Received: 19 July 2024 | Revised: 7 November 2024 | Accepted: 27 November 2024 | Published online: 24 December 2024

RESEARCH ARTICLE

International Journal of Changes in Education
2024, Vol. 00(00) 1–12
DOI: 10.47852/bonviewIJCE42023868

BON VIEW PUBLISHING

How Do Non-traditional Grading Practices Promote Student Learning: Perspectives from Two Mathematics Teacher Educators

Shelli L. Casler-Failing[1,*] and Montana Smithey[2]

[1]*Department of Middle Grades and Secondary Education, Georgia Southern University, USA*

[2]*Department of Elementary and Special Education, Georgia Southern University, USA*

Abstract: This qualitative self-study, conducted as a multiple case study, investigated pre-service teachers (PSTs) experiences with non-traditional grading practices (NTGPs) in two mathematics methods courses taught by the authors – one designed for middle-grade PSTs (grades 4–8) and one designed for elementary PSTs (grades Pk-5). Archival data were analyzed through the lenses of pedagogical content knowledge (PCK), self-efficacy, and growth mindset to investigate how PSTs' experiences with the process of NTGPs supported their learning. Additionally, the course instructors sought to explore which strategies promoted the successful implementation of NTGPs. Data comprised PSTs' assignments/artifacts, reflections, and mid-semester and end-of-semester self-reflections as well as instructors' reflections, data spreadsheets, and notes/debriefs from conversations conducted as part of their critical friendship. Thematic analysis was used to examine data independently by case (each course was identified as an independent case), and then, a cross-case analysis was performed, iteratively repeating for each research question. Throughout each course, feedback on assignments was provided via verbal and written comments to support PSTs' growth in learning and understanding. Key findings describe the development of PSTs' PCK through asking questions, unpacking feedback alongside instructors, and applying feedback to initial submissions. The findings show it is necessary to provide clear expectations, timely and detailed feedback, and access to clear rubrics aligned with learning goals and success criteria. PSTs and instructors must be willing to commit to the time required to successfully implement NTGPs. Additional findings show that PSTs can be supported by NTGPs to improve their self-efficacy and growth mindset to develop their PCK due to the reduced stress felt when instructors remove grades. However, more studies investigating NTGPs are required to support these findings. The authors reiterate that NTGPs are a *work in progress* and reflection is necessary after each implementation to continue to improve one's practices for future students.

Keywords: ungrading, non-traditional grading practices, pre-service teachers, assessment, comment-only feedback

1. Introduction

Methods courses focused on specific content are designed to instruct, via modeling, research-based best practices. Bishop and Harrison [1] propose "instruction . . . [should be] active, purposeful, and democratic" and also provide "[v]aried and ongoing assessments [to] advance learning as well as measure it" to promote student learning. We posit modeling this form of instruction while incorporating non-traditional grading practices (NTGPs) via ungrading [2] and portfolios [3] can focus pre-service teachers' (PSTs) attention on feedback to promote learning and reflection at a deeper level. Investigations into NTGPs in higher education have been reported in various content areas such as ELA, mathematics, art, history, and psychology [4]. Furthermore, this can occur in courses requiring a final numerical or letter grade if the methods employed are consistent and equitable.

As colleagues teaching mathematics methods courses to students majoring in middle-grade education (Casler-Failing) and elementary education (Smithey) – courses requiring a final numerical grade – we chose to implement NTGPs. In making this change, we decided to reflect on our pedagogical practices to determine if our incorporation of NTGPs engaged our PSTs in deeper learning, developed stronger understandings of pedagogical practices, and supported their reflective practice. The focus of this self-study [5] was to analyze archival data and our experiences supporting PSTs' development of pedagogical content knowledge (PCK) [6], self-efficacy [7], and growth mindset [8, 9] through NTGPs. Our research was based upon two assumptions proposed by *The Standards for Preparing Teachers of Mathematics* [10]: "[e]nsuring the success of every learner demands a deep, integrated focus on equity in every program that prepares teachers of mathematics" and "[t]hose involved in mathematics teacher preparation must be committed to improving their effectiveness in preparing future mathematics teachers". Being firmly committed to improving our effectiveness as educators, we felt we could improve our teaching by promoting equitable assessment practices via NTGPs. Additionally, as teacher educators, it is imperative that we model equitable assessment practices in our instruction so PSTs can apply

***Corresponding author:** Shelli L. Casler-Failing, Department of Middle Grades and Secondary Education, Georgia Southern University, USA. Email: scaslerfailing@georgiasouthern.edu

the experiences and practices in their future classrooms. The questions that guided this qualitative self-study were:

1) *What strategies promote the successful implementation of non-traditional grading practices?*
2) *How do non-traditional grading practices support pre-service teachers' learning?*

2. Literature Review

We begin this section by sharing a historical review of the development of grading processes in the United States. We transition from the history of grading to more recent research, sharing literature that informed our decisions to move away from traditional grading practices in our courses. We will conclude this section by describing the theoretical framework that guided our analysis of the archival data reported in this manuscript.

2.1. Where did grades come from?

As Brookhart et al. [11] stated, "grading refers to the symbols assigned to individual pieces of student work or to composite measures of student performance on student report cards". Although grades have always been a part of our academic lives on both the receiving end and the giving end (for many of us), they have not always been a part of the school curriculum in the United States. As universities began to develop in the eastern US and the labor market began to boom in the late 19th century, there became a need to guide students into a particular track – college or work [12, 13]. Grades, once used as tools for internal communication between teachers and families, became a form of external communication for system-building, requiring them to become universal and standardized. Over time, grades shifted from a European, competition-based form of student ranking to a more consistent monthly grade report.

Throughout the early twentieth century, various forms of grading were carried out by school systems, including narrative letters to parents. Nonetheless, by the 1940s, the A-F system was widely used by most K-12 schools and colleges across the US [12]. However, a lack of criteria was a common source of variability in grading policies among institutions nationwide, creating a lack of credibility when transferring grades among schools and universities [11]. Research illuminates that grading systems have progressed over the years, beginning as a means to differentiate students between work and college tracks, then later from progress reports provided during home visits by the teacher in the 19th century to an A-F grading scale in the mid-20th century. Today, various scales are used (e.g., standards-based grading, A-F grading, percentage grading). Over time, the problem has not necessarily become who does the grading but the systems and structures that have altered the assessment goals [13–15]. How can this history inform our current practices? The following sections will share more recent research on grading practices.

2.2. What do we see and hear?

As mathematics teacher educators, we have strong connections to our surrounding school districts. We observe PSTs during field experiences, provide professional development to in-service teachers, and work closely as a liaison with one of our university's Professional Development Schools (Smithey). Through these experiences, we understand traditional grades' heavy emphasis and narrowed focus on student and teacher mindsets. In-service teachers often remind us that they must record a specific number of weekly grades for each student. It is common to hear, "I have to input three formative grades this week", or "I need to submit a summative grade each week". When grades become a "number in a system", they are less likely to be used as assessments of learning or, better yet, as assessments *for* learning [16, 17]. They are just a means to dot an "i" and cross a "t". Furthermore, research has shown that traditional grading systems have become an extrinsic motivator for students rather than an intrinsic motivator [18, 19].

Grades have become a focus of conversations at district meetings and are becoming increasingly prevalent on a state and national level. Grades matter and profoundly impact student, teacher, and administrative outcomes [12, 20]. From a student's perspective, grades determine academic standings (e.g., valedictorian), course options (e.g., AP or honors), scholarship funding, participation in extracurricular activities and sports, and in some instances, poor grades could lead to consequences at home. From a teacher's perspective, grades can impact contract renewals, mandate professional development, or limit teaching options. From an administrator's perspective, grades often reflect performance on standardized exams, which impact state and national school funding and can potentially lead to consequences if students are underperforming.

An even more alarming concern related to collecting student grades is the number of tests students are subjected to each year of their K-12 schooling. A recent report by the Georgia Department of Education stated that the increase in mandated testing is one of the leading reasons for teacher attrition in the state [21]. If teachers leave the profession because they feel burned out due to the amount of testing, what is this doing to students who do not have the option to go elsewhere? Goe et al. [22] claim, "What is measured is a reflection of what is valued, and as a corollary, what is measured is valued". Are we measuring too much? Do we truly value our students and teachers if we value and measure so much that they are burning out?

We propose all students can learn without having their learning connected to a numerical grade. Butler [23] introduced the possibility of removing grades with her research investigating feedback options – grades, grades and comments, comments, praise, or no feedback. Butler's study revealed that student performance improved the greatest when feedback was provided solely through comments, and student performance decreased when feedback was in the form of grades, praise, or no feedback. More importantly, these findings are not an anomaly. Butler's findings have been supported by the research of Bremser [24] and Guberman [25], and similar experiences have been shared by numerous educators, such as Blackwelder [26], Chiaravalli [18], Gibbs [27], and Wiliam [16, 17].

2.3. What does all of this mean?

It is time for teachers and administrators to shift their mindsets when approaching grading and assessment. Assessment should not be viewed as a need to submit a finite number of grades in "the system" each week for each student. Rather, assessment should be viewed as a way to encourage motivation and engagement in learning [17]. Kohn [28] proposes "the more [students'] attention is directed to *how well* they're doing, the less engaged they tend to be with *what* they're doing". Assessment should be viewed as a way to monitor growth in understanding over time, not in a single moment, that is averaged into a set of scores. Maehr and Midgley [29] state, "an overemphasis on assessment can actually undermine the pursuit of excellence". When students receive a poor grade on an assessment, they may regard themselves as "dumb"; they view the

grade as a measure of their character [24]. This perspective can strengthen their fixed mindset [9] and undermine their self-efficacy [7]. Additionally, they may consider themselves to be inferior to their classmates ("Half of the class scored better than me!"). However, removing grades from assessment practices can alleviate classroom comparisons [25].

Bremser's [24] research highlighted the importance of providing metacognitive opportunities while assessing tasks we want students to perform well. Equally important is providing meaningful feedback on those tasks to further their learning [30]. By modeling effective feedback strategies, teacher educators can support PSTs' development of PCK [6]. Feedback is a skill developed over time, and although PSTs will not immediately become experts, they will experience feedback from a learner's perspective that can support their development as future teachers and lay the foundation to implement such practices independently.

Based on existing research and using the lenses of PCK [6], self-efficacy [7], and growth mindset [8, 9], we sought to engage our PSTs in assessment practices that encouraged engagement with meaningful tasks inclusive of metacognitive opportunities guided by clear directions and rubrics [15, 19, 31], participation in reflective practices [2], and the experience of NTGPs.

2.4. Theoretical framework

This self-study [5] was conducted to determine if our NTGPs supported improved learning for our PSTs. As we analyzed the teaching practices implemented throughout each of our courses, we did so through the lenses of PCK [6], self-efficacy [7], and growth mindset [8, 9]. These theoretical lenses blended well and allowed us to connect PSTs' growth to specific skills and/or behaviors.

Shulman [6] proposed the concept of PCK to stress the importance of having deep content knowledge and strong pedagogical skills to support student learning. Shulman proposed, for example, that teaching mathematics requires similar, yet different, instructional approaches than teaching social studies. Furthermore, teaching mathematics at the middle-grade level requires much more content knowledge than teaching mathematics at the elementary level. Additionally, the instructional strategies at each level will vary. Shulman proposed that having strong content knowledge or excellent teaching skills does not alone create a strong educator. It is only when the two domains merge that deep learning can occur. As methods instructors, it is crucial to support our PSTs' development of PCK through modeling and instruction. We propose that self-efficacy can be an important component in that development.

Bandura [7] coined the concept of *self-efficacy* and defined it as an individual's belief in their ability to complete tasks at identified levels of success. That is, individuals can be successful despite meeting challenges [32]. Bandura identified four encounters that could develop self-efficacy: mastery experiences, vicarious experiences, verbal or social persuasion, and emotional or physiological states [33]. Mastery experiences are those in which prior success is realized, allowing a person to deepen their beliefs in their ability to be successful, such as performing well on a mathematics assessment. Vicarious experiences support self-efficacy by observing others' successes, mainly role models. When a student observes a peer's success they may see themselves as also being capable. Verbal persuasion refers to instances when an individual is supported through positive interactions. Students can be reminded of their ability to be successful through praise and sharing previous accomplishments. Finally, creating a safe, nurturing learning environment can minimize anxiety and improve self-efficacy. When educators focus on social-emotional learning supports, self-efficacy can be positively influenced. The feedback and modeling provided in our courses allowed opportunities for PSTs to experience all four origins of self-efficacy to develop and/or improve their own self-efficacies. Nevertheless, a thread evident in each of these experiences is that it depends on the individual's belief in themselves [34], which closely connects to the concept of a growth mindset [8, 9].

Growth mindset, introduced by Dweck [9], proposes that understanding and skills are developed by providing the necessary effort to the tasks. Boaler [8] applied this concept to mathematics learning to combat the fixed mindset that if you are "bad" at mathematics, you cannot be successful. By embracing a growth mindset, whether in school or life in general, an individual is willing to put forth the effort to achieve success. It does not matter if that success is based on prior success, observations of others, or persuasion – possessing a growth mindset allows one to be in a positive emotional state [7].

As documented in this section, these theoretical frameworks can be powerful in supporting student learning independently. However, when merged, they create a strong lens through which to reflect on our pedagogical practices, specifically our NTGPs. We sought to model for our PSTs how learning can be deepened through reflective and constructive feedback practices designed to support mastery. In the next section, we will describe the context of this research and how we implemented the NTGPs in our mathematics methods courses, a process applicable across educational contexts.

3. Methodology

This self-study [5] was conducted to determine if the use of NTGPs improved PSTs' development of PCK [6], self-efficacy [7], and growth mindset [8, 9] in a mathematics methods course. We utilized a self-study framework to investigate our own instructional practices through a critical friendship [35]. This self-study, presented as a collective case study [36], reports on archival data from the Spring 2021 semester in which both authors began NTGPs in their mathematics methods courses.

3.1. Setting and participants

This self-study was developed and implemented at a large, rural university in the southeastern part of the United States. Both researchers are members of the College of Education, but work in separate departments. Casler-Failing is an Associate Professor of Mathematics Education in the Department of Middle Grades and Secondary Education. Smithey is an Assistant Professor of Mathematics Education in the Department of Elementary and Special Education. Casler-Failing's course, presented as Case 1, was designed for PSTs majoring in middle-grade education (grades 4–8), with mathematics as one of their chosen content areas (in our state, middle-grade students obtain certification in two content areas). Smithey's course, presented as Case 2, was designed for PSTs majoring in elementary education (grades Pk-5).

Four participants of this study were PSTs enrolled in Casler-Failing's middle-grade mathematics methods class. These participants self-identified as female (three White and one Black), three participants were first-semester seniors, and one participant was a second-semester junior. Additional participants were 14 PSTs enrolled in Smithey's elementary mathematics methods class. All PSTs self-identified as female (twelve White and two Black). All students were first-semester seniors. All participants were enrolled in the respective courses as a requirement in their program of study. We chose these classes because these were the only classes we taught in a face-to-face environment during this semester.

3.2. Curriculum: Casler-Failing

Casler-Failing implemented ungrading practices in a face-to-face (F2F) mathematics methods course with undergraduate PSTs in an initial teacher preparation program (TPP) for middle-grade education. In this course, PSTs were informed of the assessment practice of ungrading on the first day of class. Based on the readings by Feldman [19] and Blum [2], PSTs were advised that they would receive comment-only feedback [23] on all assignments and projects. They were expected to apply the feedback to future assignments and may be required to revise and resubmit assignments if the minimum expectations were unmet; no quantitative grade would be provided. This information was also listed in the course syllabus as:

> Research shows that students tend to focus more on a given grade than utilizing feedback as a means to take their learning and understanding further. This semester I will be providing thorough feedback on each assignment and it is expected that the feedback will be incorporated into future assignments. At the end of the semester I will ask each student to reflect on their learning and then meet with me to discuss an appropriate course grade based on our conversation regarding effort and learning throughout the semester.

PSTs were required to complete a mid-semester (Google form) and an end-of-semester self-assessment (Google doc) to reflect on their learning. PSTs were required to attend a F2F meeting during finals weeks with Casler-Failing after completing the end-of-semester self-reflection to discuss their performance throughout the course. The conversations were student-driven based on the information provided in the self-assessment; during this final conference, a final grade was discussed and agreed upon by the instructor and PST.

Aside from enrollment in this course, PSTs engaged in content courses and a field placement during the semester. Within the field placement, PSTs developed a 5-day learning segment in their field placement course. Examples of activities PSTs completed in this mathematics methods course were weekly reading reflections, components of a learning segment (e.g., rationale and concept map), and a robotics lesson (based on learning about LEGO robotics throughout the course). The developed learning segment was supported by various activities in this methods course, such as creating concept maps, formative assessments, and an "I have, who has?" card game. The culminating assignment was an unessay [36] in which PSTs were required to research a topic of their choice, chosen from the topics covered throughout the semester that they wanted to continue learning. They also had the opportunity to choose how they presented the information (e.g., music video, PowerPoint, research paper).

3.3. Curriculum: Smithey

Smithey implemented portfolio assessment practices in a F2F methods course with undergraduate PSTs in an initial TPP for elementary education. This course was designed for a 15-week semester but was taught in 8 weeks (twice a week, 3-hour classes) to allow students to engage full-time in their assigned field placement during the second half of the semester after the course had concluded. In this course, Smithey informed PSTs of the portfolio grading structure [3] on the first day of class. Three portfolios documented PSTs' learning over the semester: (1) Professional Dispositions Portfolio, (2) Learning Portfolio, and (3) Knowledge for Teaching Portfolio. The course syllabus described this information to PSTs as:

> Learning is a complex endeavor that is not bound by time. Therefore, you are expected to make mistakes along the way. In order to create a safe space for these mistakes, portfolios will be utilized to collect evidence of growth over the semester. The most extensive portfolio will include formative assessments in which you have the option to revise and resubmit completed work to reach each assignment's learning goals/criteria based on written feedback. Each task or assignment will be accompanied by a rubric that lists criteria to meet the assignment's learning goals.

In the first portfolio, a professional disposition rubric was used mid and end-of-semester as a tool for self-reflection (by PSTs) and informal feedback (from the instructor) based on observations. For example, one criterion included effective communication and collaboration with peers. The second portfolio, the learning portfolio, comprised 50% of all coursework (formative assessments), and the third portfolio, the knowledge for teaching portfolio, comprised 30% of all coursework (summative assessments). Assignments included small tasks, such as short reflective journal entries, and larger tasks, such as course-long projects separated into manageable chunks. Regardless of the assignment size, all assignments had a rubric defining learning goals within these portfolios. The instructor shared written feedback for each assignment, devoid of traditional points. However, numbers indicated the level of mastery for each submission's criteria on the learning management system for communicating the progress of their portfolios – 0 (no submission), 1 (some learning goals met, major revisions needed), 2 (most learning goals met, minor revisions needed), and 3 (all learning goals met, complete). In the second portfolio only (learning portfolio), PSTs had the opportunity to revise/resubmit to progress toward mastery of learning goals within two weeks of receiving feedback, but it was not required. Collectively, a learning portfolio grade was assigned based on the number of assignments categorized as no submission (0), some learning goals met (1), most learning goals met (2), and all learning goals met (3). Revise and resubmit were only offered in the learning portfolio because it was the only portfolio that included formative assessments and was meant to be an assessment for learning [16, 17].

The course assignments included in the learning portfolio consisted of four types: (1) reflective journal entries based on the module focus (i.e., developing positive math identities, eliciting and using evidence of student thinking, cognitive demand, and lesson design), (2) choice boards to explore resources (i.e., technology, ongoing professional development), (3) examining student work samples (i.e., focus on mathematical content knowledge, student strategies, questioning student thinking), and (4) lesson design (i.e., critiquing the quality of math tasks, creation of math lesson plans).

3.4. Data and analysis

We used archival data to report on the research in this manuscript. After concluding the Spring 2021 semester, while engaging in conversation, we realized what we experienced throughout the semester as we implemented these assessment practices was valuable and needed to be shared with other educators interested in NTGPs. Therefore, we sought IRB approval from our institution to utilize the archival data from our courses after the semester concluded. The archival data collected consisted of all student assignment submissions, student reflections, and our written feedback paired with assignment rubrics. Additional data sources that were instrumental in the data analysis phase were the spreadsheets each author created to capture student mastery of learning before and after revisions. A limitation of using archival data, however, was the inability to secure PSTs' consent to use actual samples of student work or their exact words from reflections and self-assessments. Additionally, we could not contact the PSTs if questions arose during the analysis phase.

Instructor reflections regarding observations and personal experiences with NTGPs were completed by both authors. Casler-Failing completed reflections at the end of each class and after providing assignment feedback. For example, after reviewing the first submission of the content research project, Casler-Failing reflected on common omissions and misconceptions. The reflection informed instruction for the next class by allowing time to review the project's requirements and where to locate the information. The reflections were also instrumental in informing future course iterations and, thus, were included in data sources. Smithey completed reflections at the end of the course and made "notes of thoughts" throughout the semester. For instance, Smithey made a list of changes needed for future course iterations to improve the success of portfolios based on conversations with students and observations throughout the semester. The data we analyzed are shown below in Table 1. These data sources were chosen to capture a holistic view of NTGPs that included multiple perspectives (i.e., PSTs, instructors), types of data sources, and data from different points of time throughout the semester. Utilizing multiple data sources for each research question within each case supports the validity and reliability of the findings through triangulation [36, 37].

All data were analyzed qualitatively using an inductive thematic analysis [38]. This process involved individually reading through each data source to generate initial codes or important ideas that repeated or were relevant to NTGPs from multiple points of view (PSTs and instructor) such as extensive feedback, assignment descriptions, learning, or emotions. Then, we shared initial codes across all data sources and developed general categories (e.g., PCK, stress), converging to create overarching themes (e.g., expectations, feedback, time). To elaborate, once codes were identified, critical conversations [35] were conducted to develop general categories. This stage involved questioning one another's data and asking for justification with evidence from data sources to minimize biases. Each author then revisited the data through the lens of the mutually agreed-upon categories to develop the overarching themes in each data set. After independently developing overarching themes, the authors engaged in a critical conversation to discuss themes common to both data sets. Additionally, differences between the cases were discussed to determine if they were due to variances in teaching philosophies, student populations, or both. An important aspect of the data analysis process was the critical conversations both authors engaged in. Reflecting on our implementation of NTGPs across departments, programs, and courses added richness to the findings related to future teacher preparation. We felt this comparison was necessary since both authors instruct at the same institution – a large, rural university in the southeastern US. We acknowledge that our students come from a range of prior schooling experiences, including diverse grading systems, which were not controlled for in this study. Partnering as critical friends to reflect on our teaching practices allowed us the space to analyze our data sources through multiple lenses (PCK, growth mindset, and self-efficacy) to improve our enactment of equitable assessment practices. Our conversations reflecting on our experiences and our archival data from Spring 2021 illuminated the fact that NTGPs are not easily implemented into courses. In particular, our PSTs were unfamiliar with the process and had difficulty "letting go" of grades, and at times, we sometimes shared similar feelings. Since these NTGPs were something new that we implemented with little prior experience – we designed and implemented course content in response to our unique teaching philosophies – the critical conversations allowed us to discuss the purpose of the assignments in each course and the aligned feedback. Furthermore, the conversations allowed us the opportunity to question one another's findings as we performed our cross-case analyses [36, 37], thus minimizing bias.

Table 1
List of data sources aligned with research questions

Research question	Data source
What strategies promote the successful implementation of non-traditional grading practices?	PSTs' assignments/artifacts PSTs' reflections PST's mid-semester self-reflection PSTs' end-of-semester self-reflection Instructor reflections Notes/debrief from critical conversations
How do non-traditional grading practices support pre-service teachers' learning?	PSTs' mid-semester self-reflection PSTs' end-of-semester self-reflection Instructor data spreadsheets Instructor reflections Notes/debrief from critical conversations

4. Findings

This section will reveal our qualitative findings related to each research question. We will share the overarching themes identified from the data per case and then provide a cross-case analysis for each research question.

4.1. What strategies promote the successful implementation of non-traditional grading practices?

4.1.1. Case 1

In Case 1, the overarching themes focused on assignment expectations, rubrics, feedback, revisions, and time. The findings revealed that it is necessary to make the instructor's expectations clear and explicit. Examples were evidenced in the mid-semester self-reflections when PSTs reflected on their effort expended during learning (e.g., completing out-of-class readings, participating during in-class discourse) vs. when demonstrating learning (e.g., completing tasks and assignments). Furthermore, these expectations should be repeated often and presented in multiple modalities (e.g., written, verbal, visual). For instance, Casler-Failing shares her expectations in the course syllabus, verbally reviews them during the first class of the semester and intermittently throughout the semester, and includes applicable expectations when assigning tasks. Providing rubrics with all assessed assignments was necessary for supporting PSTs' ownership of learning, self-efficacy, and self-assessment, while also reminding PSTs of the expectations.

As with all courses, the content covered in a mathematics methods course moves quickly. Therefore, it was important to provide PSTs with timely feedback, but more importantly, the feedback had to be descriptive to support their continuous growth and improvement. For instance, simply stating "good job" is not adequate feedback; it is necessary to inform the PST of what they did well and what areas could be improved for future submissions. Each PST was required to create a learning segment;

one of the PSTs chose the topic of beginning statistics (i.e., scatter plots, lines of best fit) and, as part of the learning segment, was tasked with creating a student version of a concept map. After the submission, the feedback provided by the instructor was, "This is great. There is space for students to add pictures. You could support them by having some blank graphs on here that they add to when showing positive and negative associations and trend lines. Well done". This type of feedback not only provides praise but also moves the learner forward to apply their learning at a deeper level. Additionally, if the PSTs did not meet the expectations of the task, they needed to understand what was required in the revision process – and revisions were required when the minimum expectations of the task were not met. An example of this type of feedback was provided in response to a formative assessment submission as part of a learning segment. The following example comes from the topic of expressions and equations. The feedback stated, "Your objective has two verbs, therefore it is not specific. You should write this as 2 learning objectives. Your objective references expressions, but your examples are equations. Expressions do not contain an equal (or inequality) sign. Please revise and resubmit". Although this required revision could be deemed "extra work", it was viewed as a reflective process of learning and growth. PSTs reported through their reflections that they realized revisions resulted from their lack of understanding, lack of effort, or a combination of both. By implementing the revision requirement at the start of the semester, growth and learning developed throughout the semester to reduce or eliminate the need for revisions by the end of the semester (this idea is discussed further when sharing the findings for the next research question). Students reported the feedback allowed them to "grow" as learners. Some PSTs even reflected on experiencing productive struggle [39] throughout their learning as they described the revision process.

The final theme that emerged in the data was time. This theme was the most important because it encompassed all aspects of the non-traditional graded course. Additional time was demanded when creating clear expectations, repeating them often in multimodal formats, creating rubrics for all tasks, providing immediate and detailed feedback, requiring revisions and providing additional feedback, meeting with PSTs outside of class to discuss feedback, and finally, conferencing with PSTs at the end of the semester. The instructor always needed to be available for PSTs in the non-traditional graded environment. An additional time commitment not accounted for in this iteration of the course, but realized upon reflection after the semester's conclusion, was that PSTs needed the opportunity to participate in a mid-semester conference in addition to the end-of-semester conference. Completing the Google Form mid-semester self-reflection did not provide the instructor with enough information about how each PST felt about their demonstration of learning to date – an authentic F2F mid-semester conversation with the instructor would have been more beneficial for PSTs' self-reflection of learning and growth.

4.1.2. Case 2

In Case 2, four themes emerged when analyzing data sources and considering strategies that support the successful implementation of portfolio grading: explaining grading expectations, feedback, rubrics, and time. In the first theme, the findings revealed that PSTs need consistent and explicit conversations specific to the portfolio structure, emphasizing the revise/resubmit process beyond the first day of class, especially given that the process was optional. In this case, on the first day of the course, the instructor explained NTGPs, including the reasoning behind this type of grading, and described what the process would look like throughout the semester. After the first week of the course, the instructor considered it the responsibility of PSTs to review the course syllabus or ask questions if they needed clarification on grading processes. Thus, the instructor spoke little about the overall structure and the revision/resubmit process, with only a few reminders. By the end of the semester and upon instructor reflection, it was clear PSTs needed frequent reminders, encouragement, and additional discussions of grading practices after PSTs engaged with the feedback. For example, this course announcement was the only written reminder sent to students, and the announcement was sent four weeks into the course because the instructor noticed many students opted out of the revise/resubmit process:

> As I am posting the feedback today, I wanted to go ahead and send a reminder. We aren't doing traditional grades in this course because we are focused on the learning, reading, and application of feedback to assignments in your learning portfolio. A 0 indicates you are missing a submission for an assignment in your learning portfolio. A 1 or a 2 helps you see that you have feedback available and are encouraged to engage in either minor or major revisions to meet the learning goals of the assignment. In the feedback, you will see a note when the revision is due, so it is important to monitor your feedback frequently. Revisions are not required but strongly encouraged. When you submit a revision, make sure to resubmit to the submission area. I can then look at prior feedback as I re-examine the assignment based on the rubric. A 3 means your feedback has been posted and revisions are not needed because all learning goals have been met for the assignment and the assignment is complete. I hope that helps to clarify my expectations and I will see you this coming week!

Discussions throughout the semester centered around the learning goals for assignments, including rubrics, but less on the portfolio structure and the importance of the revise and resubmit process on PST learning. In particular, the instructor must be explicit about the importance of revision and its relationship to PSTs' learning to teach mathematics and mastery of the learning goals within the course. This was evidenced by the data that shows PSTs' engagement with the revision process. The learning portfolio consisted of 11 assignments (or smaller tasks). On average, across assignments, 7 out of 14 PSTs were encouraged to revise and resubmit because they demonstrated some or most learning goals met. Of the PSTs encouraged to revise and resubmit, on average, 29% chose to engage in the process to make progress toward mastery. In the following section, an illustrative example is shared from this case to highlight successful revisions leading to evidence of learning and to describe the other themes that emerged from the data – rubrics, feedback, and time.

Considering strategies that support the successful implementation of portfolio grading, findings revealed specific rubrics with clear expectations of learning goals within the assignment, detailed and timely feedback, and space for revisions on top of regular coursework are needed. For instance, one major project within the course included a "high-quality mathematics task project" broken down into three smaller tasks over the semester. In task 1, PSTs critiqued activities through the lens of cognitive demand [40] and justified their thinking. In task 2, PSTs modified an activity to increase the cognitive demand for Pk-5 students and summarized modifications connected to course learning. Finally, in task 3, PSTs wrote their activity as a complete lesson plan. Consistent across all assignments, feedback was provided using a rubric devoid of traditional points with criteria listed for "complete". No other indicators were provided, as those submissions not meeting the requirements for "complete" would receive detailed written comments on the rubric to guide their "minor" and "major" revisions (see Figure 1).

For example, in one submission, the rubric indicated "complete" for the first two criteria and minor revisions for the second two. In the submission, the PST justified an activity of

Figure 1
Criteria for parts 1 and 2 of the high-quality math task project

Criteria
Identify High/low quality tasks: At least 5 out of the 6 tasks are classified correctly
Identify High/low quality tasks: Ranking of tasks in order from higher to lower quality, shows understanding of the more specific details (demand levels) within high-quality tasks and within low quality tasks
Justification: Thoughtful reflections and/or critiques (beyond surface level) for highest and lowest ranking task, shows accuracy in understanding of characteristics of higher/lower quality tasks
Justification: Clear understanding of what constitutes a high quality and low quality task based on what has been learned (and read) in class
Task Modification: Revised task includes clear changes to add or expand use of research-based theories and effective teaching practices from in-class learning and out of class readings
Task Modification: Summary includes justification for the changes made to the task, a) how the changes increase the cognitive demand (or the quality) of the task, b) Clear connection to in-class learning and out of class readings

high cognitive demand because children could self-monitor and explain their thinking without drawing specific connections to the activity or course ideas. Within five days, feedback for the justification criteria was returned, among other written comments:

> You successfully identified two characteristics of high cognitive demand tasks: children self-monitoring their progress and explaining their reasoning for their math work to a peer. Elaborate on what you have written based on your learning of levels of cognitive demand. If it would be helpful, refer back to the class [PowerPoint] from week 4 to draw on more specific language and use some of the readings to support students engaging in classroom discussions. Another resource to pull from may be the beginning of the semester when we learned about the National Council of Teachers of Mathematics' effective teaching practices and standards of mathematical practices.

Although the rubric encourages connection to course content and readings, in the written feedback, the instructor highlighted specific resources to which the PST could refer to guide their revision and began with positive feedback. The PST revised the justification and submitted the assignment within a few days of receiving feedback (although PSTs had a two-week window) that described how the activity allowed children to self-monitor, including a more robust description of higher-order thinking skills in which children engaged. Further, the PST wrote more in-depth about making sense of problems and persevering in solving them, critiquing the reasoning of other children, and the activity's emphasis on children describing their conceptual understanding (sharing the why behind the steps of the problem). In this revision, the PST described details within the activity and incorporated effective mathematics teaching practices into the justification. The criteria were marked complete to indicate to the PST that the learning goals for the assignment were met. Upon reflection, the instructor noted that most rubrics within the course should be revised to include measurable and specific indicators at every rubric level to increase clarity of progress toward mastery and provide an additional layer of specific feedback. Further, the learning goals should be more explicit within the rubrics, as PSTs reported that the relationship between criteria and the learning goals was unclear. Although the learning goals were described as part of the assignment directions and the criteria aligned, having the success indicators on the rubric aligned with the criteria visually would support PSTs' understanding of that connection. Finally, in this illustrative example, the PST submitted the assignment, feedback was provided by the instructor, and the PST engaged in the revision process within a week. However, this time frame was atypical across PSTs as the instructor also reflected that often the revision process from start to finish was much longer. For instance, the instructor provided feedback within one week of the original submission, but PSTs had a two-week window to submit a revision. At the end of the semester, PSTs shared the challenge of the time-lapse associated with initially starting work on an assignment to finalizing a revision. Thus, timely feedback from the instructor and a smaller window for revision is needed. Regarding time, PSTs also reported a lack thereof to engage in meaningful revisions on top of regular coursework. They suggested minimizing repetitive assignments to allow space for more in-depth focus on one assignment. Four PSTs reflected that they had been interested in submitting a revision for various assignments, but had prioritized initial submissions across their courses. Throughout the semester, the instructor also observed the condensed nature of the course, along with the heavy workload of this course (and other courses) may have impeded PSTs' ability to engage in *optional* revision.

4.2. Cross-Case analysis for RQ #1

After engaging in a critical conversation to discuss our overarching themes, we quickly realized that several similarities existed. The strategies we found that promote the successful implementation of NTGPs are the necessity to provide clear expectations, the importance of timely and detailed feedback, the need for rubrics, and the time commitment (both for students and instructors). During our conversation, we discussed how we had provided the expectations to our PSTs. Since we both taught our PSTs the importance of multimodal instruction, we were not surprised to learn from one another that this area was also modeled for our PSTs. Our expectations were shared in written, verbal, and visual form, often through assignment exemplars.

The second similarity existed in our commitment to feedback. We both focused on providing detailed feedback and realized the importance of timely feedback, but we differed in our ability to provide timely feedback. Casler-Failing ensured feedback was provided to PSTs within 72 hours of each assignment submission, as stated in her course syllabus, and Smithey provided feedback within 10 days of each assignment submission. However, a statement was not provided in her course syllabus regarding feedback practices. Upon reflection, Smithey grappled the most between the number of students in the course and the length of feedback provided. Further, she realized that in a non-traditional grading environment, timeliness of feedback must become a priority, and has since improved her feedback practice to strike a balance between detailed feedback and timeliness. Although providing detailed feedback can be a major time commitment, we understood the importance of feedback in supporting learning [23]. Therefore, we were committed to supporting the growth and development of our PSTs. Furthermore, we viewed this feedback as a means to engage our PSTs in reflective practice and a way to support the accuracy of their revisions when required, so to us, this was a "win-win" – they were developing into future teachers.

Rubrics were another area that was common to both classes. We both deemed rubrics beneficial in supporting our expectations for

assignments, along with the verbal and visual guidance we provided. The rubrics aided PSTs in understanding our expectations and supported the feedback process, guiding PSTs in the revision process. Casler-Failing provides an example from a Content Research Project submitted as part of a learning segment; the criterion was described as "Identifies specific target learners and instructional strategies that support teaching the content to all learners". The feedback to the PST for this criterion as part of the revise and resubmit process was, "More information is needed about the target learners and how the instructional strategies will support all of their needs". Further, we discussed the importance of using specific language within rubrics to guide: feedback for the PSTs and support PSTs in making connections between the learning goals of the assignments, the criteria of rubrics, and our written feedback.

Finally, we discussed the time commitment in-depth, a major theme across both cases. Time is a significant component, from developing a non-traditional graded course to its implementation. Assessing assignments and reviewing subsequent revisions is not a quick process. For instance, both authors shared spending a fair amount of time supporting PSTs through multiple revisions with the goal of PSTs understanding the content. Through conversations, we also noticed a considerable amount of the semester was spent meeting with students, whether it was an end-of-semester conference or debriefing feedback together in more detail. To successfully implement the course while ensuring each PSTs' success, instructors must be willing to commit to the endeavor.

During one of our critical conversations, the main difference recognized across cases was that Casler-Failing required revisions when minimum expectations were not met, and Smithey made revisions optional. During the conversations, we concluded that PSTs' level of ownership of their learning and self-efficacy [7] increased at a higher rate when revisions were required. We posit that it may be too easy to say, "I will do it later" or "I don't have time right now" when the choice is optional. However, when PSTs must revise to meet expectations, they receive feedback allowing them to view their growth. PSTs witness how effort can develop their understanding, which supports their self-efficacy through mastery experiences and a growth mindset [8, 9]. In Case 1, revisions were required when minimum expectations were not met, and as the semester progressed, the depth of the revisions required for each assignment was reduced. PSTs still revised assignments in the second half of the semester, but the revisions were minor (e.g., APA formatting, rewording objectives, minor mathematical errors). Smithey noticed PSTs who consistently engaged in the revision process also needed to revise less often as the semester progressed.

4.3. How do non-traditional grading practices support pre-service teachers' learning?

4.3.1. Case 1

In Case 1, the overarching themes included PST reflection, ownership of learning, teacher feedback supported PST development of PCK, and reduced stress. The mid-semester and end-of-semester self-reflections allowed the PSTs to reflect on their learning. Casler-Failing conjectures that this opportunity to overtly reflect on their learning may have supported PSTs' ownership of learning as the data showed that one PST increased their effort in completing the course readings while two PSTs remained consistent with completing 75–100% of the readings. During the final F2F conference, PSTs reported benefitting from the numerous opportunities to reflect on their learning, whether it be through the questions the instructor posed in feedback (e.g., Does discourse always have to be about the debate? or Which mathematical practice standards would be more applicable for hands-on learning as opposed to an interactive lecture?), the required reading reflections, or the self-reflections. The feedback-only component of the course required PSTs to revise and resubmit assignments that did not meet the minimum requirements of the assignments. This feedback and revision requirement meant that PSTs were required to take ownership of their learning; they could not just complete an assignment to "turn it in" – the assignment had to be completed at the "acceptable level". PSTs' development of PCK was evidenced throughout the semester by the quality of the revise and resubmits required as the semester progressed. Although revisions were requested of all PSTs for nearly all assignments, by the end of the semester, required revisions were primarily minor for 3 out of 4 PSTs (e.g., improving grammar or aligning an objective with an assessment). No PSTs were required to complete a revise and resubmit for the final robotics lesson plan or unessay assignment [41], providing evidence of PCK [6]. For the final theme, all PSTs reported that not worrying about grades throughout the semester alleviated their anxiety. PSTs felt they could complete coursework without added pressure, knowing that if they did not "get it right" they would have opportunities to discuss and improve what they did incorrectly. One PST expressed the belief that removing grades would simplify the work, only to discover that it did not make the work easier but rather less stressful. Another PST reported that she could provide her best effort in an ungrading environment without feeling the pressure of what grade she would receive. However, one of the PSTs, who was excited about not having grades at the start of the semester, realized they were a motivating factor for her, and the feedback-only format, although beneficial to her growth as a future teacher, did not support her motivation.

4.3.2. Case 2

In Case 2, three themes emerged in analyzing data sources to consider how NTGPs support PSTs' learning: perception of feedback, the relationship between emotion and grades, and the development of PCK. In the first theme, PSTs shared how their perception of the role of feedback altered over the semester. For instance, during the discussion to introduce NTGPs, PSTs described typical instructor feedback as corrective, indicating what was "wrong" in their assignment or validating that they did a "good job". In class at the end of the semester, PSTs described using feedback as a reflective tool that supported their learning concepts more in-depth and attributed this idea to the specificity of feedback and the opportunity to apply the feedback. PSTs also described portfolios as a way to examine a collection of work for connections between assignments, visualize their growth over time, and see how assignments strategically build upon one another. In reflection, the instructor attributed some of these connections to the written feedback that guided its application from one assignment to the next. Additionally, opportunities for PSTs to reflect on their portfolio were provided periodically during the semester. To utilize the example from the "High-Quality Math Task Project" shared previously, the instructor drew PSTs' attention to the assignment of analyzing sample tasks through the lens of cognitive demand (part 1) and describing ideas for modification (part 2) to lesson design (part 3). For example, one PST's feedback on their revised version of part 1 and part 2 read:

> From your updated justifications, I have a much better sense of what you understand about high- and low-cognitive demand tasks with the specific details you added. In the future, when analyzing tasks you find through online resources, consider the characteristics of cognitive demand and National Council of Teachers of Mathematics' effective teaching practices. This process will be particularly important when you

complete Part 3, where you will simulate locating and modifying your activities to align with best practices.

In the written feedback, connections and future feedback applications were made explicit.

In the second theme, PSTs reported experiencing less worry about their grade throughout the semester because the emphasis was placed on their understanding with options to revise assignments as needed. End-of-course self-reflections revealed comments specific to the instructor's care about their learning and understanding of course content. However, many PSTs described increased anxiety toward the end of the semester because they had not fully understood exactly how their final grade would be entered into their transcript at the university level. Over half of the class began to understand the purpose of the revise and resubmit process, a way to demonstrate a more robust understanding toward the end of the semester. Unfortunately, this was after revision windows closed. Those PSTs who did not engage in the revise and resubmit process immediately began to express concern about their assigned grade at the end of the course.

In the third theme, NTGPs supported the development of PSTs' PCK. It is important to highlight that 85% of resubmitted assignments did reach mastery of all learning goals for the assignment. Further, 100% of the assignments resubmitted were improved by one or more rubric levels compared to the original submission. To offer an example from the "High-Quality Math Task Project", one assignment submission described a task modification for a place value activity that asked children to practice several problems rounding a number, given the underlined digit. The PST described how they modified the task to include students explaining their reasoning for each problem. After specific feedback to elaborate on the proposed modification to include increased attention to The National Council of Teachers of Mathematics' effective math teaching practices [39], the PST provided a more detailed description of their modification: children creating problems (including real-world situations), exchanging created problems with peers, and children engaging in discussions surrounding similarities and differences in their responses and explanations. The revise and resubmit process allowed the PST to revisit course material to extend their understanding of what it might look like in practice and provide space for children's ideas. PSTs also reported significant improvement in their understanding when they applied feedback to their work, instead of reading written feedback and moving on to the next assignment.

4.4. Cross-Case analysis for RQ #2

We engaged in another critical conversation to discuss how our NTGPs supported the learning of our PSTs. When sharing our findings, it was clear that a common similarity across both cases was the PSTs' responsiveness to the feedback we provided, an important aspect in developing each PSTs' PCK. Our PSTs were willing to ask questions when they did not understand the feedback or sit down to unpack the written feedback alongside us, to increase clarity in what they understood and where they needed support. Asking questions, unpacking feedback, and having chances to apply feedback to initial submissions developed their PCK. Furthermore, PSTs knew we would expect the feedback to be applied to future assignments, which added to the importance of being responsive and asking for clarification when unsure how to apply the feedback. Through our critical conversation, we shared examples of improved understanding of assignments. Since we saw a drastic change in their PCK, neither of us could imagine returning to grading practices that did not allow for a revision and resubmit process. Another similarity we discovered that surfaced across both courses was that the stress levels of our PSTs were minimized, or even eliminated, with our NTGPs throughout the semester. PSTs reported they did not need to stress over grades in our classes; they only needed to focus on putting forth their best effort. PSTs understood that if they made a mistake, it was just part of the learning process. However, PSTs in Smithey's course experienced a wave of anxiety at the very end of the course as they began to consider how NTGPs fit into the norms of the university, and one PST in Casler-Failing's course realized that grades were motivational. Ultimately, across cases, we found a variety of emotions connected to traditional and NTGPs. A major difference between the two cases was the level of ownership PSTs took over their learning, which we posit was due to communicated expectations about revisions. In Case 1, a revise and resubmit was required when the minimum expectations of the assignment were not met, whereas in Case 2, a revise and resubmit was optional. This dichotomy connects to our theoretical frameworks of self-efficacy [7] and growth mindset [8, 9]. When students are required to revise and resubmit they are supported through a mastery experience. This mastery experience may provide vicarious experiences for their peers. Both experiences can support increased self-efficacy, which we propose also supported growth in PSTs' mindsets; they realized that putting forth effort produces results in their learning, which, in this course, was improved PCK.

5. Conclusion

Our self-study of NTGPs identified five important strategies that promote successful implementation to support the success of PSTs. We argue these strategies are not limited to our field and can be adapted for educators implementing NTGPs and assessing student learning across disciplines and Pk-20 settings. Successful implementation requires a significant time commitment, clear expectations, detailed and timely feedback to engage students in reflective practice, detailed rubrics, and required revisions when minimum expectations are unmet. An instructor undertaking any form of ungrading must understand that this practice is extremely time-consuming. One must be committed to this endeavor to provide detailed feedback on all assignments in a timely manner. When revisions are required [27], you provide additional feedback while also assessing new assignments as they are submitted, not to mention, if a 3rd revision is required.

As we continued our critical conversations, and reviewed the analyzed data, we realized that to be successful with NTGPs, the course and instructor expectations need to be explicit and consistent from day 1. PSTs, or students in any course, need to be supported in this process; it may be just as new to them as it is to the instructor [15]. The expectations should be provided in many forms – verbal, written, and visual. Feedback may be provided in multiple formats as well, but it is important that it is detailed and timely, but not so specific that critical thinking and problem-solving are removed from the task. Remember, students are the learners and the task should be reflective. Pose questions to invoke thinking about the content. Provide detailed rubrics to support self-assessment and require students to self-assess before the first submission of the assignment. Although we believe in the "less is more" ideology of rubric creation, we learned that PSTs need enough information to gauge if they are at the minor revision, major revision, or complete stage. Finally, revisions must

be a requirement; they cannot be optional, which means contemplating course workload and time spent on revisions should be considered. Smithey's reflections revealed the need for PSTs to see the importance of and power in revising their work. The instructor considered that it may be helpful to showcase PSTs' work during class to highlight the important differences in understanding from one submission to the next and to reinforce key ideas of the course. PSTs may not be exposed to mastery or vicarious experiences [33] if they are not required to revise and resubmit. However, a necessary component in this experience, we suggest, is a growth mindset [8, 9, 42] among instructors and PSTs. Furthermore, by conferencing with students to review feedback and support their revisions, their self-efficacy is further supported by verbal persuasion and through promoting an emotional state for learning to occur [15, 33]. Additionally, three significant outcomes of our experiences during Spring 2021 were that PSTs were receptive to the feedback they received, PSTs' stress levels were reduced without grades, and PSTs' PCK improved throughout the semester in both classes. Although self-assessment is an important formative assessment practice, we found that PSTs were not well-versed in this area, so this became an opportunity to model best practices while enabling PSTs to reflect on their performance and advocate for their learning needs. Both forms of NTGPs used by the authors are a form of "assessment [that] must be a conversation, a narrative that enhances ... understanding of what [PSTs] know, what they can do, and what needs further work" [3]. Additionally, analyses of PSTs' assignments – specifically our feedback provided to the PSTs – provided evidence that their PCK [6] improved throughout the semester, particularly when participating in the revise and resubmit phase.

Increases in self-efficacy [7] and growth mindset [8, 9] were also apparent in the PSTs' artifacts, aligning with the *Standards for the Preparation of Mathematics Teachers* [10]. The artifacts reflected the PSTs' "...pedagogical knowledge, effective and equitable mathematics teaching practices, and positive and productive dispositions toward teaching mathematics..." Additionally, utilizing these NTGPs allowed our PSTs to experience methods of assessment that were focused on constructive feedback while allowing us, the instructors, to implement "[e]ffective assessments of ... candidates' development of mathematical knowledge relevant to teaching..." Furthermore, the NTGPs allowed us to evaluate PSTs' "dispositions related to mathematics teaching... and [their identity] as a mathematics teacher and learner". These NTGPs were "student-centered, unbiased, and fairly implemented" and used to "foster purposeful learning and meaningful relationships" with each instructor [1]. The feedback we provided to PSTs encouraged a deeper understanding of course content and connection-making in an environment built to promote the success of every PST. We propose the understanding gained through the self-assessment, feedback, and revision processes are skills the PSTs will carry forward into future learning and, hopefully, into their future classrooms.

We would be remiss if we did not discuss the grade aspect of our courses, after all, we were required to report letter grades for our PSTs at the end of the semester. In Case 1, the final topic of conversation in each final conference was the "grade discussion". More specifically, the self-reflection required PSTs to give themselves a letter grade and justify the grade. As the instructor, Casler-Failing agreed or disagreed with a justification; a disagreement would require evidence. During this first semester of non-traditional grading, each PSTs' self-imposed final grade aligned with Casler-Failing's final assessment of their performance. In Case 2, the final grade was determined through a weighted average of all three portfolios – 20% Dispositions Portfolio, 50% Learning Portfolio (the non-traditional grading aspect of the course), and 30% Knowledge for Teaching Portfolio. The Learning Portfolio grade was determined by reviewing the final documents submitted within the portfolio. This grade was determined by the documents' "completeness". Unfortunately, due to the low numbers of revisions submitted, this produced lower final grades for several PSTs than anticipated.

As with all research, there are limitations. First, this research reports only two courses from a large, rural southeastern US university. This manuscript shares our first iteration of implementing the NTGPs, and the demanding time commitment was unexpected, particularly alongside our numerous other obligations as tenure-track faculty (in Spring 2021, neither author was tenured). Utilizing archival data, especially PST self-reported data, was a limiting factor of the data collection and data analysis. It limited the richness of the reporting phase as we could not provide participant quotes since we could not obtain participant consent, nor could we follow up with participants for further clarification during analysis. Finally, this research, although rich in data sources, includes a small number of participants in each case, especially Case 1. However, the findings show promise for NTGPs to support the enhancement of self-efficacy [7] and a growth mindset [8, 9] while developing PCK [6]. Nonetheless, future studies investigating equitable grading practices with more diverse student populations and course content would enhance the applicability of these findings.

Recommendations

The implications this research presents for all educators (Pk-20) are important for several reasons. First, assessment is an important topic in all realms of education, and NTGPs should be included in the conversation to support student success, no matter the domain. The lessons we learned throughout this ungrading exploration, reported in this manuscript, are important to share – but none as important, we argue, as that of providing detailed and timely feedback [15, 30]. This idea is paramount to supporting students (in our cases, PSTs). Students need to know how to improve (or further develop their knowledge) and then have opportunities to do so [42]. However, those opportunities to revise their work must be provided sooner, not later. We understand this endeavor depends on time and acknowledge this could be a barrier to implementation. One strategy to minimize instructor workload, particularly for a large number of students such as Smithey's experience, is to keep a running document of descriptive feedback that can be applied to multiple submissions based on student work or utilize comment coding [43]. Another strategy, realized through our critical conversations, is to require PSTs to highlight their revisions upon resubmission to minimize the time required for review. Yet another avenue for reducing the time barrier would be to incorporate more peer assessment and self-assessment. Requiring students to self-assess their work [13], using the same rubric used by the instructor, may help them realize weak or missing content areas before submission [17]. If the assignment does not have a rubric, including a checklist [16] for students to review before submission could aid in meeting expectations. Furthermore, requiring students to provide a brief justification of their self-assessment could result in an opportunity to realize misconceptions about course content. If a revision is required, students could be required to share their initial feedback and revision with a peer to ensure the feedback has been addressed. These practices can reduce the instructor's overall workload while supporting feedback practices in large classes.

Another very important piece of advice, especially for someone new to this type of assessment who is implementing NTGPs for the first time, is to start small, with a project or paper. Do not try to overhaul your entire class at once. We overhauled our courses but Casler-Failing practiced ungrading in her K-12 classroom teaching so she could draw from those prior experiences. Smithey had no prior experience and, although she had some success, faced several challenges. NTGPs are a *work in progress,* and we continue to improve our practices every semester. If you are new to implementing these ideas, note the successes and challenges and move toward the next small change. Even when you think you have figured out the obstacles, new ones arise – be prepared to not only stumble but also grow in reflection.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

For information about the data, please contact the first author.

Author Contribution Statement

Shelli L. Casler-Failing: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review and editing, Visualization. **Montana Smithey:** Conceptualization, Methodology, Validation, Formal Analysis, Investigation, Data curation, Writing – original draft, Writing – review and editing, Visualization.

References

[1] Bishop, P. A., & Harrison, L. M. (2021). *The successful middle school: This we believe*. USA: Association for Middle Level Education.

[2] Blum, S. D. (2020). 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. 53–73). West Virginia University Press.

[3] Sackstein, S. (2015). *Hacking assessment: 10 ways to go gradesless in a traditional grades school*. USA: Times 10 Publications.

[4] Pai, G., Corby, J., Kras, N., Podlucká, D., & Yamamura, M. (2023). The dialectic transformation of teaching and learning in community colleges through ungrading. *Zeal: A Journal for the Liberal Arts*, *1*(2), 117–126.

[5] Diacopoulos, M. M., Gregory, K. H., Branyon, A., & Butler, B. M. (2022). Learning and living self-study research: Guidelines to the self-study journey. *Studying Teacher Education*, *18*(2), 175–196. https://doi.org/10.1080/17425964.2021.1992859

[6] Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14. https://doi.org/10.3102/0013189X015002004

[7] Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

[8] Boaler, J. (2016). *Mathematical mindsets: Unleashing students' potential through creative math, inspiring messages and innovative teaching*. USA: Jossey-Bass.

[9] Dweck, C. S. (2006). *Mindset: The new psychology of success*. USA: Random House.

[10] AMTE. (2020). *Standards for preparing teachers of mathematics*. USA: Information Age Publishing and the Association of Mathematics Teacher Educators. https://amte.net/sites/default/files/SPTM.pdf

[11] Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., . . . , & Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research*, *86*(4), 803–848. https://doi.org/10.3102/0034654316672069

[12] Schneider, J., & Hutt, E. (2014). Making the grade: A history of the A–F marking scheme. *Journal of Curriculum Studies*, *46*(2), 201–224. https://doi.org/10.1080/00220272.2013.790480

[13] 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). West Virginia University Press.

[14] 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

[15] Stommel, J. (2023). Do we need the word "ungrading"? *Zeal: A Journal for the Liberal Arts*, *1*(2), 82–87.

[16] Wiliam, D. (2018). *Embedded formative assessment*. USA: Solution Tree Press.

[17] Wiliam, D., Fisher, D., & Frey, N. (2024). *Student assessment: Better evidence, better decisions, better learning*. USA: Corwin.

[18] Chiaravalli, A. (2020). Grades stifle student learning. Can we learn to teach without grades? In S. D. Blum (Ed.), *Ungrading: Why rating students undermines learning (and what to do instead)* (pp. 82–88). West Virginia University Press.

[19] Feldman, J. (2023). *Grading for equity: What it is, why it matters, and how it can transform schools and classrooms*. USA: Corwin.

[20] Richmond, G., Salazar, M. D. C., & Jones, N. (2019). Assessment and the future of teacher education. *Journal of Teacher Education*, *70*(2), 86–89. https://doi.org/10.1177/0022487118824331

[21] Georgia Department of Education. (2022). *Teacher burnout in Georgia: Voices from the classroom*. Retrieved from: https://cviog.uga.edu/_resources/documents/training/teacher-burnout-georgia.pdf

[22] Goe, L., Bell, C., & Little, O. (2008). Approaches to evaluating teacher effectiveness: A research synthesis. *National Comprehensive Center for Teacher Quality*. Retrieved from: https://files.eric.ed.gov/fulltext/ED521228.pdf

[23] Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology*, *79*(4), 474–482. https://doi.org/10.1037/0022-0663.79.4.474

[24] Bremser, P. (2021). Assessment, evaluation, and grading: A reflection on my radicalization. *Journal of Humanistic Mathematics*, *11*(2), 291–297. https://doi.org/10.5642/jhummath.202102.15

[25] Guberman, D. (2021). Student perceptions of an online ungraded course. *Teaching and Learning Inquiry*, *9*(1), 86–98. https://doi.org/10.20343/teachlearninqu.9.1.8

[26] Blackwelder, A. (2020). What going gradeless taught me about doing the "actual work." In S. D. Blum (Ed.), *Ungrading: Why rating students undermines learning (and what to do instead)* (pp. 42–52). West Virginia University Press.

[27] Gibbs, L. (2020). Let's talk about grading. In S. D. Blum (Ed.), *Ungrading: Why rating students undermines learning (and what to do instead)* (pp. 91–104). West Virginia University Press.
[28] Kohn, A. (2020). Foreward. In S. D. Blum (Ed.), *Ungrading: Why rating students undermines learning (and what to do instead)* (pp. xiii–xx). West Virginia University Press.
[29] Maehr, M. L., & Midgley, C. (1996). *Transforming school cultures*. USA: Westview.
[30] Winstone, N. E., & Boud, D. (2022). The need to disentangle assessment and feedback in higher education. *Studies in Higher Education*, *47*(3), 656–667. https://doi.org/10.1080/03075079.2020.1779687
[31] Burns, E., & Frangiosa, D. (2021). *Going gradeless, grades 6–12: Shifting the focus to student learning*. USA: Corwin.
[32] Akhtar, M. (2008). What is self-efficacy? Bandura's 4 sources of efficacy beliefs. *Positive Physchology UK*. Retrieved from: https://positivepsychology.org.uk/self-efficacy-definition-bandura-meaning/
[33] Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71–81). Academic Press.
[34] Schunk, D. H., & DiBenedetto, M. K. (2021). Self-efficacy and human motivation. In A. J. Elliot (Ed.), *Advances in motivation science* (Vol. 8, pp. 153–179). Elsevier. https://doi.org/10.1016/bs.adms.2020.10.001
[35] Costa, A. L., & Kallick, B. (1993). Through the lens of a critical friend. *Educational Leadership*, *51*(2), 49–51. https://educandojuntos.cl/wp-content/uploads/2017/12/through_the_lens_of_a_critical_friend.pdf
[36] Yin, R. K. (2018). *Case study research: Design and methods*. USA: SAGE Publications.
[37] Patton, M. Q. (2002). *Qualitative research and evaluation methods*. USA: SAGE Publications.
[38] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa
[39] NCTM. (2024). *Principles to actions: Ensuring mathematical success for all*. USA: The National Council of Teachers of Mathematics.
[40] Stein, M. K., Smith, M. S., Henningsen, M., & Silver, E. A. (2009). *Implementing standards-based mathematics instruction: A casebook for professional development*. USA: Teacher College Press.
[41] Denial, C. (2019). *The unessay*. Retrieved from: https://catherinedenial.org/blog/uncategorized/the-unessay/
[42] Sackstein, S. (2020). Shifting the grading mindset. In S. D. Blum (Ed.), *Ungrading: Why rating students undermines learning (and what to do instead)* (pp. 74–81). West Virginia University Press.
[43] Keeley, P., & Tobey, C. R. (2017). *Mathematics formative assessment, volume 2: 50 more practical strategies for linking assessment, instruction, and learning*. USA: Corwin and National Council of Teachers of Mathematics.

How to Cite: Casler-Failing, S. L., & Smithey, M. (2024). How Do Non-traditional Grading Practices Promote Student Learning: Perspectives from Two Mathematics Teacher Educators. *International Journal of Changes in Education*. https://doi.org/10.47852/bonviewIJCE42023868