieved this outcome, ideally by citing specific assignments/projects/in-class work, etc.**
2. What are the highlights of your learning in this course? That is, what are things you've learned that have stuck out to you or that are particularly meaningful to you? Think about skills as well as knowledge.
3. Are there any areas in which your learning and/or work for this course didn't meet your personal goals? What could you have done differently to have met those goals, and how can you continue to work toward improving in the future?
4. Based on your answers to the previous prompts, **use the grade guidelines in the syllabus to choose the letter grade on the A-F scale that best represents your learning in this course and explain why.** This will be your official final grade for the course.
5. **Optional:** Do you have any thoughts about this course, the grading system, or anything else that you'd like to share?

Example 4

The final self-assessment students completed in my courses in Spring 2024.

Changes from Fall to Spring

Self-assessment is the area where I made the most changes between Fall 2023 and Spring 2024 semesters. In the fall, students began the semester by filling out a "Learning Goals Questionnaire" that asked them to think about what they wanted to get out of the course based on their personal goals related to music. My goal was to get students thinking about the relevance of the course for them and to provide them with a means of assessing their progress over the course of the semester. I do think this assignment initially provided a bit more buy-in, but I found that students tended to set learning goals that were either too vague, too specific, or not directly related to the course. Choosing ineffective learning goals made it difficult for some students to assess their progress toward their goals at the end of the semester.

In the spring, I opted to drop the learning goals questionnaire for my undergraduate courses and instead tried to foster buy-in by having frequent discussions in class about the relevance of what we were learning. This seemed to make students more engaged in they're learning and still left space for students to think about how things were meaningful for them without having to navigate the complexities of having students set and assess realistic goals.

¹¹ This self-assessment was largely the same as the one I used in Fall 2023, with only a few tweaks in phrasing and the order of the questions. Question 1 in this assessment was actually a series of questions, each directed at one of the course learning outcomes.

In my graduate-level course in the spring semester (Analysis of Tonal Music), I used a modified version of the idea behind the learning goals questionnaire, where I set vague learning outcomes for the class as a whole and then asked each student to either personalize those outcomes or add additional outcomes. This worked better because the goals students came up with were more specific. I also intentionally left the content of this particular course somewhat vague at the beginning of the semester and then chose topics that aligned with student goals. Still, though, by the end of the semester we had not addressed some goals. Ultimately, I like the idea of having students set their own goals when the nature of the course allows it, but I need to do a lot more thinking and careful planning to make this work as well as it ought to.

The other change I made from fall to spring was offering more guidance on how to choose final grades. In the fall, I did not provide any guidance on how to choose a grade. I simply asked students to reflect on the progress they had made toward the course learning outcomes and their own personal learning goals, and then asked them to choose the letter grade that they felt best represented their learning in the course. Some students understandably found this difficult. In the spring semester, I tried a few adjustments to address this difficulty. My institution does not publish official criteria or descriptions of each letter grade, so I wrote up a generic set of criteria for each letter grade that focused on learning. These criteria are reproduced in Example 5. These guidelines are not particularly profound, but students seemed to find them helpful. Only a few students had a hard time choosing a grade at the end of the spring semester.

Butler University does not publish official descriptions of each letter grade. Here are some general guidelines you can use in determining your grade:

- D–F: Unsatisfactory; you have not achieved one or more of the course learning outcomes and/or you do not have a passable understanding of and facility with all the course material and relevant skills.
- C: Satisfactory; you have achieved all the course learning outcomes and have a passable understanding of and facility with all the course material and relevant skills.
- B: Good; you have achieved all the course learning outcomes and have a solid understanding of and facility with all the course material and relevant skills.
- A: Excellent; you have achieved all the course learning outcomes and have mastered all the course material and relevant skills.

Example 5
Grade guidelines for my courses in Spring 2024.¹²

¹² In the future, I would avoid the word “mastered” because of its negative connotations. For a discussion of reasons to avoid this word, see Talbert (2021).

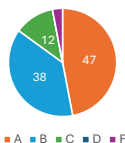
Results: Grades, Attendance, and Submissions

In this section I report on the outcomes of my experiment with ungrading. I begin with some rough data about final grades, attendance, and submissions, and then turn to thoughts from students and my own experience.

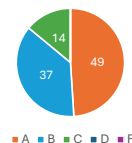
Grades

Teachers using self-assessment for final grades are often asked something like: “If students choose their final grades, will they all give themselves As?” In short, no. Example 6 provides data about grade distributions from my courses to support this. These charts show the grade distributions for all my courses during the 2023–24 academic year. Note that I have aggregated all varieties of grades with the same letter. For example, the top-left corner of the chart shows that 47% of final grades in Theory 3 were some kind of A, 38% were some kind of B, 12% were some kind of C, and about 3% were Fs. While it is true that As are the majority of the grades in all courses, some students did give themselves Bs and Cs. All of the Fs were given to students who submitted no work and had stopped attending class by the end of the semester.

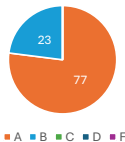
% of Grades | Theory 3 | Fall 2023



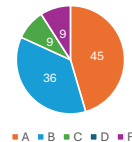
% of Grades | Theory 2 | Spring 2024



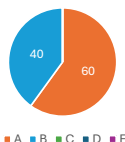
% of Grades | Form | Fall 2023



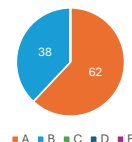
% of Grades | Theory 4 | Spring 2024



% of Grades | Grad Review | Fall 2023



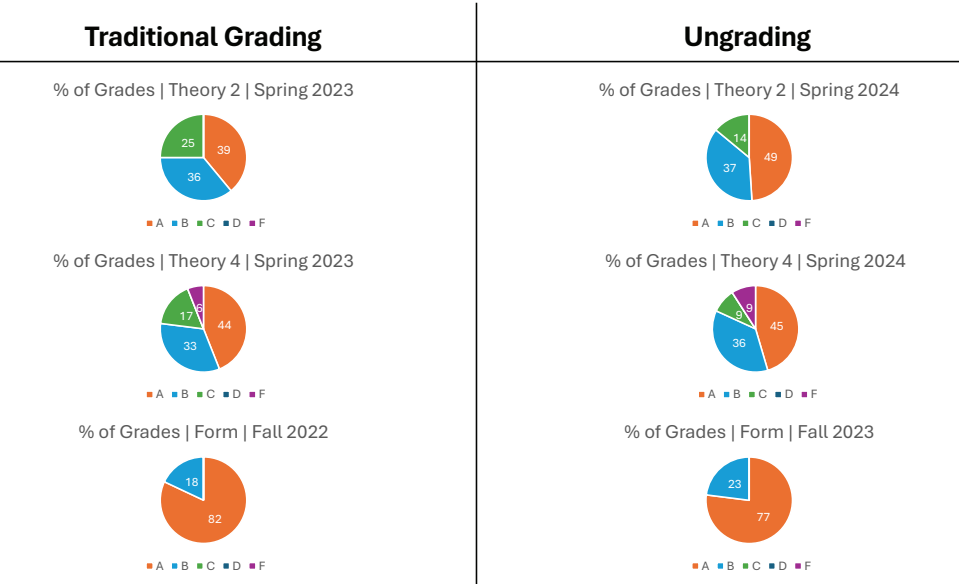
% of Grades | Tonal Analysis | Spring 2024



Example 6

Final grade distributions for all my courses during the 2023–2024 academic year when I used an alternative assessment model. +/- grades are lumped in with whole letter grades.

Even if every student did not give themselves an A, one might still think that the grade distribution seems a bit high. I initially thought so too. For context, Example 7 compares the grade distributions for three courses using my alternative assessment model from the 2023–24 academic year with the distributions from the year before when I used a traditional points-based grading model. Although the distributions are not identical, I am struck by just how similar they are. In my own courses at least, it seems that student-chosen grades were not significantly higher than grades calculated from points-based grading models. I do not know of any rigorous studies that have explored such comparisons with larger sample sizes, but ultimately, I am not sure the comparison is important, because I do not think that traditionally assessed grades measure learning.



Example 7
A comparison of final grade distributions for courses using a traditional points-based grading model (left column) and those same courses using my alternative assessment model (right column).

Attendance

I do not have any attendance records from previous years to compare with 2023–24, and I have not usually implemented formal attendance policies in my courses even when I used point-based grading systems. My intuitive sense is that attendance in 2023–24 was not much worse than previous years. There were definitely some

exceptions, however. Most of my courses had a few students who missed a significant number of classes. Theory 4 (post-tonal theory) had the most attendance issues. Only about half of the students attended consistently, and there was one day where *no one* showed up until 5 minutes after the start of class. I will say, though, that I heard from several colleagues that attendance was a problem in their courses too.

Submissions

Example 8 compares the number of missing submissions in courses where I used a traditional points-based grading model versus those same courses where I used my alternative assessment model. Courses where I used my alternative assessment model had more missing submissions, with Theory 4 being the most striking.

Traditional Grading	Ungrading
Theory 2: 13%	Theory 2: 19%
Theory 4: 14%	Theory 4: 38%
Form and Analysis: 5%	Form and Analysis: 10%

Example 8

Percentage of possible submissions not submitted for courses using a traditional points-based grading model (left column) and those same courses using my alternative assessment model (right column).

While I do not have any data on the timeliness of submissions from 2023–24 versus previous years, my sense is that there were *far* more late submissions during the 2023–24 academic year. For example, some students did not turn in midterm projects until the end of the semester. Theory 4 was particularly bad in this regard as well: *none* of the 13 students submitted their final projects on time.

I now turn to some qualitative commentary derived from student comments and my own experience. My sense of students' thoughts about the assessment model comes from comments in their final self-assessments and my course evaluations.

Student Comments

Of those students who made comments about assessment, most thought the assessment model had positive aspects. The most common positives students cited were that it helped them focus on learning, fostered a growth mindset, caused them to learn more deeply, and created a less stressful environment that was more conducive to learning; they also appreciated the flexibility it allowed.

Some students definitely did not like the assessment model. Some felt that it simply did not work for them; others thought it had potential but needed some adjustment. By far the most common complaint was that students felt that the lack of grades made it difficult to complete the assigned work. Students cited a few reasons for why they thought this was the case. Some students said that they really needed the extrinsic motivation to get themselves started and to get things in on time. That is, without any kind of penalty for not doing the work, they found it hard to find the motivation to do it. Others felt similarly, but specifically cited their busy schedules as a problem. That is, because their schedules were absolutely full and all their other courses had strict deadlines that were enforced with grades, they opted, consciously or not, to put work for my course on the back burner. Sometimes this only meant a delay of a few days until they had a gap in their schedule; at other times this meant they never did the work for my course because they never had time to do it, or it had been so long they forgot about it. A few students also mentioned that the assessment model made them anxious. I did not get the sense that they did not trust me to give them the grade they chose, but rather that they did not have a strong sense of how they were doing in the course day-to-day. On a related note, a few students also mentioned that they found it difficult to assess themselves, specifically when it came to choosing a grade for themselves at the end of the semester. Some students mentioned initially feeling anxious about the assessment model, but then feeling better about it as the semester went on.

My Thoughts

From my perspective, my assessment model was a mixed success. I was delighted to hear that it made some students less stressed and helped them learn. I also saw positive changes in my relationships with students, my own feelings toward the work of teaching, and the way I thought about running courses.

I felt more like a coach to students than the adversary I had often felt like before. In the past, I often felt distant from students because of the amount of power I held in the form of grades. When students were involved in determining their final grade, I felt that this dynamic changed. I of course still had a role to play in whether they learned, but ultimately the responsibility was theirs. I was there to help them, not judge them.

I also felt that my work as a teacher had more meaning for me. By the end of the 2022–23 academic year, I was burned out and had enough doubts about the value of my work that I was seriously thinking of leaving academia. While there were hard

days in the 2023–24 academic year and I continue to have ethical concerns about academia, I did not feel that I was actively harming my students when I was using my alternative assessment model.

I also felt I could be much more flexible in how I designed and implemented my courses. I did not have to have rubrics for every assessment, so I could let students design their own projects and have options for how they completed assignments. While my courses had learning objectives, I did not have to always think in terms of how I would assess them. I also found that ungrading made it easy to accommodate all the complexities of real life during the semester. I did not have to make any accommodations for students with documented disabilities because almost every aspect of my courses was inherently flexible.

Finally, there were students who were able to pass my courses who likely would have failed a course that used a traditional assessment model. Attendance and/or assignment submission was an issue for all of these students. But in each case, I was confident that they had met the learning outcomes based on the work they were able to submit and their self-assessments. Could these students have learned more from attending more classes and completing more assignments? Certainly. But ultimately, the whole point of assessment is to see whether learning is happening, and it did in these cases.

Now for some things that were challenging for me. First, there were the issues with attendance. I am not worried about students missing a day here or there, but when students miss several classes, their learning is obviously impacted. Moreover, absences impact those who do attend class because it makes discussions harder and there are fewer perspectives to learn from. It is also demoralizing for me and other students when half of the class is absent.

Another issue was the number of late or missing submissions. Personally, I do not think it is a problem if students submit assignments or projects a few days late. If anything, it is less overwhelming for me when submissions are somewhat spread out rather than arriving all at once. I also do not think it is necessarily a problem if some students do not submit every assignment. But I do need to address students' self-identified issue of wanting to do the work but not having the bandwidth and/or motivation to do it without a grade.

Lastly, particularly at the end of the spring semester, there were some students who I was not sure had met the course learning outcomes, which led to some tricky situations and self-doubt on my part. In particular, some students did not submit much work, and what they had done was not particularly good; however, they gave

themselves passing grades in their final self-assessments. This presented me with a challenge: do I trust the students and give them the grade they chose, or is it my job to intervene? I was in a lot of turmoil over this. In cases like this, Sackstein and others recommend having frank conversations with students.¹³ So, I met briefly with each of these students, expressed my concerns, and asked honestly whether they felt they had enough of a handle on the material to pass the course. Each student said they felt conversant with the material and committed to doing some work over the summer to review, so I gave them all the passing grades they had chosen.

Some readers might be thinking: “Of course those students said they were ready to move on! Who is going to hold themselves back?” I am worried about this too. But I believe that kind of mistrust is a byproduct of the grading mindset. An ungraded course is different. In my experience, ungrading creates an environment where students are honest about their shortcomings. I have been amazed by students’ frankness and honesty in their self-assessments. And so maybe I am naïve, but I would like to think that a student in an ungraded course would give themselves a failing grade if they were not ready to move on.

Conclusions and Future Directions

After finishing this one-year experiment with ungrading, I spent the summer of 2024 reflecting on what I learned and thinking about what I would do differently in the future. Writing this article was a big part of that reflection process. I wish I could say that I had arrived at a definite conclusion as to whether ungrading “works,” or that I discovered the perfect way to tweak my model in order to fix all the problems I encountered. Alas, by the final weeks of summer, I still had no idea what I would do for the Fall 2024 semester. Even now, as I make the final round of revisions to this article in the last weeks of the Fall 2024 semester, I remain ambivalent about ungrading. On the one hand, I still believe that traditional points-based grading models fundamentally do not assess learning and can actively harm it. Every time I read the work of progressive educators like Kohn, Stommel, and Morris, I am reminded of the harms of grading, and remain inspired to keep trying different things. On the other hand, if I trust students to assess themselves, then I ought to trust the students who said that my assessment model did not work for them. And, of course, there were the many other issues I described above.

¹³ Sackstein (2015, 109–19).

Ultimately, for the Fall 2024 semester I decided to continue using an alternative assessment model based on gradeless feedback and self-assessment but made some changes that I hoped would address some of the challenges I faced during the 2023–24 academic year. I had not yet finished the Fall 2024 semester when I submitted the final version of this article, and so I cannot provide a comprehensive report on how it went. However, in the interest of sparking ideas for those interested in alternative assessment models, I provide a brief summary of the changes I made and my thoughts about their success in the section that follows.

Changes in Fall 2024

The syllabus statement on assessment and grades I used in Fall 2024 appear in Example 9, and the grade criteria for one of my courses during the Fall 2024 semester appear in Example 10. In order to address the problem of borderline cases at the end of the semester, I added a list of essential tasks that students had to complete in order to earn a C (the minimum passing grade for music majors) and one-on-one conferences with me at the end of the semester to collaboratively decide on final grades. The essential tasks for passing included things like completing all self-assessments in good faith and submitting a final project that meets a minimum set of requirements. In all my courses, the final project is designed to address all the course learning outcomes. Theoretically, a student who completed their final project and submitted the self-assessment assignments should have met the course learning outcomes.

My thinking behind the end-of-semester conferences is that they will allow me to dialogue with students about their grades, which I hope will help students and me feel better about their final grades. In order to prepare for these conferences, I am asking students explicitly to make a case for the final grade that they choose for themselves in their final self-assessment. The prompt for this self-assessment is shown in Example 11. The conferences will give me an opportunity to follow up on students' self-assessments, ask for more details about their learning and why they chose the grade they did, and/or to raise concerns if I feel that their chosen grade is unfair for some reason. By the end of the conference, students and I will have collaboratively chosen the final grade that we both think is fair.

Feedback is an essential part of the learning process. You need some way of knowing if you're making progress and what you still need to work on. This is the point of assessment. You don't need letter grades, points, or percentages to do this. Moreover, studies show that grades harm learning in several ways. As much as possible, I'd like to separate assessment and learning from grades. My goal is for you to forget about grades entirely and focus on learning.

For the time being, though, our entire educational apparatus is built around grades, so we need an equitable way for you to earn a final grade in this course that won't hinder your learning. Below is the system I propose we use for assessment and grades this semester. I'm happy to discuss any concerns you have about this system and to figure out ways we could tweak it to work better for you.

You and I will assess your learning primarily in two ways:

1. Gradeless feedback. I'll give you qualitative feedback on all major assignments. There won't be any points or grades attached to this feedback. In some cases, we'll collaboratively come up with a set of "specifications" for what a successful submission looks like. In these cases, I'll also mark each submission "complete" or "incomplete" according to whether it has met all the specifications. You are encouraged to revise and resubmit all work.
2. Self-assessment. You'll list some goals at the beginning of the semester and complete guided self-assessments at three points in the semester (weeks 5, 10, and 17) focusing on your personal growth and progress toward the learning outcomes. These will help you know if you're on track and identify changes you can make moving forward.

You and I will determine your final grade collaboratively in a short conference during finals week. The point of the conference is not for you to convince me to give you a particular grade. Rather, these conferences are intended to be a conversation about your learning that will help you determine the grade that best represents your learning in the course.

To earn a C, you simply need to complete a list of tasks. By completing those tasks successfully, you will have achieved the course learning outcomes at a passable level. Grades higher than C depend on your self-assessment of your learning. You can make a case for a grade higher than C in your final self-assessment, which you and I will discuss in your conference. If you've met the requirements for a C and can present evidence of learning that you think warrants a higher grade, you'll receive that grade. I reserve the right to suggest a different grade, but you will ultimately receive the grade that we both agree on during your conference.

Example 9

Syllabus statement on assessment and grades I used in Fall 2024.

Grade	Requirements
A	1. All the requirements for C 2. Evidence of deep learning and excellent work
B	1. All the requirements for C 2. Evidence that you went beyond the bare minimum
C	1. Provide feedback to a classmate on <i>each</i> stage of the Exhibition Project 2. Submit a final draft of your poster that meets the specifications 3. Present your poster 4. Submit goals and 3 self-assessments in good faith
D	Evidence of some progress toward the requirements for C
F	No evidence of progress toward the requirements for C

Example 10

Grade criteria for one of my courses in Fall 2024.

Write me a letter in a word document in which you do the following:

1. Tell me about your learning and growth related to music theory and music analysis this semester and how this has been and/or will be relevant to you
2. Explain how you have met each of the course learning outcomes listed below with concrete examples
3. Choose your final letter grade for the course and make a compelling case for why you believe this grade is appropriate based on the quality of your learning and work in the course (see the criteria for final grades in the syllabus)

Example 11

Prompts for the final self-assessment in Fall 2024.

To address the issues that I encountered with missing and late submissions, I added more frequent check-in points for projects that involved peer-assessment. On the due date for each project check-in, students were required to come to class to meet with assigned peer-review groups and give each other feedback on their drafts. If they were not able to attend class on these days, students were required to track down each of their group members and give them feedback outside of class. My hope was that these peer-review days would add a sense of accountability for getting work done on time. As an added bonus, these peer-review sessions should help students develop the valuable skill of providing helpful feedback and also give them an opportunity to see their peers' projects.

Submission rates for project drafts seemed better, although I cannot say for sure that the peer-review sessions were the cause. One piece of evidence that suggests that peer reviews helped is that submissions remained low for smaller assignments without a required peer-review component. Even with the pressure of the peer

reviews, though, some students still never submitted any project drafts, but this also happened with a few students when I used point-based grading systems. I will say that attendance at the peer-review sessions dwindled over the course of the semester; I imagine that many students simply opted to stay home and give their peers feedback asynchronously. Not allowing asynchronous feedback as an option might have helped with attendance at these peer-review sessions.

I still did not grade attendance or participation in any of my courses in Fall 2024, and I continued to have issues with attendance. This was most pronounced in my 8 a.m. course; courses that met later in the day generally had good attendance. In all of my courses, though, a few students missed a significant number of class periods. While attendance is not specifically one of the criteria for a final grade, I will address attendance issues in the final grade conferences.

At the beginning of the semester, I also asked students to decide as a group whether they wanted to include things like attendance and assignment submissions as a part of the final grade criteria. I dedicated one class period at the beginning of the semester to collaboratively creating some course policies. One such policy had to do with late or missing work. I specifically asked students if they wanted missing or late work to impact final grades in some way. Some students did, while others did not. Ultimately, most of my classes decided that assignment submission *could* be a factor for grades higher than C but did not specify a minimum number of submissions for a C. That is, students must discuss the number of assignments they completed and whether they submitted them on time in order to make a case for a grade higher than a C, but there are no specific numerical requirements for the number of assignments that must be submitted on time for each letter grade.

While the adjustments to the assessment model I made in Fall 2024 have helped in some respects, I continued to have many of the same issues that I did during the 2023–24 academic year, and I am still not happy with the model. Attendance and submission rates for smaller assignments in particular became a problem in my course that met at 8 a.m. At one point in the semester, I shared my frustrations with students and asked them for their thoughts on what I could do to address these issues. A common suggestion was to grade attendance in some way. While I value students' suggestions, it is important to keep in mind that this suggestion came from the students who *were* in class that day and attended class consistently. I would thus also need to gather input from students who *did not* attend regularly. For those students for whom attendance was an issue, I will ask them what would have helped them attend class consistently during their grade conference.

Closing Thoughts

As of this writing, I have not yet decided what kind of assessment model I will use in the Spring 2025 semester. In this section, I outline some ideas I am considering.

One option is to continue to tweak the current model—for example, to continue to use gradeless feedback and self-assessment, but add more specific requirements for final grades. One of the reviewers for this article had helpful suggestions in this regard. For instance, to address missing work, I could impose hard deadlines for assigned work, after which students cannot receive feedback and/or count this work toward their final grade. Or, alternatively, I could continue to add more specific requirements for each grade, specifying the maximum number of assignments and/or class periods students can miss and still earn each grade.

Adding more specific requirements for grades in an ungrading model is something I have considered before, but I have at least one concern: the limits one chooses for missing assignments and/or absences will necessarily be arbitrary. Say, for example, that I allow no more than three absences for an A. Is the amount of learning lost by a student that misses four classes significant enough that they ought to earn a lower grade than someone who missed three classes? Put differently, is one class period worth one third of a letter grade? And what if the student who missed four classes did great work? What if their work was better than that of the student who only missed three classes? Part of the benefit of an ungrading model without criteria for final grades is that it allows flexibility for these kinds of nuances. It seems to me that the more specific the requirements for grades become, the more likely it becomes that grades will measure things that are *proxies* for learning. That is, grades stand for the number of classes students attended or the amount of work they did, rather than how much they have learned or grown. This is not necessarily a problem. Indeed, at some point a model like this essentially becomes a form of contract grading.¹⁴ But this shift from assessing learning to assessing tasks completed is something that ought to be articulated clearly to students and decided upon intentionally.

Other alternative assessment models include (among others) standards-based grading and specifications grading.¹⁵ While it is beyond the scope of this article to

¹⁴ As with most forms of alternative assessment, contract grading is a difficult thing to define precisely. It generally involves some sort of list of tasks that students must complete in order to earn a specific final grade. For more on contract grading, see Inoue (2022, 2023) and Melzer et al. (n.d.).

¹⁵ Again, both of these practices vary substantially from person to person. Specifications grading usually involves paring grades with bundles of projects/assignments/etc. that must be completed to a

provide a full consideration of the merits of these and other alternative assessment models, I do think that contract grading, specifications grading, and standards-based models improve upon traditional assessment models in various ways. These models are good compromises and are probably the most realistic options for alternative assessment within the current educational establishment. But, it seems to me that none of these models addresses the deeper concerns raised by Kohn, Stommel, Morris, and many others about the effects that grading can have on motivation, learning, students' ability to assess themselves and be self-directed learners, and so on. Students in a course using contract grading might do a lot of work, but not learn anything. Students in a course using standards-based grading can still lose sight of learning in pursuit of meeting minimum standards for a particular grade. If we are truly concerned about helping students learn and if we believe that the reasons people do things matter, then these are ideas we ought to be concerned about.

All of this leads me, in conclusion, to the possibility of a more radical path forward. As I have reflected on the 2023–24 academic year, I have begun to wonder if some of the issues I encountered were a result of trying to implement ungrading within the educational status quo. Stommel makes a striking claim on this point:

There is nothing ideologically neutral about grades, and there is nothing ideologically neutral about the idea that we can neatly and tidily do away with grades. We can't simply take away grades without re-examining all of our pedagogical approaches.¹⁶

Stommel's "all" is what I want to focus on. I often felt like I was rethinking everything during the 2023–24 academic year, but I have begun to realize how little actually changed about my courses. A visitor to my classroom likely would not have noticed anything particularly different. I removed grades, but almost everything else was the same. Perhaps more importantly, though, I was only one person within a large system. No matter how radically I might rethink assessment in my courses, students will still be taking courses that use traditional assessment models at the same time. As students themselves identified, when only one course makes these kinds of changes, that is the course where corners will be cut. And beyond this, by the time students are in college, they will likely have spent twelve or more years in an educational system steeped in standardized testing, ranking, and grades.

specified minimum standard. In standards-based grading, final grades are tied to the completion of a specific set of standards that specify particular skills and/or knowledge. More information about both of these assessment models is available in Clark and Talbert (2023).

¹⁶ Stommel (2023).

So yes, perhaps ungrading *is* impractical. Perhaps one could even say that it does not work within the current educational system. If this is true, it is because ungrading is fundamentally at odds with the ways that we think about and practice education in the United States. Perhaps the best thing for my students would be to transition to a more practical alternative assessment model. But this does not mean that I am giving up on ungrading.

According to Clark and Talbert, grades as we know them are products of the American university during the nineteenth century.¹⁷ And the timing is no accident. This is when education was becoming industrialized, and we continue to operate on that industrial model today. As Clark and Talbert put it:

[Traditional grading practices] have remained, not because they have stood the test of time and are the logical endpoints of centuries of mindful refinement, but primarily because they are useful—efficient, scalable, and interoperable—particularly so for the large bureaucracies that educational systems have become.¹⁸

Grades are thus tied up with the idea that education can be standardized and mass produced. With the idea that education is a commodity that can be sold for profit. With the idea that knowledge is something that can be deposited neatly from the mind of the teacher into the minds of the students, to borrow Paulo Freire's banking metaphor.¹⁹ And even more troubling, grades are associated with the idea that a college education is only meant for certain types of people. As Cathy Davidson and Christina Katopodis point out, many men who were instrumental in shaping the American university were card-carrying eugenicists.²⁰

Assessment is only one aspect of a much larger conception of education, and there are limits to what a single person can realistically change on their own. But I believe real change can happen if teachers everywhere began interrogating their assessment practices and *all* the other pedagogical practices that are wrapped up in them. Ungrading is not a silver bullet that will fix all that is wrong with education in the United States, but perhaps the end of grading could be the start of something truly transformational. As Morris writes:

¹⁷ Clark and Talbert (2023, 13–16); Smallwood (1935).

¹⁸ Clark and Talbert (2023, 16).

¹⁹ Freire (1970, 58).

²⁰ Davidson and Katopodis (2022, 21–24).

At the foundation of ungrading lies something that should change school entirely: a suggestion that ranking and evaluation, and the concomitant expertise of the ranker or evaluator, is entirely an optional way of viewing things. Ungrading should, when it's done, rearrange the room, placing the teacher in amongst students, dismantling the podium, rewriting textbooks such that the class itself becomes the text, and the rubric for comprehension of that text no less than a steady engagement with one another.²¹

While I am not going to continue using the assessment model described in this article, I hope that my experience will be productive for readers. I am not suggesting that everyone needs to come around to my way of thinking so that I can do what I want. What I *am* suggesting is that the issues I encountered with ungrading might be indicative of something wrong with the educational system, rather than the assessment model. My goal for this article is not to persuade readers to adopt my assessment model, but rather to inspire others to ask questions about how we assess learning and to think about how we could do it differently.

21 Morris (2021).

Bibliography

- Alegant, Brian. 2018. "Teaching Post-Tonal Aural Skills." In *The Norton Guide to Teaching Music Theory*, edited by Rachel Lumsden and Jeffrey Swinkin, 147–60. New York: W. W. Norton.
- Alegant, Brian, and Barbara Sawhill. 2013. "Making the Grade (or Not): Thoughts on Self-Design, Self-Assessment, and Self-Grading." *Engaging Students: Essays in Music Pedagogy* 1. <http://flipcamp.org/engagingstudents/alegantSawhill.html>.
- Blum, Susan D., ed. 2020. *Ungrading: Why Rating Students Undermines Learning (and What to Do Instead)*. Morgantown: West Virginia University Press.
- Clark, David. 2023. "An Alternative Grading Glossary." *Grading for Growth* (blog). January 30, 2023. <https://gradingforgrowth.com/p/an-alternative-grading-glossary>.
- Clark, David, and Robert Talbert. 2023. *Grading for Growth: A Guide to Alternative Grading Practices that Promote Authentic Learning and Student Engagement in Higher Education*. New York: Routledge.
- Davidson, Cathy N., and Christina Katopodis. 2022. *The New College Classroom*. Cambridge, MA: Harvard University Press.
- Duker, Philip, Anna Gawboy, Bryn Hughes, and Kris P. Shaffer. 2015. "Hacking the Music Theory Classroom: Standards-Based Grading, Just-in-Time Teaching, and the Inverted Class." *Music Theory Online* 21, no. 1. https://www.mtosmt.org/issues/mto.15.21.1/mto.15.21.1.duker_gawboy_hughes_shaffer.html.
- Freire, Paulo. 1970. *Pedagogy of the Oppressed*. Translated by Myra Bergman Ramos. New York: Herder and Herder.
- Guskey, Thomas R. 2019. "Grades Versus Comments: Research on Student Feedback." *Phi Delta Kappan* 101, no. 3. <https://journals.sagepub.com/doi/full/10.1177/0031721719885920>.
- Inoue, Asao B. 2022. *Labor-Based Grading Contracts: Building Equity and Inclusion in the Compassionate Writing Classroom*. 2nd ed. Denver: University of Press of Colorado.
- . 2023. *Crippling Labor-Based Grading for More Equity in Literacy Courses*. Denver: University Press of Colorado.
- Johnson, Vicky. 2015. "Proficiency-Based Learning with Muscle in a Music Theory Classroom." *Engaging Students: Essays in Music Pedagogy* 3. <http://flipcamp.org/engagingstudents3/essays/johnson.html>.
- Kohn, Alfie. 2013. "The Case against Grades." *Counterpoints* 451: 143–53.
- Melzer, Dan, D. J. Quinn, Lisa Sperber, and Sarah Faye. n.d. "Contract Grading—So Your Instructor is Using Contract Grading . . ." Writing Commons. Accessed December 12, 2024. <https://writingcommons.org/article/so-your-instructor-is-using-contract-grading/>.
- Michaelsen, Garrett. 2019. "Teaching with Radical Optimism: Mastery Learning in Music Theory." *Engaging Students: Essays in Music Pedagogy* 7. <https://doi.org/10.18061/es.v7io.7370>.
- Morris, Sean Michael. 2021. "When We Talk about Grades, We Are Talking about People." *Sean Michael Morris* (blog). June 9, 2021. <https://www.seanmichaelmorris.com/when-we-talk-about-grading-we-are-talking-about-people/>.
- Sackstein, Starr. 2015. *Hacking Assessment: 10 Ways to Go Gradeless in a Traditional Grades School*. Cleveland, OH: Times 10.

- Smallwood, Mary Lovett. 1935. *An Historical Study of Examinations and Grading Systems in Early American Universities: A Critical Study of the Original Records of Harvard, William and Mary, Yale, Mount Holyoke, and Michigan from Their Founding to 1900*, vol. 24. Cambridge, MA: Harvard University Press.
- Stommel, Jesse. 2023. *Undoing the Grade: Why We Grade and How to Stop*. Denver, CO: Hybrid Pedagogy. <https://pressbooks.pub/thegrade/>.
- Talbert, Robert. 2021. "A Word about Words: Why We are Moving on from the Term 'Mastery Grading.'" *Grading for Growth* (blog). November 15, 2021. <https://gradingforgrowth.com/p/a-word-about-words>.

