

Growth-based grading: a non-traditional approach to grading in leadership education

Hannah M. Sunderman

*Department of Agricultural, Leadership, and Community Education, Virginia Tech,
Blacksburg, Virginia, USA*

Received 6 June 2025
Revised 27 June 2025
Accepted 27 June 2025

Abstract

Purpose – A principal pursuit of higher education in the United States is fostering critical thinking, application and empowerment among students. However, a gap exists between this pursuit and traditional grading structures, which can restrict students' thinking and motivation in pursuit of grade achievement. Therefore, the current pedagogy manuscript describes the development, implementation and evaluation of one system of non-traditional grading, growth-based grading (GBG), within leadership education.

Design/methodology/approach – GBG emphasizes students' self-assessment of their learning in collaboration with the instructor. Specifically, GBG utilizes complete/incomplete assignments, heavy feedback and reflective portfolios to document course engagement and encourage students' metacognition. GBG was implemented over five courses at two large, land-grant institutions. Participants ($N = 83$) completed a survey evaluating their perceptions of GBG in comparison to traditional grading.

Findings – Quantitative survey results revealed a significant difference between perceptions of GBG and traditional grading with participants perceiving GBG more favorably than traditional grading. Open-ended survey responses emphasized that GBG reduced stress while enhancing focus on class content, learning and creativity. Participants also shared critiques of and suggestions for GBG, including providing more feedback on performance.

Practical implications – Curricular leadership educators are urged to regard grading as foundational to pedagogy, interrogating if and how assessment structures align with content, learning objectives and overarching beliefs about leadership education.

Originality/value – The field of leadership education has actively explored instructional practices and pedagogy; yet, grading structures, although a significant component of students' curricular experience, have largely been absent from inquiry.

Keywords Leadership pedagogy, Leadership assessment, Leadership education application and outcomes

Paper type Policy brief

Issue statement

The aspirational goals of higher education within the United States include promoting lifelong learning and critical thinking ([Council for the Advancement of Standards in Higher Education \(CAS\), 2015](#); [Golden, 2023](#)); yet, a foundational practice of higher education, traditional grading, can hinder students' abilities to realize these goals by shifting their focus from learning to grading ([Hayek, Toma, Oberlé, & Butera, 2015](#); [Kohn, 1999, 2006](#)). Bridging the gap between the goals of higher education and outcomes of traditional grading is essential in leadership education given that leadership educators seek to enhance students' capacity for critical thinking, reflection, application, self-understanding and empowerment ([Andenoro, 2013](#); [CAS, 2015](#); [Seemiller, 2016](#)).

Therefore, the current paper describes the creation, implementation and evaluation of growth-based grading (GBG), a method of non-traditional grading in which learners self-assess their development. Ultimately, a final letter grade is determined at the end of the

© Hannah M. Sunderman. Published in *Journal of Leadership Education*. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at [Link to the terms of the CC BY 4.0 licence](#).



Journal of Leadership Education
Emerald Publishing Limited
1552-9045
DOI 10.1108/JOLE-06-2025-0060

semester when students utilize a rubric to suggest a grade based on evidence of their learning and then engage in a discussion with their instructor. Drawing on practices common in non-traditional grading, which has also been described as ungrading or alternative grading (Blum, 2020a), GBG utilizes complete/incomplete assignments, heavy feedback and reflective portfolios to document course engagement and encourage students' metacognition. The current article will discuss the background and history of grading, highlight pertinent learning, outline the practices and processes of GBG, detail the findings that emerged from survey data, offer reflections and recommendations from implementation and, finally, share implications and future research directions for the field.

Literature review

Before exploring a non-traditional way of grading, it is important to begin by understanding the history of grades. For more than thirteen hundred years, imperial China utilized a rank-conscious system and structured examinations to assign governmental and career placement (Elman, 2000). In Europe, grading began at medieval and early modern universities as a means of denoting outstanding performance in public oral examinations (Blum, 2020b). In the eighteenth and nineteenth centuries, written exams became increasingly popular as the number of students enrolled in higher education grew. However, these exams were often pass or fail with students allowed to retake their exams until they passed. In the United States, more structured, numerical systems of grading took root. Today's common letter-grade structure (i.e. A, B, C, D, F) reportedly began at Mount Holyoke College in 1897 before spreading to most schools in the United States. As highlighted by the brief history on the development of today's common letter-grade structure, although grades are often utilized to assess student learning, grades and assessment are two distinct constructs (Tannock, 2017). Assessment, in general, is the evaluation, estimation or measurement of identified outcomes (e.g. growth, knowledge or skills). Grading (i.e. today's traditional letter-grade structure) is a specific means of evaluating student performance utilizing a standardized metric.

While today's common letter-grade structure spread (i.e. traditional grading), so, too, did opposition to traditional grading. Over a century ago, some believed in applying the principles of scientific management, such as maximizing efficiency, norming outcomes and sorting students, to education while others tended towards what is known today as "progressive education" or "new education" (Blum, 2020b), which focused on child-centered curriculum. Tension around the purpose of grading and assessment in schools continues today, as demonstrated by the 21st-century push for nationwide K-12 standardized testing (Alcocer, 2019) and the recent movement away from utilizing standardized test scores in determining college acceptance (Barshay, 2022, October 31).

Research within the past few decades (as well as research that is a century old; e.g. Starch & Elliott, 1912) has demonstrated that the negative effects of grades include decreasing students' interest in the actual learning, reducing students' preference for challenging tasks, lessening the quality of students' thinking, wasting time that could be spent on learning, encouraging cheating and hindering teachers' relationships with students (Kohn, 1999). Kohn (2006) shared that there are three consistent outcomes when students are graded: they (1) think less deeply, (2) avoid taking risks and (3) "lose interest in the learning itself" (p. 12). Butler (1988) examined the impact of only comments, comments and grades and only grades on student outcomes. The results revealed that students receiving only feedback reliably outperformed the other two groups of students. Lipnevich and Smith (2008) found that detailed, descriptive feedback was most effective for student improvement on essays when given without grades or praise. Hayek *et al.* (2015) discovered that groups of students primed as being graded for their individual, visible contribution offered less unshared information to other group members, indicating that grading can limit cooperation and collaboration in groups.

Within the field of leadership education, little scholarly work has been devoted to understanding the effect of grading on student learning, although significant recent scholarship

has studied leadership education pedagogy (Allen & Hartman, 2008, 2009; Cress, Astin, Zimmerman-Oster, & Burkhardt, 2001; Jenkins, 2012, 2016, 2018; Odom, 2015). While Jenkins (2016, 2018) researched instructional and assessment strategies in graduate and undergraduate leadership education courses, specifically analyzing the weight placed on assessment strategies (e.g. discussion boards, term papers) in students' grades, grading structures specifically require further investigation, revealing a gap in the leadership education scholarship. In sum, the field of leadership education has actively explored instructional practices and pedagogy; yet, grading structures, although a significant component of students' curricular experience, have largely been absent from inquiry. Therefore, the current article shares a non-traditional grading structure, GBG. The description of the practice and results have implications for the study and practice of curricular leadership education in institutions of higher education.

Theoretical and conceptual frameworks

The development and implementation of the GBG structure was influenced by two frameworks: constructivist learning and transformative learning. As other leadership educators contemplate their own grading structures, there are additional relevant frameworks to consider (e.g. feminist, critical; Blum, 2020b).

Constructivist learning

With theoretical roots among writers such as Dewey and Piaget (Phillips, 1995), constructivist learning assumes that learning is emergent and co-constructed among a group of learners (Bright *et al.*, 2016). More broadly, constructivism "stresses all knowledge is context bound, and individuals make personal meaning of their learning experiences" (Knowles, Holton, & Swanson, 1998, p. 142). As students face new knowledge and ways of understanding, they engage in meaning making, which culminates in learning (Piaget, 1977). Therefore, the goal of an instructor utilizing constructivist learning is to guide student participation and reflection towards new understandings. Rather than serving as a "sage on the stage," constructivist learning educators are a "guide on the side" (Ramsey & Fitzgibbons, 2005, p. 336), viewing all class components as potential learning experiences and opportunities for motivation (Bright *et al.*, 2016). The current work applies the concept of meaning making by co-constructing grades and utilizing grading as a learning opportunity that aligns with course content, pedagogy and my beliefs about leadership education (i.e. a belief that leadership education ought to be a learning relationship among all members of the classroom grounded in reflection and inclusivity).

Transformative learning

Developed by Mezirow (1978), transformative learning is defined as "the process by which we transform problematic frames of reference (mindsets, habits of mind, meaning perspectives) – sets of assumption and expectation – to make them more inclusive, discriminating, open, reflective and emotionally able to change" (Mezirow, 2008, p. 92). Specifically, transformative learning is a student-centered approach that asks learners to engage in critical reflection. Instructors focus on fostering active learning by encouraging students to thinking critically and creatively – a capacity that lasts beyond college (Dyson, 2010). Regarding the current practice, GBG sought to reframe students' perspective on grading and center students' learning and reflection, enhancing their ability to self-evaluate their development. Constructivist learning and transformative learning view learning as a process to be engaged in relationally and through critical self-reflection and metacognition. All involved must be active participants in the process of learning.

Positionality

In the fall of 2019, I taught my first class. Throughout the semester, we discussed topics like listening, empathy and values. During the fall of 2020, I taught the class again, this time

including topics such as growth mindset and vulnerability and tweaking the exams to be open book, open notes and applied to a student's context (e.g. an on-campus organization). However, the assessment tweaks felt insufficient; the more I taught this course, the more uncomfortable I became with utilizing a traditional grading structure – sensing incongruity between the content I was teaching and the way I was assessing learning. While I worked to promote “progress over perfection” and intellectual risk taking, the grading mechanisms (e.g. evaluating student reflections on a ten-point scale) appeared contradictory to the content.

By the end of the 2021 spring semester, it was clear that traditional grading structures were not serving my students' learning, nor were they serving me, the instructor. I was spending more time meeting with students about their grades than meeting with students about their learning – discussing the nuances of an A or an A– rather than discussing the nuances of topics such as rational compassion. I was not the first to feel discontent with traditional grading; educators have been querying about the harms of grading since the system began. After reading *Ungrading* by Blum (2020a), an edited book on non-traditional grading with chapter authors spanning areas of expertise and ages of learners, as well as consulting a handful of published experts (i.e. journal articles) and practical experts (i.e. former students), I drafted a non-traditional grading system, describing it as “growth-based grading,” given its emphasis on students self-assessing their learning. The remainder of the current paper discusses the implementation and evaluation of GBG.

Description of the practice

Context

GBG was implemented with two classes in the fall of 2021 ($n = 97$), one class in the spring of 2022 ($n = 30$) and two classes in the fall of 2022 ($n = 46$) across two large, land-grant institutions. The two institutions span regions (i.e. Midwest and Mid-Atlantic), and both are predominantly white institutions. Students ranged from first year to senior standing, and most were traditional college students. Students also spanned a diverse array of majors and colleges, although all classes were taught within colleges of agriculture and departments of agricultural leadership.

The two classes in the fall of 2021, one class in the spring of 2022 and one of the classes in the fall of 2022 had similar curriculum focused on interpersonal development with topics including listening, vulnerability, growth mindset, verbal and non-verbal communication, values, perceptions, inclusion and conflict. Learning objectives included: (1) Develop and/or enhance personal skills in leadership through critical reflection and (2) Make decisions that reflect personal values. The other class in the fall of 2022 was a public speaking course that employed a variety of speeches (i.e. narrative, persuasive, demonstration, debate and special occasion) within the context of agriculture. Learning objectives included: (1) Develop interpersonal communication skills and (2) Describe the different types of speeches and their intended purposes. All the classes utilized a small group structure where students sat in assigned teams and worked through questions and activities as a group. The classes in the fall of 2021 and spring of 2022 had Undergraduate Teaching Assistants that led small groups and provided feedback on students' assignments (i.e. reading reflections and reflection journals). Notably, some of the courses were reserved for a specific student population. One of the classes in the fall of 2021 was a special section for college students mentoring K-12 youth; the other class in the fall of 2021 was largely reserved for students in their first semester of college. The course in the spring of 2022 was a special section for students involved in a large, comprehensive college transition program with most students identifying as first-generation. All courses were options within the university's undergraduate core curriculum and taken primarily to fulfill the core requirements.

Syllabus and structure

Inspired by the work of Cheney (2019, August 11), the syllabi in all five courses sought to be welcoming and learner-centered, including sections like “what you will gain from this course”;

“expectations for me, the instructor”; “course assumptions” and “how I can help you succeed”, which included information on topics such as accessing accommodations for accessibility, mental health resources and on-campus support for basic needs. Assignments were assessed as complete/incomplete (i.e. without points, numbers or a letter grade) with feedback from the teaching team, and there were no exams. Additionally, students had the opportunity to redo one incomplete assignment (i.e. an assignment that was not turned in on time or that was marked incomplete because it did not fulfill the requirements).

One section of the syllabus was devoted to the topic of grading. The grading component stated:

We will discuss topics like growth mindset and focus on progress over perfection. To have congruence between the content and the grading structure, we will use a “growth-based” grading system. The goal of this system is to decenter grades and, instead, center learning. . .we’re shifting away from a traditional grading structure to ease stress and make more room for curiosity. Within our “growth-based” grading system, you will self-assess your growth through two portfolios and, subsequently, portfolio conferences. At midterms and finals, you will complete a portfolio document demonstrating your growth in the course. In the midterm portfolio, you will set goals for the rest of the semester. You will meet with [Author] for five minutes to go through the portfolio. In the final portfolio, you will suggest the grade you think you have earned and then meet with [Author] again for five minutes. We will discuss and decide upon your final grade together.

Drawing heavily from the reflection structure utilized by [Blum \(2020c\)](#), the portfolio document encouraged metacognition by guiding students through a series of prompts at the midterm and final. The midterm portfolio asked students to reflect on a variety of course aspects (i.e. reflection journals, class content, additional assignments, participation and self-evaluated grade).

Building on the midterm portfolio, the final portfolio had seven sections. The first section encouraged students to reflect on specific components of the course (e.g. “Choose 3 of your favorite or best or most fabulous paragraphs [or statements for those who did videos] from your reflection journals and type them in here. Why did you choose them?”; “Compare and contrast the first half of the semester to the second. When and how did you grow?”). In the second section of the final portfolio, students were asked to make general reflections about their experience in the course (e.g. “What do you know now that you did not know in August?”; “If you were going to write a leadership myth – something people generally believe, but which is false – what would it be? Why?”; [Blum, 2020c](#)). The third section focused on student feedback, including how students utilized the course content in their interactions with each other (e.g. “What have you learned about the way you learn?”; “What was your initial reaction to the lack of focus on grades? How do you feel about it now? How, if at all, did it impact your learning?”).

In the fourth section, students answered questions based on the “numbers” (e.g. “How many classes did you miss?”; “How many missing or incomplete assignments do you have?”). In the fifth section, students self-evaluated their in-class participation (e.g. “I had read and was prepared to discuss assigned readings”; “I engaged in class activities and discussions with authenticity”). The last section of the final portfolio asked students to suggest a grade for themselves utilizing a rubric (see [Figure 1](#); [Chiaravalli, 2020](#)). Students were guided to select a grade and defend their selection based on the rubric and their performance in the course. After students completed their midsemester and final portfolios, I met with them individually in a portfolio conference to discuss their performance in the course and their self-evaluated grade, indicating whether I, the instructor, agreed or disagreed with their assessment and why. Notably, I reserved the right to alter students’ self-assessment of their final grade in line with the grading rubric as appropriate. During the portfolio conferences, I also highlighted students’ strengths in the course, such as positive peer relationships, enhanced effort on particular assignments and application of course concepts to real-life situations.

Please suggest a grade for yourself with justification.

Descriptive Grading Criteria:
Use phrasing from the Descriptive Grading Criteria throughout your explanation with *evidence* from above and your work. Please respect your classmates by evaluating yourself honestly.

<i>A</i>	• Demonstrates a high level of understanding of concepts on reading reflections • Shows frequent evidence of growth, turning weaknesses to strengths • Consistently incorporates feedback in reactional journals, challenging oneself to grow • Self-assessment of participation averaged ≥ 4.5 • Attended class almost always (one absence or less) • Turned in almost all assignments (has one incomplete or less)
<i>B</i>	• Demonstration of a good grasp on concepts on reading quizzes on reading reflections • Shows some evidence of growth on reaction journals • Sometimes incorporates feedback in reaction journals • Self-assessment of participation averaged ≥ 4.0 • Attended class most of the time (two unexcused absence NOTE: missing more than three classes may result in a failure of the course [see syllabus]) • Turned in most assignments (has 2-3 incompletes)
<i>C</i>	• Demonstrates a satisfactory acquisition of concepts on reading reflections • Shows little evidence of growth on reaction journals • Rarely incorporates feedback in reaction journals • Self-assessment of participation averaged ≥ 3.5 • Attended class some of the time (2-3 unexcused absences, NOTE: missing more than three classes may result in a failure of the course [see syllabus]) • Turned in some assignments (has 4-5 incompletes)
<i>D</i>	• Demonstrates little acquisition of concepts on reading reflections • Shows little evidence of growth on reaction journals • Rarely incorporates feedback in reaction journals • Self-assessment of participation averaged < 3.0 • Has four or more unexcused absences (NOTE: missing two classes may result in a failure of the course [see syllabus]) • Turned in some assignments (has 6-7 incompletes)
<i>Incomplete</i>	• Does not show satisfactory acquisition of concepts on reading reflections • Provides too little evidence of learning to decide • Shows no evidence of growth on reaction journals • Does not incorporate feedback in reaction journals • Self-assessment of participation averaged < 3.5 • Has five or more unexcused absences • Has eight or more incomplete assignments

- a. Grade:
- b. Explain how you arrived at this grade using the above criteria:

Figure 1. Section six of the final portfolio: grading rubric. Source: Figure by author

Introducing GBG

On the first day of class each semester, I set aside 30 minutes to discuss the grading structure because the structure is unfamiliar to most students and can be a source of apprehension (Sorensen-Unruh, 2020). Even while college students note the stress and anxiety that results from traditional grading, they often feel comfortable with the traditional grading structure (Stommel, 2020). In line with the transformative learning theory (Mezirow, 1978, 2008), I introduced GBG through critical reflection on students’ emotions and experiences regarding the term “grades.” To introduce the topic of grades, I followed the process of Kirr (2020) and asked students to write down the thoughts that came to mind when I said “grades.”

After students finished writing their thoughts, I asked them to identify their sticky note as largely as “positive,” “negative,” or “both/neutral” and then place them in groups around the classroom. Repeatedly, most sticky notes landed in the “negative” pile. Examples of “negative” sticky notes include “pressure,” “stress!!!,” “grades do not = comprehension” and

“narrow/limiting.” Examples of “both/neutral” sticky notes include, “content,” “hoping the assignments are not too difficult” and “rubrics/points.” Examples of “positive” sticky notes include “progress and success,” “demonstration of effort” and “see my growth.” After debriefing students’ sticky note responses, I shared research on grading and student learning and asked students to reflect on how the research does or does not align with their perceptions and experience. After overviews of the research, I shared aspects of GBG, including portfolio conferences. Next, I offered tips and expectations, which include the importance of attending class and turning in assignments, as well as the notion of being growth-oriented versus achievement-oriented. Finally, students asked remaining questions.

In sum, the process of grading (i.e. the portfolios and self-evaluations) became a mechanism for student learning, aligning with course content and pedagogy. Specifically, the act of metacognition, or thinking about thinking, benefitted the process of learning “as individuals increase awareness of personal attitudes, beliefs, and values” (Lumpkin & Achen, 2019, p. 98). Komives, Dugan, Owen, Slack, and Wagner (2011) regarded contemplative practice, the purpose of the portfolios, as one of seven powerful pedagogies. Blum (2020b) notes that students who make a case for their final grade show heightened levels of reflection and comprehension.

Evaluation methods and results

Participants

GBG was evaluated with a survey. The study was approved by the Institutional Review Board at both institutions, and anonymity was emphasized. Survey questions were based on previous research comparing traditional and non-traditional grading systems (Kohn, 1999). Data were anonymously collected at the end of each semester. The survey included demographic questions, quantitative questions and open-ended questions. Data from the five courses were collected and aggregated. Of the 173 undergraduate students enrolled in the five courses utilizing GBG, 83 participated in the survey. Demographic data are presented in Table 1.

Quantitative survey methods and results

The quantitative survey questions were divided into two parts. The first part was about GBG. Participants were given the following prompt: “For each of the following statements, think about the GBG structure utilized in [class] this semester. Then, please select the answer that matches your position most, according to the following scale: 1 – Strongly Disagree; 2 – Disagree; 3 – Undecided; 4 – Agree; 5 – Strongly Agree.” The second part of the quantitative questions was about traditional grading. Participants received the following prompt: “For each of the following statements, think about a traditional grading structure (‘with the traditional grading system, many elements are combined to determine [a] grade – test scores, quizzes, homework, participation, coming to school on time, extra credit – then, the average of the semester’s work equates into a percentage that correlates with a letter grade’ (Stringer, 2014)). Please select the answer that matches your position most, according to the following scale: 1 – Strongly Disagree; 2 – Disagree; 3 – Undecided; 4 – Agree; 5 – Strongly Agree.” The questionnaire items, frequencies and averages are presented in Table 2. Traditional grading numbers are in parentheses.

The data from part one on GBG were compared to the data from part two on traditional grading to determine if participant responses revealed a significant difference in perception between the two grading structures. The data were compared using the Wilcoxon signed-rank test, which is the nonparametric equivalent to the paired samples *t* test. The Wilcoxon signed-rank test was selected because it compares the means of items taken from the same individual and does not assume normality in the data. Sample size was sufficient ($N = 83$; Dwivedi, Mallawaarachchi, & Alvarado, 2017). Results of the Wilcoxon signed-rank test are reported in Table 3. There was a statistically significant difference on all items with participants marking GBG more favorably than traditional grading.

Table 1. Demographic data

Demographic	N	Percent
Gender		
Female	57	68.7%
Male	26	31.3%
Location		
Mid-Atlantic University	18	21.7%
Midwestern University	65	78.3%
Race/Ethnicity		
American Indian or Alaska Native	1	1.2%
Asian	4	4.8%
Black or African American	6	7.2%
Hispanic or Latinx	11	13.3%
White	61	73.5%
Year in School		
First Year	24	28.9%
Second Year	37	44.6%
Third Year	19	22.9%
Fourth Year	3	3.6%
College		
Agriculture	27	32.5%
Architecture	1	1.2%
Arts and Sciences	14	16.9%
Business	23	27.7%
Engineering	9	10.8%
Natural Resources and Environment	6	7.2%

Source(s): Table by author

Open-ended response methods and results

In addition to the quantitative component of the survey, participants also completed three open-ended response questions: (1) “In comparison to traditional grading, what ways, if any, did GBG enhance your learning?”; (2) “In comparison to traditional grading, what ways, if any, did GBG hinder your learning?” and (3) “What changes would you recommend to the GBG structure to maximize student learning?” Participant responses were analyzed using open coding to segment the data into categories (Creswell & Poth, 2018).

When asked to share what ways, if any, GBG enhanced their learning, participants ($N = 67$) articulated that GBG allowed them to focus on content, growth and learning ($n = 42$); emphasized creativity and individualization ($n = 11$); helped them enjoy class by reducing stress ($n = 6$); encouraged interaction with classmates ($n = 5$) and supported trying new things ($n = 3$; see Table 4). Two participants did not think growth-based growth enhanced their learning.

When asked to share what ways, if any, GBG hindered their learning, participants ($N = 65$) articulated that GBG decreased motivation ($n = 14$); was ambiguous about student performance in class ($n = 8$); required effort from students and instructors ($n = 3$); caused confusion at the beginning ($n = 3$) and could be subjective ($n = 2$; see Table 5). Thirty-five participants did not think growth-based growth hindered their learning.

When asked to share what changes they would recommend to the GBG structure to maximize student learning, participants ($N = 64$) articulated that the GBG structure could be enhanced by providing more feedback ($n = 15$); ensuring accountability ($n = 4$); giving partial credit on assignments ($n = 3$) and further explaining GBG at the start of the semester ($n = 3$; see Table 6). Twenty-seven participants did not have recommendations for changing the structure. Seven participants offered general feedback on the class that was not related to GBG (e.g. shorter class sessions). One participant responded with “make it traditional grading.”

Table 2. Questionnaire items and respondent averages for GBG and (traditional grading)

Question	Strongly disagree	Disagree	Undecided	Agree	Strongly agree	GBG (traditional grading) average
[Grading structure] increased my interest in learning	2 (12)	1 (30)	4 (21)	15 (8)	54 (5)	4.55 (2.53)
[Grading structure] increased my preference for challenging tasks	1 (20)	2 (24)	7 (19)	21 (9)	45 (4)	4.41 (2.38)
[Grading structure] increased my motivation	1 (11)	4 (14)	7 (18)	22 (25)	42 (8)	4.32 (3.07)
[Grading structure] increased the quality of my thinking	0 (19)	3 (22)	4 (26)	16 (5)	53 (4)	4.57 (2.38)
[Grading structure] was valid, reliable and objective	2 (11)	3 (15)	7 (19)	19 (21)	45 (9)	4.34 (3.03)
[Grading structure] enhanced the curriculum	1 (20)	2 (23)	5 (23)	23 (6)	45 (4)	4.43 (2.36)
[Grading structure] wasted time that could have been spent on learning*	3 (10)	3 (14)	5 (15)	15 (27)	50 (10)	4.39 (3.17)
[Grading structure] encouraged cheating*	3 (7)	1 (4)	1 (8)	8 (29)	63 (27)	4.67 (3.87)
[Grading structure] spoiled your relationship with your teacher*	4 (11)	2 (12)	4 (24)	8 (20)	58 (9)	4.50 (3.05)
[Grading structure] spoiled your relationship with other students*	2 (13)	3 (17)	3 (25)	9 (16)	59 (5)	4.59 (2.78)

Note(s): Numbers in parentheses are the traditional grading responses. Starred (*) items were “negatively” phrased whereas the other items were “positively” phrased and, therefore, have been reverse coded, 1 – Strongly Disagree; 2 – Disagree; 3 – Undecided; 4 – Agree; 5 – Strongly Agree

Source(s): Table by author

Table 3. Results of Wilcoxon signed-rank test comparing growth-based and traditional grading

Question	Z	p
Traditional grading increased my interest in learning	−6.76	<0.001
Traditional grading increased my preference for challenging tasks	−6.64	<0.001
Traditional grading increased my motivation	−5.23	<0.001
Traditional grading increased the quality of my thinking	−7.04	<0.001
Traditional grading was valid, reliable and objective	−5.09	<0.001
Traditional grading enhanced the curriculum	−6.70	<0.001
Traditional grading wasted time that could have been spent on learning*	−6.05	<0.001
Traditional grading encouraged cheating*	−6.80	<0.001
Traditional grading spoiled your relationship with your teacher*	−5.94	<0.001
Traditional grading spoiled your relationship with other students*	−5.68	<0.001

Note(s): Starred (*) items were “negatively” phrased whereas the other items are “positively” phrased. The “negatively” phrased items were computed based on negative ranks, while the “positively” phrased items were computed based on positive ranks

Source(s): Table by author

In sum, the overarching evaluation findings revealed a positive perception of GBG. The quantitative results demonstrated a significant difference between opinions of GBG and opinions of traditional grading related to student learning with participants perceiving GBG more favorably than traditional grading. In response to open-ended survey questions,

Table 4. Themes for GBG enhancing learning

Theme	# of responses	Evidence
Focusing on content, growth and learning	42	“I was able to take the time to learn rather than memorize or meet rubrics. I was naturally remembering based of my understanding rather than feeling forced to know something” “It made me want to produce high quality work and relieved the pressure of being right” “I felt motivated to learn out of genuine interest compared to feeling the need to get a good grade”
Creativity and individualization	11	“It offered more room for exploration. It also allowed me to explore the topics more as I chose to, not as the teacher deemed necessary” “Growth based learning allowed me to focus on the creative side of class and not just studying for tests”
Enjoying class and reducing stress	6	“Reduced stress” “I think GBG made sure I was actually paying attention while in class and that I was engaged”
Interacting with classmates	5	“It enhanced it because I actually felt like I was learning by collaborating with other students”
Trying new things	3	“I wanted to participate in class regardless of the grade” “I was able to practice, fail, and come back stronger the next time”
Did not enhance learning	2	“It did not”
Source(s): Table by author		

Table 5. Themes for GBG hindering learning

Theme	# of responses	Evidence
Did not hinder learning	35	“It did not hinder my learning at all, I think all classes should be taught this way” “In [class name], it did not hinder my learning”
Lacking motivation	14	“Growth-based learning decreased my motivation slightly as traditional grading always stresses me out” “It didn’t push me to learn all the content as strictly, but I was able to apply it instead of just know it for a test” “I do think that I able to slack on an assignment or two because they were all weighted the same”
Ambiguity about performance	8	“Yes, as I had no specific measure of how I was doing. It created so much more stress for me because I became overly critical of myself” “When compared with traditional grading, sometimes it felt very difficult to gauge my position in the class and if I needed to be worried about my grade”
Requiring effort	3	“GBG won’t work if someone does not want to learn/is not open to challenge their own perspective”
Uncertainty about structure	3	“At first, it was hard to understand the criteria for each grade for GBG compared to traditional grading”
Subjective	2	“I think there is a slight lack of accountability and even the opportunity for misplaced grading (since subjectivity can be inconsistent)”
Source(s): Table by author		

Table 6. Themes for recommendations to change GBG structure

Theme	# of responses	Evidence
None	27	"I haven't been in any other growth-based classes so I don't really know what you can change" "For a class like that it is great, but I would not do the same for a science or math class"
More feedback	15	"Making check-ins more common in order that we know what to work on more clearly throughout the course" "I would recommend giving a grade on the [reflection assignments] to give the students a baseline to improve upon but still give full credit regardless. This way, there is a way to show their improvement" "Maybe weekly evaluations to stay on track"
Accountability	4	"More accountability in the sense of meeting the few due dates and deadlines expected" "I think GBG can sometimes make it hard to motivate students. The students must remain challenged and in an encouraging, passionate environment"
Partial credit	3	"Have partial credit- instead of an all or nothing approach, try having each aspect of the rubric worth one point and then discussing at the end for the grade based on growth. . . the stress the 0 [incomplete] causes hinders learning"
Further explanation	3	"I think giving the grading criteria early on can help students know what to stick to and find ways to take risks in class"

Source(s): Table by author

participants expressed that GBG reduced stress while enhancing their focus on class content, learning and creativity. Participants also shared critiques of and suggestions for GBG, including providing more feedback throughout the semester on their performance.

Reflections and recommendations

Overarchingly, the transition to GBG has brought energy and connection to the students, our classroom and me. As with others who have implemented forms of non-traditional grading, I spend less time on tasks such as discussing the difference between an 8/10 and a 9/10 on a reflection paper and answering emails about grading concerns. Rather, I find myself spending more time on tasks such as talking with students about how they can make their reflections more meaningful, even if it will not affect their grade, and learning about their lives outside of the classroom. Of course, there are still some interactions about late assignments and absences, but they are less frequent.

Further, a particular source of fulfillment as a leadership educator has been final portfolio conferences. Rather than devoting my final weeks of the semester to leaving comments on papers via a learning management system, which many students ignore (Wiliam, 2014, November 29), I get to spend one-on-one time fostering connection and providing feedback to students. Not only do the final portfolio conferences save time when compared to providing written feedback, I end the semester feeling energized from the personal interactions. A continued source of learning is that GBG has required me to become more comfortable receiving feedback (or at least more willing – it has yet to feel comfortable) to ensure that the structure is truly learner-centered. I have collected midsemester feedback in each course, asking students to identify things we should keep doing, start doing or stop doing to support their learning. In response to students' ideas, we adjust the course (e.g. time to work within pairs within their small groups).

As I work to value my students' growth, so, too, I work to value my own growth. The way I implement non-traditional grading now is different than how I implemented it in the fall of

2021. During my first semester of implementing the GBG structure, I discovered that some students heard “growth-based grading” and equated it with a lack of structure, which created confusion. Therefore, I began highlighting during the first day of the semester the importance of attendance due to the interactive nature of the course, as well as turning of assignments on time so that the reflection is completed in proximity to content. Additionally, I found that for some students the portfolio conferences were a source of apprehension because of their unfamiliarity. In response, I discuss the structure of the portfolio conference at the start of the semester and before both conferences. Next, I deliberately highlight the importance of dialogue about questions related to the grading structure throughout the semester by emphasizing that I am not a mind reader and need students to share their questions and confusions related to the grading structure, as well as class in general, so that I can help. Finally, in classes with Undergraduate Teaching Assistants, the teaching team practices providing feedback on students’ assignments to sharpen each other’s capacity and ensure students receive quality, in-depth feedback.

In response to the current findings, I added a “one month check-in” for students to ask questions related to GBG and discuss their performance and grade during the spring of 2023. When working with Undergraduate Teaching Assistants, I put additional effort into ensuring that students talk with me about topics related to their grade (e.g. attendance) to promote consistency of implementation and clarity. Finally, I now share the final portfolio grade rubric at the start of the semester to decrease uncertainty.

Implications

The GBG structure presented in the current manuscript, as well as the survey results, have three primary implications for leadership educators: (1) consider grading structures as a component of pedagogy; (2) analyze our grading structures for congruence or incongruence with our identified course outcomes, class content and values and (3) recognize that there are a variety of ways to incorporate non-traditional grading into leadership education classrooms.

First, participants’ perceptions of the connection between grading and learning provide preliminary evidence that grading ought to be considered an essential component of pedagogy in leadership education. When discussing instructional strategies and assessment processes, grading structures ought to be included in the conversation given their perceived impact on student learning, motivation and engagement. Recognizing grading as pedagogy is a step towards critically reflecting on how grading practices enhance or hinder student learning.

Second, leadership educators are encouraged to consider if and how grading practices align with and contribute to learning outcomes, pedagogy and course content. Given that leadership educators seek to enhance students’ capacity for critical thinking, reflection, application, self-understanding and empowerment (Andenoro, 2013; CAS, 2015; Seemiller, 2016), interrogating grading structures is a worthy and essential pursuit. For example, leadership educators may consider the following questions:

- (1) Why do we grade?
- (2) Why do we grade the way that we grade?
- (3) How, if at all, do our grading practices align with our course content and pedagogy?
- (4) How, if at all, does our grading system contribute to student learning?
- (5) What alternate grading practices or subtle shifts to our grading might we consider that align with our course content, pedagogy and beliefs about leadership education?

The field of leadership education recognizes that effective signature pedagogies promote student participation, encourage student engagement and foster a learning environment where students are seen and heard (Jenkins, 2012). The results of the current study encourage us to ask: are the signature pedagogies of leadership education reflected in grading structures?

Stated another way, do grading structures promote student participation, encourage student engagement and foster a learning environment where students are seen and heard? Backwards design presents one method of enhancing intentionality in grading. Described as necessary for effective curriculum (Wiggins *et al.*, 2005), backwards design establishes learning outcomes and desired results before building curriculum. The process of backwards design can also be applied to assessment and grading. By intentionally considering the desired outcomes of courses before implementing an assessment and grading process, leadership educators may be better equipped to have congruence between outcomes, pedagogy, content and grading structures.

Third, it is necessary to acknowledge that there are a variety of ways to incorporate non-traditional grading into leadership education classrooms. While utilizing GBG was a good match for the content and pedagogy of the courses in the current article, other leadership educators might prefer different non-traditional grading structures, such as contract grading or mastery-based systems. Even when educators use a similar method of non-traditional grading to GBG, specific practices may vary. For example, I tend to be firm on deadlines and attendance. Not all non-traditional graders align with this way of thinking. Some prefer to eliminate firm deadlines and attendance policies and allow resubmission of all assignments to further center student choice. There are those who, justifiably, go “further” than I do or are more radical in moving away from traditional grading, only allowing their classes to be taken as pass/no pass. On the other hand, there are those who, justifiably, do not go as “far” as I do in moving away from traditional grading, retaining aspects of traditional grading (e.g. reading quizzes or exams) while incorporating aspects of non-traditional grading (e.g. providing only feedback on written reflections). Leadership educators may find Nilson, Clark, and Talbert’s (2023) text, *Grading for Growth*, useful to explore a variety of alternative grading systems. Of note, while Nilson *et al.*’s (2023) text and GBG are similarly titled, Nilson *et al.* (2023) do not discuss GBG, as outlined in the current article, as an alternative grading system.

Future research and limitations

The current article is limited by its context and approach. Regarding context, GBG was implemented by one instructor at two universities and, therefore, the results are not generalizable to all instructors at all universities. Correspondingly, I encourage leadership education scholars to build on the current article by exploring grading structures within a variety of contexts (e.g. graduate courses and non-traditional students). Regarding approach, the current evaluation focused only on students’ perceptions of their experience with two grading approaches. Future research is encouraged to measure knowledge acquisition and capacity growth in relation to grading structures with grade earned as a control variable. Future research may specifically utilize experimental or quasi-experimental research designs to examine causal relationships between assessment practices/grading structures and key outcomes (e.g. leadership capacity growth).

Additionally, there is tension in critiquing the use of numbers to grade learning while also using numbers to partially evaluate a grading system. Therefore, I recommend that future scholars and practitioners explore the effectiveness of grading structures through alternative qualitative methods, such as narratives, which elucidate depth of reflection and foster meaning making among participants (Creamer, Magolda, & Yue, 2010). It is also recommended that advanced quantitative methods with larger sample sizes be employed to research grading practices in leadership education. Further, I encourage leadership educators who implement non-traditional grading to disseminate the details of their practice and to explicitly analyze the connections between metacognition, grading and learning. Finally, beyond analyzing how teachers evaluate students, there is work to be done on how students evaluate teachers, particularly given that higher evaluations are correlated with more generous grading (Bunge, 2018, November 27; Ewing, 2012). Future scholars are encouraged to explore how various evaluation practices of both teachers and students impact learning and classroom environments.

In conclusion, the current pedagogy paper described the development, implementation and evaluation of GBG, a method of non-traditional grading. Developed in response to the incongruence I felt between course content/outcomes and traditional grading structures, GBG emphasizes students' self-assessment of their learning. Evaluation results revealed a positive perception of GBG in comparison to traditional grading among students, while also offering suggestions for improvement. In response, leadership educators are urged to regard grading as central to pedagogy and consider if our assessment structures align with our content, learning objectives and pedagogical beliefs.

References

- Alcocer, P. (2019). History of standardized testing in the United States. *National Education Association*, 66(1), 13–19.
- Allen, S. J., & Hartman, N. S. (2008). Leadership development: An exploration of sources of learning. *SAM Advanced Management Journal*, 73(1), 10.
- Allen, S. J., & Hartman, N. S. (2009). Sources of learning in student leadership development programming. *Journal of Leadership Studies*, 3(3), 6–16. doi: [10.1002/jls.20119](https://doi.org/10.1002/jls.20119).
- Andenoro, A. C. (2013). The national leadership education research agenda: Strategic priorities and deepened perspectives. *Journal of Leadership Education*, 12(3), 1–9. doi: [10.12806/V12/I1/C1](https://doi.org/10.12806/V12/I1/C1).
- Barshay, J. (2022). Proof points: Research on increasing diversity in college admissions. The Hechinger Report. Available from: <https://hechingerreport.org/proof-points-research-on-increasing-diversity-in-college-admissions/>
- Blum, S. D. (2020a). *Ungrading: Why rating students undermines learning (and what to do instead)*. Morgantown, WV: West Virginia University Press.
- Blum, S. D. (2020b). Introduction: Why grade? Why ungrade?. In S. D. Blum (Ed.), *Ungrading: why rating students undermines learning (and what to do instead)* (pp. 1–24). Blackwell.
- Blum, S. D. (2020c). Just one change (just kidding): Ungrading and its necessary accompaniments. In S. D. Blum (Ed.), *Ungrading: why rating students undermines learning (and what to do instead)* (pp. 54–73). Blackwell.
- Bright, D. S., Caza, A., Turesky, E. F., Putzel, R., Nelson, E., & Luechtefeld, R. (2016). Constructivist meta-practices: When students design activities, lead others, and assess peers. *Journal of Leadership Education*, 15(4), 75–99. doi:[10.12806/V15/I4/R6](https://doi.org/10.12806/V15/I4/R6).
- Bunge, N. (2018). Students evaluating teachers doesn't just hurt teachers. *It hurts students*. Washington, DC: The Chronicle of Higher Education. Available from: <https://www.chronicle.com/article/students-evaluating-teachers-doesnt-just-hurt-teachers-it-hurts-students/>
- Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology*, 58(1), 1–14.
- Cheney (2019). *Toward cruelty-free syllabi [power point slides]*. Google Slides. Available from: <http://bit.ly/cfsyllabi>
- Chiaravalli (2020). The point-less classroom. A math teacher's ironic choice in not calculating grades. In Blum, S. D. (Ed.), *Ungrading: why rating students undermines learning (and what to do instead)*. Blackwell (pp. 161–172).
- Council for the Advancement of Standards in Higher Education (2015). CAS learning and development outcomes. In Wells, J. B. (Ed.), *CAS professional standards for higher education* (9th ed.). Washington, DC: Author.
- Creamer, E. G., Magolda, M. B., & Yue, J. (2010). Preliminary evidence of the reliability and validity of a quantitative measure of self-authorship. *Journal of College Student Development*, 51(5), 550–562. doi: [10.1353/csd.2010.0010](https://doi.org/10.1353/csd.2010.0010).

-
- Cress, C. M., Astin, H. S., Zimmerman-Oster, K., & Burkhardt, J. (2001). Developmental outcomes of college students' involvement in leadership activities. *Journal of College Student Development*, 42(1), 15–27.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches*. Thousand Oaks, CA: Sage.
- Dwivedi, A. K., Mallawaarachchi, I., & Alvarado, L. A. (2017). Analysis of small sample size studies using nonparametric bootstrap test with pooled resampling method. *Statistics in Medicine*, 36(14), 2187–2205. doi: [10.1002/sim.7263](https://doi.org/10.1002/sim.7263).
- Dyson, M. (2010). What might a person-centred model of teacher education look like in the 21st century? The transformism model of teacher education. *Journal of Transformative Education*, 8(1), 3–21.
- Elman, B. A. (2000). *A cultural history of civil examinations in late imperial China*. Berkeley, CA: Univ of California.
- Ewing, A. M. (2012). Estimating the impact of relative expected grade on student evaluations of teachers. *Economics of Education Review*, 31(1), 141–154. doi: [10.1016/j.econedurev.2011.10.002](https://doi.org/10.1016/j.econedurev.2011.10.002).
- Golden, B. (2023). Enabling critical thinking development in higher education through the use of a structured planning tool. *Irish Educational Studies*, 42(4), 949–969. doi:[10.1080/03323315.2023.2258497](https://doi.org/10.1080/03323315.2023.2258497).
- Hayek, A. S., Toma, C., Oberlé, D., & Butera, F. (2015). Grading hampers cooperative information sharing in group problem solving. *Social Psychology*, 46(3), 121.
- Jenkins, D. M. (2012). Exploring signature pedagogies in undergraduate leadership education. *Journal of Leadership Education*, 11(1), 1–27.
- Jenkins, D. M. (2016). Teaching leadership online: An exploratory study of instructional and assessment strategy use. *Journal of Leadership Education*, 15(2), 129–149.
- Jenkins, D. M. (2018). Comparing instructional and assessment strategy use in graduate-and undergraduate-level leadership studies: A global study. *Journal of Leadership Education*, 17(1), 73–92.
- Kirr, J. (2020). Conference musings and the g-word. In S. D. Blum (Ed.), *Ungrading: why rating students undermines learning (and what to do instead)*. Blackwell (pp. 188–203).
- Knowles, M. S., Holton, E. F., & Swanson, R. A. (1998). *The adult learner*. Gulf.
- Kohn, A. (1999). From degrading to de-grading. *High School Magazine*, 6(5), 38–43.
- Kohn, A. (2006). The trouble with rubrics. *English Journal*, 95(4), 12–15.
- Komives, S. R., Dugan, J. P., Owen, J. E., Slack, C., & Wagner, W. (2011). *The handbook for student leadership development*. San Francisco, CA: John Wiley & Sons.
- Lipnevich, A. A., & Smith, J. K. (2008). Response to assessment feedback: The effects of grades, praise, and source of information. *ETS Research Report Series*, 2008(1), 1–57.
- Lumpkin, A., & Achen, R. M. (2019). Leadership education: Assessment of learning in a sport leadership course. *Journal of Leadership Education*, 18(2), 95–110. doi:[10.12806/V18/I2/R7](https://doi.org/10.12806/V18/I2/R7).
- Mezirow, J. (1978). Perspective transformation. *Adult Education*, 28(2), 100–110.
- Mezirow, J. (2008). An overview on transformative learning. In Illeris, K. (Ed.), *Contemporary theories of learning* (pp. 40–54). London: Routledge.
- Nilson, L. B., Clark, D., & Talbert, R. (2023). *Grading for growth: A guide to alternative grading practices that promote authentic learning and student engagement in higher education*. Routledge.
- Odom, S. F. (2015). Undergraduate student perceptions of the pedagogy used in a leadership course: A qualitative examination. *Journal of Leadership Education*, 14(2), 17–29.
- Phillips, D. C. (1995). The good, the bad, and the ugly: The many faces of constructivism. *Educational Researcher*, 24(7), 5–12. doi: [10.3102/0013189X024007005](https://doi.org/10.3102/0013189X024007005).
- Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures*. (Trans A. Rosin). New York, NY: Viking.

- Ramsey, V. J., & Fitzgibbons, D. E. (2005). Being in the classroom. *Journal of Management Education*, 29(2), 333–356. doi: [10.1177/1052562904271144](https://doi.org/10.1177/1052562904271144).
- Seemiller, C. (2016). Assessing student leadership competency development. *New Directions for Student Leadership*, 2016(151), 51–66. doi: [10.1002/yl](https://doi.org/10.1002/yl).
- Sorensen-Unruh, C. (2020). A STEM ungrading case study: A reflection on first-time implementation in organization chemistry II. In S. D. Blum (Ed.), *Ungrading: why rating students undermines learning (and what to do instead)* (pp. 140–160). Blackwell.
- Starch, D., & Elliott, E. C. (1912). Reliability of the grading of high-school work in English. *The School Review*, 20(7), 442–457.
- Stommel, J. (2020). How to ungrade. In S. D. Blum (Ed.), *Ungrading: why rating students undermines learning (and what to do instead)* (pp. 25–41). Blackwell.
- Stringer, C. (2014). Standards-based grading vs. traditional grading using Missouri MSIP scores, (Doctoral dissertation). Available from: <https://www.nwmissouri.edu/library/researchpapers/2014/Stringer,%20Christie.pdf>
- Tannock, S. (2017). No grades in higher education now! Revisiting the place of graded assessment in the reimagination of the public university. *Studies in Higher Education*, 42(8), 1345–1357. doi: [10.1080/03075079.2015.1092131](https://doi.org/10.1080/03075079.2015.1092131).
- Wiggins, G., Wiggins, G. P., & McTighe, J. (2005). *Understanding by design* (2nd ed.). Alexandria, VA: ASCD.
- William, D. (2014). *Is the feedback you're giving students helping or hindering*. Learning Sciences International. Available from: <https://www.dylanwilliamcenter.com/2014/11/29/is-the-feedback-you-are-giving-students-helping-or-hindering/>

Corresponding author

Hannah M. Sunderman can be contacted at: hsunderman@vt.edu