

De-weaponizing Grades Through Teacher Tinkering and Student Insight

Lance Piantaggini

University of Massachusetts Amherst, USA

ABSTRACT

When high-stakes assessments increase pressure and anxiety for students and teachers, they become weaponized. Jack Schneider and Ethan Hutt (2023) identified the permanence of grades as the root cause of weaponization in the American grading system. In this article, the author documents how they de-weaponized grades in their 9th-grade urban public school Latin classroom by using ungraded portfolios combined with self-grading. Seeking evidence of student experiences with grading throughout the school and looking to identify other practices that de-weaponize grades, the author conducted an exploratory qualitative study, blending a document analysis of 43 teachers' syllabi with in-depth interviews involving three 18-year-old senior class student participants. Participants shared negative experiences with procedures and policies related to grade permanence, including assessments lacking feedback and grading based on completion, accuracy, and effort. These procedures and policies appeared in 74% of teacher syllabi, showing that it was common practice to weaponize grades. Participants each suggested practices that could address grade permanence, including flexible deadlines, retakes, self-grading, and ungrading. These practices appeared in very few teacher syllabi, underscoring the need for de-weaponizing grades. This paper both presents solutions for de-weaponizing grades and shows how practitioners can engage in small-scale research at their schools to support reform efforts by using local, empirical findings.

Keywords: grading, ungrading, self-grading, portfolio assessment

In their book *Off the Mark: How Grades, Ratings, and Rankings Undermine Learning (but Don't Have To)*, authors Jack Schneider and Ethan Hutt (2023) tracked the development of the American grading system to its current problems. They identified one problem as "weaponization," referring to high-stakes assessments that increase pressure and anxiety for students and teachers (p. 161). The root cause of weaponization is the permanence of standardized test scores and grades. Therefore, classroom practices that make grades "overwritable" serve as teachers' and students' main defense towards de-weaponizing grades (p. 232).

In Part I of this article, the author identifies weaponized grading practices found throughout the literature and various attempts to address grade permanence. In Part II, they document the overwriting of grades in their 9th-grade public school Latin classroom by using ungraded portfolios and self-grading. Part III features a small-scale exploratory study involving 43 teachers' syllabi and interviews with three 18-year-old senior class student participants guided by the research questions: What are students' perceptions of teachers' assessment and grading practices? How prevalent were practices used to overwrite grades at the school?

Findings show weaponization present throughout the school and how student perceptions of teachers' assessment and grading practices were mostly negative. Finally, the author concludes with promising solutions for de-weaponizing grades, based on students' suggestions in the study and supported by the literature.

PART I: LITERATURE REVIEW

Weaponization of Grades

Weaponization of grades in classrooms involves assessments that increase pressure and anxiety for students through permanence. An assessment is considered permanent when a grade is recorded in the school's learning management system (LMS) and has no opportunity to be overwritten, such as through revising and resubmitting (Schneider & Hutt, 2023). Grades are literally documented on a student's permanent record, whether a series of "50s" and "60s" on individual assignments or a "D" for the overall grading term. Unfortunately, the literature shows that students become overly concerned with avoiding low grades out of fear of this documented failure (Beck et al., 1991; Butler & Nisan, 1986; Chamberlin et al., 2018). An assessment could also be considered permanent if students perceive it to have lasting effects, such as constant low scores as a reflection of permanent traits, resulting in low self-esteem and disengagement from learning (Black et al., 2004; Harlen & Deakin-Crick, 2002; Koenka, 2022; Koenka et al., 2021; Pulfrey et al., 2011; Schinske & Tanner, 2014; Shim & Ryan, 2005). Thus, practices that overwrite grades, such as revisions and retakes, portfolios, and self-grading, and that create an environment in which their effects are not considered lasting, can address this permanence.

Addressing Permanence

Revision & Retakes

Overwriting assessment results can be accomplished by students revising their work and retaking assessments, which scholars advocate. As Jung (2024) said, "denying chances to revise or try again undermines the very purpose of education" (p. 179). Similarly, Townsley (2022) wrote, "educators should provide students with multiple opportunities spread out over time, rather than a single point in time, to demonstrate their learning" (p. 87). New assessment results can replace old ones in the LMS after further instruction, revision, and/or retakes. Feldman (2023) recommends that these revisions and retakes replace old ones without penalty.

Portfolios

Portfolios typically involve students collecting their own work and providing rationales for their chosen evidence, encouraging more ownership of their learning (Mahmood & Jacobo, 2019; Wormeli, 2018). For example, Bulgakov Cooke et al. (2022) reported the status of implementation goals at a middle school using digital portfolios over two years. The authors found that 72% of students felt that they had a say about what happened to them at school when using portfolios to curate their own work. This was up from 57% the year before. Since students are more likely to add their best work to portfolios, there is a low chance that they contain assessments of poor quality that would otherwise have lasting negative effects. Challenges to implementing portfolios, however, include concerns about a lack of standardization across teachers and inconsistent scoring of student work (Koretz, 1998; Schneider & Hutt, 2023). Systems to ensure uniformity can address some of these concerns (see, for example, Black & Wiliam, 2018).

Self-grading

Grading is a summative function used to report achievement at the end of learning (Black & Wiliam, 1998; Shepard et al., 2018). Self-grading promotes responsibility and accountability, and it can disrupt power relations over what becomes permanent (Nieminen, 2022). Still, some have advocated against self-grading, instead recommending ungraded self-assessment (Bourke, 2018). For example, Andrade (2019) identified self-assessment as a self-regulating skill that must occur before self-grading, and that this skill must be developed through an ungraded, formative process, not a graded, summative event. Andrade's position is that formative self-assessment is more productive than summative self-grading. Andrade also noted that students' self-assigned grades tend to be higher than grades given by their instructors (see Strong et al., 2004; Tejeiro et al., 2012), yet some studies have shown no significant differences between students' and teachers' grades (Darrow et al., 2002; Falchikov & Boud, 1989; Falchikov & Goldfinch, 2000; Gorichanaz, 2024; Guberman, 2021; Kitchen et al., 2006; Sadler & Good, 2006; Sanchez et al., 2017). Beyond self-grading's validity, Boud (1989) made a case for self-grading on the grounds of "expediency," freeing up teacher time to provide more feedback rather than grading (p. 22).

PART II: TINKERING WITH OVERWRITING GRADES

Tyack and Cuban (1997) identified teachers as key actors in reform, tinkering with their practices towards positive change. Schneider and Hutt (2023) suggested recalibrating systems rather than reinventing them. These authors share the idea that small adjustments to practice can have a big impact on teaching and learning. In this section, the author documents recalibrations made to their classroom systems while tinkering towards de-weaponizing grades.

Ungraded Portfolios

While revision and retakes are overwritable options when assessments are graded, the alternative is to not grade individual assessments at all. Thus, the author used portfolios as collections of ungraded learning evidence showcasing student understanding and performance. Ungraded portfolios, then, effectively bypassed the need for revisions or retakes since grades were not being recorded outside of progress reports and quarterly report cards. Furthermore, students could choose to add early work to their portfolio and replace it later, or to wait until they had a final, more polished product to add. In terms of addressing permanence, it was not so much that grades were overwritable: the assessments themselves were overwritten.

Students added work to portfolios each week and were given time at the end of class to do so. The teacher encouraged students to add some assignments, while other assignments were of the students' choosing. It is important to note that the use of portfolios did not hinder the formative feedback process. If anything, the author observed that portfolios helped the process by preserving written feedback, available for students to access and review in class. Furthermore, since reviewing work was part of the self-grading procedure, portfolios supported the development of self-regulated learning skills.

Self-grading

While Schneider and Hutt (2023) noted problems with large-scale portfolio use, the author's experience was different at the classroom level. One practice used in combination with ungraded portfolios that may have mitigated negative effects was self-grading. During each term, students graded themselves using holistic rubrics that listed broad descriptions of the characteristics defining each level of the standard. These rubrics were presented in a Google Form, in which students selected one of six grades keyed to each level and provided a rationale with supporting evidence (e.g., "My grade should be ___ because ___ shows that I ___"). Figure 1 is one of two holistic rubrics used throughout the year.

Figure 1

Holistic Reading Rubric

A LOT of growth <i>Texts feel super easy to read, and they feel at or below my reading level.</i>	100 (A+)
Growing <i>Older texts seem easier to read, and/or current ones feel like my reading level.</i>	95 (A)
Staying the same <i>Older texts seem just right, and/or current ones feel a little above my reading level.</i>	85 (B)
Fading <i>Older texts seem a little hard to read, and/or current ones feel above my reading level.</i>	75 (C)
Really fading <i>Older texts seem hard to read, and/or current ones feel far above my reading level.</i>	65 (D)
Falling behind <i>Texts feel very hard to read, and/or I feel like I've never seen these words before.</i>	55 (F)

Note. The revised design of numerical values with midpoint cutoffs (i.e., 55, 65, 75, 85, 95) was used after the author observed how a previous grading scale (i.e., 50, 60, 70, 80, 90) became a source of unnecessary frustration for students who were stressed by corresponding letter grades below the midpoint with minuses (i.e., D-, C-, B-, A-). Furthermore, seeking a perfect grade point average (GPA) drove some students to focus on their grades more than on learning. Since grades above "93" were equal to a GPA of "4.0" in

the district, a “95(A)” achieved the maximum GP. The “100(A+)” on this revised scale represented achievement beyond district and classroom expectations.

Students tended to under-grade themselves more often than to over-grade, perhaps a sign of low confidence, as documented in traditional point-accumulation classrooms (Black et al., 2004; Chamberlin et al., 2018; Townsley, 2022). In the few instances when a student’s grade and/or rationale misaligned with the teacher’s records and observations, the teacher and student met to discuss the discrepancies. Differences were resolved by the teacher showing the student how their learning evidence met or did not meet expectations, and by recording the student’s new grade, which now aligned with the teacher’s.

In sum, the author de-weaponized grades by using an ungraded portfolio system combined with self-grading. No permanent grades were reported until the end of the grading term. Students had autonomy to replace early learning evidence with recent learning evidence, and their self-assessment and self-grading of evidence could be adjusted through discussion. From the author’s perspective, these recalibrated practices were fluid and lacked the permanence that weaponizes grades. Yet questions remained: Were teachers de-weaponizing grades in other ways throughout the school? What did other students think about their teachers’ assessment and grading practices? To address these questions, attention was given to the most important interest holders: students themselves.

PART III: EXPLORATORY STUDY

Purpose & Research Questions

The purpose of this exploratory qualitative study was to explore students’ experiences with grading across the school and to identify practices to de-weaponize grades. This study was guided by the research questions: What are students’ perceptions of teachers’ assessment and grading practices? and How prevalent were practices used to overwrite grades at the school?

Setting & Participants

This study took place at a small, urban, public high school in New England. Participants were a convenience sample of students at the author’s school who were not taught by the author and were able to give consent independently. An interest survey was sent to all 18-year-old seniors. Three of these students agreed to participate.

METHODS

Data Collection

Given the purpose of exploring participants’ lived experiences, interviews were the most appropriate choice. These were conducted late in the fall semester during students’ study hall to avoid disrupting class time. An interview protocol was developed in consultation with a university professor of qualitative methods (see Appendix A). Interviews lasted between 15 and 30 minutes.

Data Analysis

Interviews were thematically coded by the author in Google Sheets, using initial coding followed by focused coding (Saldaña, 2013). This coding was done inductively to explore new ideas free from preconceived ones. Initial codes were refined during focused coding to identify recurring patterns in the interviews representing common experiences across participants. After identifying shared experiences with grading, 43 syllabi across all 21 teachers were collected and reviewed. The grading policy section of each syllabus was thematically coded using a procedure similar to that used for the interviews. For these documents, however, a deductive approach was used, based on student responses, to determine the prevalence of de-weaponizing grades at the school.

FINDINGS

Focused coding resulted in six categories, shown in Table 1. Based on these categories, participant experiences were categorized into two themes: *negative experiences* and *wanting more ownership*.

Table 1

Interview Categories and Their Reference Counts

Categories	References
Motivation & Interests	130
Procedures & Policies	111
Grades & Performance Expectations	86
Navigating Demands of High School	42
Meaning of Grades	18
Social Impact of Grades	17

Negative Experiences

Students felt their learning was restricted by assessments lacking feedback and graded on completion, accuracy, and effort. Their descriptions evoked a sense of compliance rather than curiosity about learning. All three students used remarkably similar phrasing regarding the lack of feedback on their work. Lina questioned, “I don't know why I'm getting a C or something.” Sable remarked, “There was just one C . . . I was like ‘what did I do wrong?’” Nigel commented, “I took a test and then ended up getting a bad grade . . . ‘Oh what did I do wrong?’” While letter grades provide information about achievement, they offer no actionable steps for students. When asked about the lack of feedback, Lina wished her teachers “had said things more like ‘here's what you can do to practice more to make sure you don't get those problems wrong in the future.’”

Assessments graded on completion, accuracy, and effort were also associated with a lack of meaningful purpose. For example, Nigel described “being fed information,” and Lina found a lack of feedback unhelpful:

A lot of times . . . your teacher will just mark your answer wrong and that's your grade. If you're just grading them on the number instead of how much they understand it, it doesn't really help them. I think that's the base of how most grading is, like the number of things you got right or wrong.

Syllabi showing assessments graded on completion, accuracy, or effort were quite common in the school. Such methods appeared on 74% of the syllabi, including the science course Sable was taking. Table 2 shows the overview of the teacher syllabi document analysis.

Wanting More Ownership

The second theme was participants wanting more ownership of their learning. Students suggested practices that would overwrite grades, including flexible deadlines, retake opportunities, self-grading, and courses with an ungraded approach.

Flexible Deadlines

Sable spoke about allowing students to slow down when concepts become challenging or speed up when thorough understanding is reached:

Yeah it might take longer. Like, you will learn one topic at a time and then you move on to the next topic. That would definitely raise a lot of grades. That's what I feel like most teachers are missing out on. Like, learning for yourself not for the teacher.

Sable's description suggests that flexible deadlines could address low grades resulting from slow learning rates. She was in a learning environment with mostly fixed deadlines, as evidenced by 35% of syllabi listing late-work penalties and none explicitly indicating the possibility of extensions.

Table 2***Practices Found in Teachers' Syllabi***

Practice	# Course Syllabi	% Course Syllabi	# Teachers	% Teachers
Grading Completion, Accuracy, Effort	32	74	16	76
Undefined Standards	16	37	10	48
Late Work Penalties	15	35	8	38
Retakes	7	16	6	29
Self-grading	4	9	3	14
Mention of Grading Scale	2	5	2	10
Ungrading	2	5	2	10

Retakes

Sable also thought students should have more opportunities to learn. This included more ways to show improvement and involve teachers guiding students through greater understanding via reteaching:

I feel like they should go like second rounds. I feel that's the purpose of learning...you know, like, "Oh, you did this wrong, but I'll give you another chance to redeem yourself, but I will re-teach you before I give you the chance to redeem yourself or have a retake." If there was more chances teachers give to students to learn more . . . like for redemption.

When asked about her experience with retakes and teachers reteaching content, Sable shared, "It doesn't really happen often, especially in some courses that are considered more severe, like AP classes, for example. Those are pretty big classes, so yeah." Syllabi data showed seven instances of retake policies, sometimes conditional (e.g., optional retakes if the first score is below 80), making retake policies an uncommon grading practice in the school (See Table 2).

Self-grading

Nigel wanted students to have more responsibility in determining their course grade. He acknowledged that this would require teachers to relinquish some control and require trust. Nigel also thought it would hold students accountable as they develop into self-regulated learners:

I would say let the students grade themselves, but you don't know if they're gonna be honest or not . . . I think letting students take responsibility for what they've learned . . . to see how honest they are, and to see if they actually learned what was taught to them . . . if they're mature enough to take responsibility and show the teacher they can really do it, and show they learned or maybe they didn't learn so they'll give themselves a bad grade, right?

Syllabi data shown in Table 2 indicate that self-grading was an uncommon grading practice found in just three syllabi, ranging from partially to fully determining the course grade.

Ungrading

Ungrading is the elimination or reduction of grades throughout the learning process (Kohn, 2020; Sorensen-Unruh, 2020). Lina described how learning from mistakes in a classroom without grades best supported her learning. Mistakes were not only allowed but even welcomed by her teacher as part of class procedures:

We were going over a test we did, and I got most of it wrong, which was fine because we were going over how to fix that. That's one thing I like about class; she doesn't grade us on . . . how many questions we got wrong, she grades us on our understanding . . . we review it and we go over and we learn to solve the problems . . . I just find that's a lot better than just like "oh this was wrong."

Syllabi data shown in Table 2 indicate that ungrading was rare in the school, found in only two syllabi. Furthermore, Lina was the only student suggesting a practice she had experienced.

DISCUSSION

While negative effects of grading on student motivation have been documented in the past (Beck et al., 1991; Butler & Nisan, 1986; Chamberlin et al., 2018), findings from the current study provide deeper insight into how students experience those negative effects through their responses. All students felt their learning was restricted by assessments graded on completion, accuracy, and effort. Townsley (2022) lists these factors as contributing to a "point-accumulation classroom" favoring assessments as gradebook events, rather than formative assessments designed to improve learning (p. 14). Assessments graded on completion, accuracy, and effort were found in 74% of teacher syllabi in this study, making the point-accumulation classroom the school's standard. These findings are similar to those of Harland et al. (2015), who found that "in some situations it was possible to experience hundreds of graded assessments in an academic year" (p. 528), with students focused on doing work for grades and points rather than for their own learning. Therefore, in the current study, the lack of clarity about how students could improve their learning due to insufficient feedback is unsurprising. Students found themselves in a grade-heavy environment and suggested ways to address the permanence of grades.

Sable's desire for retakes was strong. Her use of words like "redeem" and "redemption" gives us a glimpse into what a teenager can be experiencing as a result of a teacher's procedures and policies. Retake policies can harness the power to neutralize the onslaught of grades and have support in the previously mentioned literature (Feldman, 2023; Jung, 2024; Townsley, 2022). Furthermore, Nigel's suggestion of self-grading is supported by the literature (Darrow et al., 2002; Falchikov & Boud, 1989; Falchikov & Goldfinch, 2000; Gorichanaz, 2024; Guberman, 2021; Kitchen et al., 2006; Sadler & Good, 2006; Sanchez et al., 2017).

Lina's experience in a classroom that welcomed mistakes and had no grades was positive. Ungrading has been described as "an attempt to minimize the damaging effects of the final grade" (Kohn, 2020, p. xvi), and a way of "purposely eliminating or minimizing the use of points or letters to assess student work" (Sorensen-Unruh, 2020, p. 140). Given the point-accumulation classroom as the school's norm, ungrading would address this permanence. Researchers recommend replacing grades with feedback (Gorichanaz, 2024; Koenka et al., 2021; Newton et al., 2020; Rapchak et al., 2023), though not all educators see ungrading as a worthwhile alternative. Davis (2023) was unconvinced that grades are the problem, citing results from her study comparing three grading methods for first-year undergraduate writing students: traditional grading, portfolio grading, and ungrading. For the ungrading approach, Davis had students self-grade at the end of the term. Davis' main argument was that summative self-assessments used to arrive at grades did little to build the metacognitive skills leading to self-regulated learners. Still, a single instance of summative self-assessment at the end of the grading term may be beneficial if it allows for grading to be reduced throughout the rest of the learning process, as Lina experienced. Dyer (2024) was also concerned about the negative impact of implicit bias in both the instructor and the student when determining final grades. While implicit bias is likely present in any grading approach, Dyer claimed that the impact would be greater if self-grading were used for final course grades. The logic behind this claim is understandable. If implicit bias skews how achievement is reported, it is best suited to low-impact assignments. Still, implicit bias issues must be weighed against other negative impacts in the current point-accumulation classroom, such as students experiencing weaponized grade permanence more often. An ungraded approach might not be immune to bias, as Dyer writes, yet its benefits might outweigh this potential drawback. More research is needed on the equitable outcomes of ungrading. If done well, this approach could address grade permanence, de-weaponizing grades to the greatest extent possible by limiting them until absolutely necessary, such as at the end of the grading term.

Sable had negative experiences with fixed deadlines set by curricular pacing, and syllabi referenced accepting late-work penalties. Bloom (1968) advocated that students learn at their own pace, with teachers using formative assessment to provide feedback until they achieve mastery. This ongoing assessment process had no deadlines and was personalized according to individual student learning rates. Hills and Peacock (2022) report findings from a study involving 43 participants in an undergraduate biology course. Every assignment had a deadline and a one-week extension option. Of the 301 possible extensions, students chose this option for just 54 (18%). This shows students opted for extensions only as needed, and what was needed varied from student to student. For example, nearly all students (95%) chose the extension for the one high-

stakes essay in the course. For the other low-stakes assignments, 40% of students chose only one extension, and 35% chose none. These findings illustrate the range of students' needs and how flexible deadlines gave students greater ownership of their learning. While teachers can reasonably expect to teach students time-management skills to meet deadlines, late-work penalties harm slower learners, and fixed deadlines have been shown to decrease performance and engagement (Miller et al., 2019). In the spirit of providing an equitable education to all, flexible deadlines could be considered the garrison of late work penalties and would answer the call-to-arms against weaponized grades. Concerns about flexible deadlines exist, such as "cautioning about an avalanche of work to grade at the end of the semester to student exploitation of this policy" and a lack of "emphasis on personal responsibility" (Kruger, 2023, p. 234). However, as Kruger (2023) stated, there is little empirical evidence that such concerns manifest when deadline extensions are offered. Therefore, schoolwide flexible deadlines could result in overwriting what would otherwise be a permanent grade representing timeliness, not achievement.

LIMITATIONS & FUTURE RESEARCH

Syllabi alone could not have captured any teacher's complete practices. The degree to which teacher syllabi matched what was happening in class is unknown. Future research could validate syllabi data with observations, teacher interviews, assessment samples, or surveys. The small student sample size and short interview lengths were also limitations. Future research could involve a larger student sample with longer, multi-session interviews.

The revised 6-point scale in Figure 1 has not been researched, though it falls within optimal scale levels found to reduce variability and error (Guskey, 2024; Lozano et al., 2008). Its design, however, was intended to address motivational issues. Future research should investigate whether this scale keeps a focus on learning by eliminating the distraction of minuses (i.e., F-, D-, C-, B-, A-) and student concerns with achieving the highest possible score of "100."

Finally, future research should investigate how student-suggested practices are used across multiple teachers, and whether combinations of these practices are more effective than others in de-weaponizing grades, such as the combination of ungraded portfolios and self-grading used by the author. Furthermore, practitioners are encouraged to conduct small-scale studies in their own buildings to explore the de-weaponization of grades.

CONCLUSION

The author sought evidence of student experiences with grading and wanted to identify practices in the school that de-weaponize grades. Students identified negative experiences with grading, including a lack of feedback and assessments graded on completion, accuracy, and effort. Weaponized practices that emphasized points and grades were found in 74% of syllabi. Grades constantly surrounded students, and their interview responses showed concern about grade permanence.

Participants wanted more ownership of their learning and suggested practices to overwrite grades, including flexible deadlines, retakes, self-grading, and ungrading. The latter two practices were used in the author's own classroom. Furthermore, the combination of ungraded portfolios with self-grading asked students to be more responsible, as Nigel requested, building important self-regulated learner skills. However, the suggested practices to address grade permanence were uncommon elsewhere in the school.

While the generalizability of these findings is limited, this study shows that grade de-weaponization was rare throughout the school, though the practices that students suggested could be promising solutions for grade de-weaponization by overwriting them, just as Schneider and Hutt (2023) suggested. Therefore, practitioners are encouraged to explore the solutions presented in this paper, including flexible deadlines, retakes, and ungraded portfolios combined with self-grading.

REFERENCES

- Andrade, H. (2019). A critical review of research on student self-assessment. *Frontiers in Education*, 4(87). <https://doi.org/10.3389/feduc.2019.00087>
- Beck, H. P., Rorrer-Woody, S., & Pierce, L. G. (1991). The relations of learning and grade orientations to academic performance. *Teaching of Psychology*, 18(1), 35–37. https://doi.org/10.1207/s15328023top1801_10
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86(1), 8–21. <https://doi.org/10.1177/003172170408600105>
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551–575. <https://doi.org/10.1080/0969594X.2018.1441807>

- Bloom, B. S. (1968). Learning for mastery. Instruction and curriculum. Regional education laboratory for the Carolinas and Virginia, topical papers and reprints, number 1. *Evaluation Comment*, 1(2), 1–12.
- Boud, D. (1989). The role of self-assessment in student grading. *Assessment & Evaluation in Higher Education*, 14(1), 20–30. <https://doi.org/10.1080/0260293890140103>
- Bourke, R. (2018). Self-Assessment to incite learning in higher education: Developing ontological awareness. *Assessment & Evaluation in Higher Education*, 43(5), 827–839.
- Bulgakov Cooke, D., Paepflow, C., Carter, J., & Belton, T. R. (2022). *Middle school competency-based grading initiative: 2018-19 to 2020-21*. Wake County Public School System. <https://eric.ed.gov/?id=ED624580>
- Butler, R., & Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal of Educational Psychology*, 78(3), 210–216. <https://doi.org/10.1037/0022-0663.78.3.210>
- Chamberlin, K., Yasué, M., & Chiang, I.-C. A. (2018). The impact of grades on student motivation. *Active Learning in Higher Education*, 24(2), 109–124. <https://doi.org/10.1177/1469787418819728>
- Darrow, A.A., Johnson, C. M., Miller, A. M., & Williamson, P. (2002). Can students accurately assess themselves? Predictive validity of student self-reports. *Update: Applications of Research in Music Education*, 20(2), 8–11. <https://doi.org/10.1177/875512330202000203>
- Davis, H. T. (2023). Ungrading: Self-Assessment, effort, and motivation. *Composition Studies*, 51(2), 143–147.
- Dyer, J. (2024, January 8). Ungrading has an equity-related Achilles heel. *Grading For Growth*. <https://gradingforgrowth.com/p/ungrading-has-an-equity-related-achilles>
- Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research*, 59(4), 395–430. <https://doi.org/10.2307/1170205>
- Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70(3), 287–322. <https://doi.org/10.3102/00346543070003287>
- Feldman, J. (2023). *Grading For Equity: What It Is, Why It Matters, and How It Can Transform Schools and Classrooms* (2nd ed.). Corwin.
- Gorichanaz, T. (2024). “It made me feel like it was okay to be wrong”: Student experiences with ungrading. *Active Learning in Higher Education*, 25(1), 67–80. <https://doi.org/10.1177/14697874221093640>
- Guberman, D. (2021). Student perceptions of an online ungraded course. *Teaching & Learning Inquiry*, 9(1), 86–98. <https://doi.org/10.20343/teachlearningqu.9.1.8>
- Guskey, T. R. (2024, April 29). Addressing inconsistencies in grading practices. *Kappan Online*. <https://kappanonline.org/addressing-inconsistencies-in-grading-practices/>
- Harland, T., McLean, A., Wass, R., Miller, E., & Sim, K. N. (2015). An assessment arms race and its fallout: High-stakes grading and the case for slow scholarship. *Assessment & Evaluation in Higher Education*, 40(4), 528–541. <https://doi.org/10.1080/02602938.2014.931927>
- Harlen, W., & Deakin-Crick, R. (2002). *A systematic review of the impact of summative assessment and tests on students' motivation for learning*. EPPI-Centre, University of London.
- Hills, M., & Peacock, K. (2022). Replacing power with flexible structure: Implementing flexible deadlines to improve student learning experiences. *Teaching and Learning Inquiry*, 10. <https://doi.org/10.20343/teachlearningqu.10.26>
- Jung, L. A. (2024). *Assessing students, not standards: Begin with what matters most* (1st ed.). Corwin.
- Kitchen, E., King, S. H., Robison, D. F., Sudweeks, R. R., Bradshaw, W. S., & Bell, J. D. (2006). Rethinking exams and letter grades: How much can teachers delegate to students? *CBE—Life Sciences Education*, 5(3), 270–280. <https://doi.org/10.1187/cbe.05-11-0123>
- Koenka, A. C., Linnenbrink-Garcia, L., Moshontz, H., Atkinson, K. M., Sanchez, C. E., & Cooper, H. (2021). A meta-analysis on the impact of grades and comments on academic motivation and achievement: A case for written feedback. *Educational Psychology*, 41(7), 922–947. <https://doi.org/10.1080/01443410.2019.1659939>
- Koenka, A. C. (2022). Grade expectations: The motivational consequences of performance feedback on a summative assessment. *The Journal of Experimental Education*, 90(1), 88–111. <https://doi.org/10.1080/00220973.2020.1777069>
- Kohn, A.F. (2020). Foreword. In S.D. Blum (Ed), *Ungrading: Why rating students undermines learning. (and what to do instead)* (p. xvi). West Virginia University Press.
- Koretz, D. M. (1998). Large-scale portfolio assessments in the US: Evidence pertaining to the quality of measurement. *Assessment in Education: Principles, Policy and Practice*, 5(3), 309–334.
- Kruger, J. S. (2023). Rethinking penalties for late work: The case for flexibility, equity, and support. *Pedagogy in Health Promotion*, 9(4), 234–236. <https://doi.org/10.1177/23733799231198778>

- Lozano, L., García-Cueto, E., & Muñiz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology*, 4, 73–79. <https://doi.org/10.1027/1614-2241.4.2.73>
- Mahmood, D., & Jacobo, H. (2019). Grading for growth: Using sliding scale rubrics to motivate struggling learners. *Interdisciplinary Journal of Problem-Based Learning*, 13(2). <https://doi.org/10.7771/1541-5015.1844>
- Miller, L. A., Asarta, C. J., & Schmidt, J. R. (2019). Completion deadlines, adaptive learning assignments, and student performance. *Journal of Education for Business*, 94(3), 185–194. <https://doi.org/10.1080/08832323.2018.1507988>
- Newton, J. R., Williams, M. C., & Feeney, D. M. (2020). Implementing non-traditional assessment strategies in teacher preparation: Opportunities and challenges. *Journal of Culture and Values in Education*, 3(1), 39–51. <https://doi.org/10.46303/jcve.03.01.3>
- Nieminen, J. H. (2022). Disrupting the power relations of grading in higher education through summative self-assessment. *Teaching in Higher Education*, 27(7), 892–907. <https://doi.org/10.1080/13562517.2020.1753687>
- Pulfrey, C., Buchs, C., & Butera, F. (2011). Why grades engender performance-avoidance goals: The mediating role of autonomous motivation. *Journal of Educational Psychology*, 103(3), 683–700. <https://doi.org/10.1037/a0023911>
- Rapchak, M., Hands, A. S., & Hensley, M. K. (2023). Moving toward equity: Experiences with ungrading. *Journal of Education for Library and Information Science*, 64(1), 89–98. <https://doi.org/10.3138/jelis-2021-0062>
- Sadler, P., & Good, E. (2006). The impact of self- and peer-grading on student learning. *Educational Assessment*, 11(1), 1–31. https://doi.org/10.1207/s15326977ea1101_1
- Saldaña, J. (2013). *The Coding manual for qualitative researchers* (2nd ed.). London: Sage.
- Sanchez, C. E., Atkinson, K. M., Koenka, A. C., Moshontz, H., & Cooper, H. (2017). Self-grading and peer-grading for formative and summative assessments in 3rd through 12th grade classrooms: A meta-analysis. *Journal of Educational Psychology*, 109(8), 1049–1066. <https://doi.org/10.1037/edu0000190>
- Schinske, J., & Tanner, K. (2014). Teaching more by grading less (or differently). *CBE-Life Sciences Education*, 13(2), 159–166. <https://doi.org/10.1187/cbe.CBE-14-03-0054>
- Schneider, J., & Hutt, E. L. (2023). *Off the mark*. Harvard University Press. <https://books.google.com/books?id=CoG3EAAAQBAJ>
- Shepard, L. A., Penuel, W. R., & Pellegrino, J. W. (2018). Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. *Educational Measurement: Issues and Practice*, 37(1), 21–34. <https://doi.org/10.1111/emip.12189>
- Shim, S., & Ryan, A. (2005). Changes in self-efficacy, challenge avoidance, and intrinsic value in response to grades: The role of achievement goals. *The Journal of Experimental Education*, 73(4), 333–349. <https://doi.org/10.3200/JEXE.73.4.333-349>
- Sorensen-Unruh, C. (2020). A stem ungrading case study: a reflection on first-time implementation in organic chemistry II. In S.D. Blum (Ed), *Ungrading: Why rating students undermines learning (and what to do instead)* (p. 140). West Virginia University Press.
- Strong, B., Davis, M., & Hawks, V. (2004). Self-grading in large general education classes: A case study. *College Teaching*, 52(2), 52–57.
- Tejairo, R. A., Gómez-Vallecillo, J. L., Romero, A. F., Pelegrina, M., Wallace, A., & Emberley, E. (2012). Summative self-assessment in higher education: Implications of its counting towards the final mark. *Electronic Journal of Research in Educational Psychology*, 10(2), 789–812.
- Townsend, M. (2022). *Using grading to support student learning* (1st ed.). Routledge.
- Tyack, D. B., & Cuban, L. (1997). *Tinkering toward utopia: A century of public school reform* (Revised edition). Harvard University Press.
- Wormeli, R. (2018). *Fair isn't always equal, 2nd edition: Assessment & grading in the differentiated classroom* (2nd ed.). Stenhouse Publishers.

LANCE PIANTAGGINI, EdS, is a PhD Candidate in the College of Education at the University of Massachusetts Amherst. His major research interests lie in the areas of classroom assessment, feedback, and grading. Email: lpiantaggini@umass.edu

Manuscript submitted: May 5, 2025

Manuscript revised: November 10, 2025

Manuscript revised: April 30, 2026

Accepted for publication: June 19, 2026

APPENDIX

STUDENT INTERVIEW PROTOCOL

Relevant Demographics

- Pronouns

Focused Life History

- What brought you to this school?
- What's your first memory of getting a grade?
 - What was that like?
 - What did that feel like?
 - Tell me about a time when you got a good/bad grade (opposite of what was shared above: use *their* terminology)

Experience

- How would you describe your academics in terms of learning?
 - How do grades align with that?

Reflection of Meaning

- Tell me what grades mean to you. What makes you want/not want to get good grades?
- If you were to change something about grading, what would it be?

Closing

- Is there anything you want to talk about that we didn't get to?