

Article

# Reimagining assessment: Exploring teachers' motivations and practices in alternative grading

Steve Verhoff<sup>1</sup>, Karen H. Larwin<sup>2,\*</sup>

<sup>1</sup> Upper Valley Career Center, Piqua, OH 45356, USA

<sup>2</sup> Department of Educational Leadership, Youngstown State University, Youngstown, OH 44514, USA

\* **Corresponding author:** Karen H. Larwin, [khlarwin@ysu.edu](mailto:khlarwin@ysu.edu)

## CITATION

Verhoff S, Larwin KH. Reimagining assessment: Exploring teachers' motivations and practices in alternative grading. *Forum for Education Studies*. 2025; 3(2): 2332. <https://doi.org/10.59400/fes2332>

## ARTICLE INFO

Received: 19 December 2024

Accepted: 10 April 2025

Available online: 3 June 2025

## COPYRIGHT



Copyright © 2025 by author(s).

*Forum for Education Studies* is published by Academic Publishing Pte. Ltd. This work is licensed under the Creative Commons Attribution (CC BY) license.

<https://creativecommons.org/licenses/by/4.0/>

**Abstract:** This mixed-method study explored the motivations, sustainability, and implementation of alternative grading practices among high school teachers in Ohio. Findings revealed that teachers adopt student-centered feedback methods such as rubrics, conferences, and self-reflection to enhance learning and support skill development. Social interactions, professional development, and peer influence emerged as significant factors shaping grading practices. Teachers often face challenges, including traditional grading pressures and limited formal training. Key limitations included a small, region-specific sample and reliance on teacher-reported data. Future research should investigate teacher preparation programs, the impact of policy mandates, and strategies for fostering systemic change. By understanding these dynamics, educational institutions can better support the adoption of innovative grading approaches that align with meaningful student learning and long-term academic success.

**Keywords:** grading practices; grading equity; student-centered feedback; assessment

## 1. Introduction

Traditional grading practices have historically been a major component of the formal educational process. Grades are used to measure student learning, motivate students to learn, and provide feedback to learners. Research indicates that flaws exist in beliefs regarding the purpose of traditional grading practices. For example, grades may not always accurately measure learning and can have adverse effects on student motivation. Furthermore, grades are not a reliable form of feedback when used exclusively. Overall, the grading process requires informed professional judgment, concern for what best serves the interests of all students, and careful examination of the tasks students are asked to complete to demonstrate their learning.

### 1.1. History of grading in the United States

The first known use of a grading system to grade or differentiate student performance in the United States was in 1785 at Yale University when 58 students were assigned a performance level for a senior examination [1,2]. By 1850, several colleges and universities in the United States had adopted some form of numerical grading system, but grading and reporting student performance was still virtually unknown in American primary and secondary schools [1,3].

Before 1850, one-room schoolhouses used narrative reports to communicate student performance to parents [2]. As school enrollment increased in the 19th century, schools began grouping students by age and using formal evaluations [3]. The growing student population made narrative reports impractical, leading to the adoption of

grading systems similar to those in higher education [3].

As compulsory education became more common, employers, universities, and society increasingly relied on grades to categorize and choose individuals for jobs, higher education, and social standing [4]. By the 1940s, the A-F grading system became widespread in K-12 and higher education [2]. Today's traditional grading practices, characterized by A-F marks, 100-point scales, and a mix of academic and non-academic factors, have a complex history [4]. Research has explored various aspects of grading, including report card composition, the impact of grades on outcomes, alternative grading practices, and the relationship between report card grades and accountability assessments [4]. The roughly 240-year history of grading practices in the United States is full of trial and error.

## **1.2. Reliability of grading**

Educators, parents, students, and researchers have long questioned the reliability and meaning of grades [2–4]. In a groundbreaking study in the early 20th century, Starch and Elliott [5] highlighted significant inconsistencies in how teachers graded identical essays. This research underscored the influence of teacher bias on grading practices, raising concerns about the fairness and accuracy of student evaluations.

Nearly a century later, Brimi [6] replicated Starch and Elliott's study, finding similarly significant inconsistencies among teachers, even those with specific training in writing assessment. This research further emphasizes the unreliability of grading practices. An "A" in one class may not equate to an "A" in another, even within the same school [6]. Brimi's [6] findings challenge the traditional notion of grades as objective measures of student achievement and highlight the need for more nuanced and equitable assessment practices.

## **1.3. Teacher bias & grading equity**

Numerous studies have identified teacher bias as a factor influencing student grades. A meta-analysis completed by Malouff and Thorsteinsson [7] reviewed 20 studies related to grading and found that bias can occur when graders are aware of irrelevant information about the students such as race, gender, and socio-economic status (SES). A 2005 report on grade inflation in higher education also supported this research, indicating that students' gender, SES, and race impact the grading of student work [8].

Further analysis has shown that racial bias can be found in teachers' evaluations of student work. Quinn [9] showed that Black students were graded lower than White students for identical work. Similar findings were observed in Sweden, where students with "foreign" names received lower grades than those with perceived "native" names [10,11].

Additional studies focused on grading bias and equity concerns found discrepancies in grades based on students' gender. Angelo and Reis [12] found that teacher grading benefits female students. Coates and Draves [13] reported that homework policies can disadvantage male students since they often submit late or incomplete assignments which is ultimately penalized in their overall grade.

Economic background, amount of adult support outside of school, prior school

experiences, and countless other variables impact student learning and can be factors that impact a student's grade when traditional grading practices are applied. Tomlinson [14] highlighted how teacher bias can impact pedagogy based on SES. Doyle et al. [15] found that teachers often judge lower-SES students more negatively than higher-SES students.

#### **1.4. Additional concerns regarding traditional grading practices**

A review of the history and research on grading practices presents a bleak outlook on traditional grading. Additional research on best practices for learning sheds more light on the issues with traditional grading and its impact on student learning. Traditional grading practices often incorporate non-academic factors such as effort, behavior, and participation, which can obscure a student's true academic achievement [4,16–21]. While these factors may be relevant in real-world settings, they can distort the assessment of academic knowledge and skills [22,23]. Elementary teachers, in particular, tend to place greater emphasis on non-academic factors compared to secondary teachers [17,23].

The inclusion of both academic and non-academic factors in a single grade, often referred to as a “hodgepodge” approach, can misrepresent a student's true academic strengths and weaknesses [1,16–22]. Combining all of these academic and non-academic factors into a single grade makes it impossible to identify the student's particular strengths and weaknesses in relation to the content. As a consequence of these traditional practices, a student's grade loses meaning by not providing an accurate reflection of student learning.

#### **1.5. Alternative grading practices**

Grading systems should do more than provide a summative grade. Grading systems should provide accurate and timely feedback to students, focusing on growth and learning rather than simply assigning grades [24,25]. In studies where schools have removed the fear of getting poor grades through alternative grading practices, students are more willing to take risks in their classes and more learning occurs [26].

Frustrations with traditional grading practices have led to reform efforts by individual teachers, schools, and school districts. Alternative grading practices, such as standards-based grading (SBG), have emerged to address these limitations. SBG provides a more accurate picture of student achievement by aligning with specific standards and competencies and offering multiple avenues for students to demonstrate their learning [24,27]. Additionally, SBG can mitigate the impact of racial bias in grading by providing specific grading criteria [9,28].

While traditional grading practices have been impacted by racial bias and stereotypes, alternative grading practices, such as SBG, diminish the impact of teachers' racial biases in student evaluation. Stereotypes have less influence on teachers' evaluations when specific grading criteria are established in advance [9]. Alternative grading practices, such as SBG and mastery-based grading with specific criteria, present potentially effective approaches to promoting racial equity in schools [9,28].

Individual teachers, schools, and school districts have explored a variety of

alternative grading practices to address inequities and problems associated with traditional grading practices. Experimentation with minimum grade policies to limit the impact of catastrophic grades and the exclusion of non-academic measures in grades are some of the ways that educators have tried to address problems with traditional grading without completely reforming traditional grading systems [16,18]. Educators have also experimented with contract grading, a grading approach based on a collective agreement, among students and instructors. Contract grading provides clarity of action steps toward learning goals and expectations [29]. In general, these alternative methods are focused on mastery and growth.

### **1.6. Impact of COVID-19 pandemic on grading practices**

The COVID-19 school closures brought on unprecedented challenges for schools. In the midst of the most disastrous pandemic since 1918, schools were presented with unprecedented challenges, including learning loss, financial strain, and increased mental health concerns [30–34]. The transition to distance learning disrupted traditional grading practices, as many districts paused traditional grading systems to avoid penalizing students without adequate access to technology or parental support [32].

Prior to the pandemic educators were focused on accountability measures of student progress, such as graduation rates and standardized test scores; modifying traditional grading practices was not largely examined. The pause on high-stakes assessments during the pandemic allowed teachers to focus on learning and mastery rather than test preparation [35]. The recent experiences of educators, students, and parents or guardians during the pandemic are prompting questions about traditional grading practices and how student learning is measured. Many educators are now experimenting with alternative grading methods, such as SBG [36].

### **1.7. Teacher training & professional development**

Assigning fair and meaningful grades to students is a challenge to novice and veteran educators at every level. The grading process requires thoughtful and informed professional judgment. Unfortunately, many teachers lack adequate training and guidance in this area, often relying on their own experiences as students [22,37]. This can lead to inconsistent and subjective grading practices, with teachers frequently incorporating non-academic factors such as behavior, attendance, and homework completion into their assessments [16,38]. Teachers are also unaware of what Carifio and Carey deem “semester killers” [16] (p. 25), which are large summative assessments such as tests or projects that can make or break a student’s quarter or semester grade.

It is crucial to provide teachers with comprehensive training on effective grading practices. Professional development can help teachers develop a deeper understanding of assessment principles, learn to identify and avoid biases, and implement more equitable and meaningful grading systems. For example, successful grading reform efforts occurred at a Minnesota high school after extensive professional development focused on grading research, classroom management, assessment, and instruction [18].

All teachers need better training and understanding of grading and assessment

practices in order to reduce teacher subjectivity [21,22]. Teachers with five years of experience or less (commonly referred to as novice teachers) receive very little training prior to entering the world of education. Novice teachers, in particular, may struggle with ethical decision-making and objective grading practices [39,40]. Bergman's [40] results suggest that novice teachers accept grading practices of veteran teachers, perpetuating detrimental aspects of traditional grading.

### **1.8. Theoretical framework**

Through collective leadership, Grogan and Shakeshaft [41] found that leaders are able to harness collective knowledge within an organization. Collective leadership, characterized by shared decision-making and collaboration, has been shown to positively impact school improvement and student achievement [42–44]. By harnessing the collective knowledge of stakeholders, leaders can foster a more supportive and innovative school culture [41].

Social network theory emphasizes the importance of interpersonal relationships in shaping individual and organizational behavior [45–47]. In educational settings, strong social networks among teachers can facilitate the sharing of knowledge, support, and innovation [45,48–50]. However, these networks can also reinforce traditional practices, including grading, if they are characterized by strong ties to individuals who resist change.

In conjunction with social network theory, collective leaders can identify with like-minded groups and individuals within and outside the organization to pursue change [41]. Collective knowledge provides a solutions-based approach to problems and fosters the development of action plans for sustained change. The development of relationships with stakeholders within and outside the organization helps leaders generate political power and the capacity for sustained change [41].

Collective leadership creates a more open organization where ideas and courses of action are developed collaboratively and not by a singular person. Due to the complexity and longstanding practices associated with grading, collective leadership is necessary for reform and sustained change with grading practices. Collective leadership would allow for more teachers to share and learn about alternative grading practices and assist in establishing new grading norms and greater recognition of the flawed aspects of traditional grading.

## **2. Methods**

The current study explored the establishment of, and changes made to grading practices as well as the sustainment and spread of alternative grading practices. The following questions were pertinent to guiding the study:

- How do teachers establish their grading practices?
- What prompts teachers to change their grading practices?
- How have teachers sustained these grading practices within the confines of a traditional grading system or policy?
- What has caused alternative grading practices to spread?

Q-methodology is a convergent mixed-methods design and combines qualitative and quantitative data to understand people's subjective viewpoints on a topic [51,52].

Q-methodology has been used in research to access the lived experiences of participants [52–54]. Additionally, Q-methodology can uncover diverse perspectives on a phenomenon in a nonthreatening way because the Q-set is already generated for the participants [55].

### **2.1. Participants**

The target population for this case study was full-time public school teachers from various Ohio high schools that have experimented with or employed alternative grading practices. The potential participants were identified through contact with school administrators from various Ohio high schools. Based on this desired sample, purposive sampling was employed for this case study to ensure a diverse representation of educators. Q-methodology typically uses a small sample size since there is a limit to the distinct viewpoints on a particular topic [55–57]. Sample sizes significantly larger than the Q-set will not likely yield additional viewpoints than those obtained by a sample size that is more in line with the Q-set size [57].

### **2.2. Instrumentation**

Q-methodology begins with defining and building a concourse [52]. A concourse is a collection of statements that reflect potential participant perspectives on the research topic [52]. The concourse can be established using informal interviews of individuals connected to the research, a review of relevant literature, or both [52]. An effective Q-set consists of concourse statements that are not redundant statements and should not be biased toward a particular viewpoint [52]. After a pilot of statements, a final Q-set was established for this investigation. The final Q-set used for this investigation is provided in the Appendix. In addition to the Q-sort, participants provided basic demographic information and responded to two open-ended items. The first question asked how the participants' graded practices reflect student-centered feedback. The second question asked participants how they know a student has shown growth. No identifying information was collected.

### **2.3. Procedure**

After receiving approval from the Youngstown State University Institutional Review Board, email invitations were sent to a purposive sample of Ohio high school and career-technical school schoolteachers experienced in alternative grading practices. The emails detailed the study, estimated a participation time of 30 min, and included a link to the Q-sort activity.

Participants completed a brief demographic and experience survey before proceeding to the Q-sort, where they reflected on statements about grading practices. Using a web-based platform, participants sorted statements into categories from “strongly disagree” to “strongly agree”. A forced distribution framework guided this process, ensuring all statements were ranked. Multimedia prompts supported participants throughout the activity.

The QMethod Software automatically recorded responses, eliminating human error and enabling immediate data availability for analysis. Participant responses were statistically examined to identify dominant belief patterns about grading development.

To ensure confidentiality, pseudonyms were used for participants, schools, and districts. Informed consent was obtained, and participants were notified of their right to withdraw at any time. The study emphasized voluntary participation and privacy protection.

The survey also included open-ended questions about teachers' grading experiences, influences on their practices, and reasons for adopting alternative approaches. These responses provided valuable qualitative insights into grading practice development.

### **3. Results**

This study aims to create a relatable framework for understanding diverse teacher perspectives on grading. Cinematic mentors have been used as identifiers for each of the five factors revealed through factor analysis. These cultural references were used to assist in making complex pedagogical ideas more accessible through familiar cultural references. The five factors were identified as Mr. Miyagi, Herman Boone, John Keating, Ms. Valerie Frizzle, and Dr. Sean Maguire.

#### **3.1. Descriptive statistics**

The Q-sort was initiated by 55 participants, but only 24 participants completed the Q-sort, generating a completion rate of 44%. Nineteen participants identified as female (79%) and five as male (21%). Five participants indicated they had earned a bachelor's degree (21%), while 19 reported having earned a master's degree (79%). Seven participants reported having 16–20 years of teaching experience (29%); five participants reported having 6–10 years of experience (21%); four participants reported having 1–5 years of experience (16%); three participants reported having 11–15 years of experience (12.5%); and three participants reported having 25 or more years of experience (12.5%).

Content areas taught and school settings varied among the participants. Nineteen participants reported teaching at a traditional high school (79%) and five participants reported teaching at a career-technical school (21%). Six participants reported their teaching concentration to be mathematics (25%); six participants reported primarily teaching English language arts (25%); four participants reported teaching a career-technical program (17%); two participants reported teaching social studies (8%); and at least one participant reported teaching in each of the following content areas: fine arts, foreign language, physical education, science, and special education (4%). Seventeen participants reported their school being located in a suburban setting (71%); five participants stated their school was in a rural setting (21%); and two participants reported their school being located in an urban setting (8%). Nineteen of the participants reported their school was located in southwest Ohio (79%); two participants reported their location to be in northwest Ohio (8%); another two participants reported being located in northeast Ohio (8%); and one participant reported their location to be in central Ohio. Twenty participants reported their school enrollment to be over 1000 students (83%) and four reported the student enrollment of their school to be between 500 and 1000 students (17%).

Most of the participants indicated their school to be highly rated according to the

Ohio Department of Education rating system and utilizing traditional school-wide grading policies. Seventeen participants reported their school was rated 4.5 stars or above (71%); one participant reported their school rating to be 3.5 stars (4%); another participant reported their school rating to be two stars (4%); and five participants reported they did not know their school rating (21%). Twenty-three participants indicated their school-wide grading policies utilized traditional grading practices (96%) and one participant reported their school utilized SBG practices (4%). The following figures represent some of the participant demographics.

### 3.2. Q-sort results

#### 3.2.1. Factor extraction

Factor extraction is a statistical technique used to analyze the data collected through Q-sorts [57]. The process identifies groups of participants who ranked Q-sort items in similar ways, grouping participants with shared perspectives into factors [57]. The 24 Q-sorts in this study were analyzed through extraction of five factors with a Varimax rotation of factors. Auto flagging was set to  $p < 0.05$ . **Table 1** represents the correlation between factor scores. Low correlation among the five factors suggests discriminating viewpoints from the participants' Q-sorts.

**Table 1.** Correlation between factor scores.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	*	0.36	0.16	0.07	0.15
Factor 2	-	*	0.14	0.31	-0.01
Factor 3	-	-	*	0.07	-0.10
Factor 4	-	-	-	*	0.06
Factor 5	-	-	-	-	*

Note: \* no significant correlations between factors.

Factor analysis was run three times to ensure the optimal number of factors among participants' Q-sorts. Initially, a six-factor model was used, resulting in a 47% variance captured with four participants not being identified by one of the factors. Next, a five-factor model was evaluated resulting in a 44% variance captured with five participants not being identified by one of the factors. Finally, a four-factor model was analyzed, resulting in a 39% variance captured with five participants not identified by one of the factors.

Additional analysis using eigenvalues indicated the best factor model. A six-factor model produced an eigenvalue for factor 6 of 0.67. A five-factor model produced eigenvalues ranging from 4.47 to 1.02, adhering to the acceptable cutoff of 1.0. The comparison between the models revealed the five-factor to have the most adherence to the law of parsimony and the best fit for this study. **Table 2** displays the determining statistics for the use of a five-factor model.

**Table 2.** Five-factor model analysis.

	<b>Factor 1</b>	<b>Factor 2</b>	<b>Factor 3</b>	<b>Factor 4</b>	<b>Factor 5</b>
No. of defining variables	7.00	3.00	4.00	2.00	3.00
Avg. rel. coef.	0.80	0.80	0.80	0.80	0.80
Composite reliability	0.97	0.92	0.94	0.89	0.92
S.E. of factor Z-scores	0.19	0.28	0.24	0.33	0.28
Eigenvalues	4.47	1.82	2.00	1.19	1.02
% Explained variance	18.6	26.2	8.3	5.0	4.2
Cumulative % explain var	18.6	26.2	34.5	39.5	43.7

### 3.2.2. Varimax rotation

This study utilized QMethod Software for data collection and analysis. QMethod Software allows researchers to utilize varimax rotation for factor rotation. Varimax rotation can be used to rotate the factors in order for the Q-sort to achieve maximum factor loading, leading to a single-factor fit [57]. **Table 3** displays the Q-sort associated with each participant following the varimax rotation and indicates the factor extraction in bold numbers in the appropriate factor column.

Nineteen of the 24 participants loaded significantly on one of the five factors. Combined, the factors account for approximately 44% of the study’s variance. Five participants did not load significantly on one of the five factors. This indicates that the five participants did not fit well with the five main factors that were extracted from the study or that these participants did not provide reliable feedback.

**Table 3.** Participant factor loadings.

<b>Participant</b>	<b>Factor 1</b>	<b>Factor 2</b>	<b>Factor 3</b>	<b>Factor 4</b>	<b>Factor 5</b>
0X0F	0.15	-0.01	-0.17	0.20	<b>0.41</b>
1MRD	0.15	0.04	<b>0.75</b>	0.28	-0.14
21D8	<b>0.68</b>	0.00	-0.10	0.08	-0.04
2JZ6	<b>0.68</b>	0.16	0.06	-0.17	0.09
325T	0.37	-0.18	<b>-0.55</b>	-0.15	-0.06
5TMU	<b>0.73</b>	0.23	0.06	0.21	0.17
76EA	<b>0.43</b>	-0.02	-0.05	0.05	0.25
A3IT	<b>0.57</b>	0.16	-0.04	0.02	0.11
BNTZ	0.05	0.13	0.05	-0.04	<b>0.43</b>
G4C9	0.44	0.46	-0.02	-0.27	0.22
G5Q5	0.51	0.52	0.04	-0.11	0.10
GVCT	0.37	<b>0.60</b>	0.08	0.28	-0.13
IIQ9	<b>0.36</b>	0.20	0.03	-0.29	-0.01
IZAH	0.25	0.06	0.38	0.48	0.13
LK7Y	0.15	0.68	0.02	0.06	0.01
MCTG	0.38	0.06	<b>0.69</b>	-0.07	-0.03
QXV0	<b>0.41</b>	0.13	0.14	0.17	-0.08
RJE7	-0.02	0.08	0.02	0.05	<b>-0.54</b>

**Table 3.** (Continued).

Participant	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
T94W	0.09	0.07	-0.03	<b>0.60</b>	-0.15
TL45	-0.12	0.25	0.07	<b>0.61</b>	0.16
UTQ7	0.01	0.50	-0.13	0.21	0.03
UUPF	0.39	0.06	-0.39	0.25	0.02
VV9Z	0.10	0.26	<b>-0.54</b>	0.22	0.07
Y4EE	0.41	0.30	-0.46	-0.16	-0.09

### 3.3. Factor arrays, identification, and interpretation

Q-methodology is used to gather participants' viewpoints in order to better understand how they think about a particular topic or issue, specifically in this study, the development of grading practices [57]. A factor array is a visual tool used to display the participants collective viewpoints [57]. A factor array does not indicate any specific individual participant's comments, statements, or views, but rather a group of individuals with similar viewpoints. The following section will provide arrays for each of the five factor models along with the factor's identification and interpretation of the participant's viewpoints. For this study, factor 1 will be referred to as Mr. Miyagi; factor 2 will be referred to as Herman Boone; factor 3 will be referred to as John Keating; factor 4 will be referred to as Ms. Valerie Frizzle; and factor 5 will be referred to as Dr. Sean Maguire. Participants that did not load significantly with any of the five factors will be referred to as Professor Minerva McGonagall.

Factor 1: Mr. Miyagi

"First learn stand, then learn fly."

Mr. Miyagi, "The Karate Kid" (1984).

In "The Karate Kid", Mr. Miyagi embodies a traditional, patient, and holistic approach to teaching. For this study, Mr. Miyagi represents a perspective that values mastery over rote memorization, emphasizing the process of learning and personal growth alongside the final outcome. These participants have a grading perspective that is less focused on rigid scores and more on the student's overall development.

**Figure 1** depicts the factor array for Mr. Miyagi, which had seven statistically loading participants, accounted for 18.6% of the variance, and an eigenvalue of 4.47. **Table 4** lists the distinguishing statements for Mr. Miyagi. The Mr. Miyagi participants ranged in number of years teaching from one year to over 25 years, with more than half of the participants having more than 10 years of experience (57%). All of the Mr. Miyagi participants indicated being female. One of the Mr. Miyagi participants was from a career-technical high school and the remaining were from traditional high schools. The Mr. Miyagi participants represented a wide range of academic content areas, including two English language arts teachers, one foreign language teacher, one math teacher, one career-technology teacher, one physical education teacher, and one special education teacher. Mr. Miyagi included the largest number of participants when compared to the other factors (29%).



**Table 4.** Distinguishing statements for factor 1, Mr. Miyagi.

Statement number	Statement	Endorsement
30	I think about different ways students can demonstrate achievement, and I try to use a number of different assessment methods.	+
31	I consider whether students are demonstrating understanding of the learning objectives being taught.	+
3	Grading can keep students informed about their progress.	+
11	When I first started teaching, my experience with grading was very procedural and a high percentage of my grade was completion-based.	-
8	Grades are an imprecise measure of learning.	-
22	My peers have influenced my grading practices.	-
42	An administrator influenced the development of my grading practices.	-

Notes: Endorsement indicated as positive (+), neutral (^), or negative (-).

### 3.4. The general viewpoint of Mr. Miyagi

Mr. Miyagi is focused on determining student mastery of content and skill acquisition. Mr. Miyagi designs learning activities and assessments that are directly tied to learning objectives and offer a variety of methods for students to demonstrate learning. Mr. Miyagi views student grades as a means to monitor student progress and determine if learning objectives have been met. These sentiments were evident in the Mr. Miyagi consensus statement indicating they are more willing to allow students to redo assignments in order to make sure a grade reflects student mastery. In the content areas taught by Mr. Miyagi, mastery of content and skill acquisition is essential to understanding and learning new content or skills.

Mr. Miyagi recognizes that traditional grades are imprecise measures of student learning. Additionally, Mr. Miyagi has relied upon their professional experience to develop and modify their grading practices. Mr. Miyagi reported little influence from peers or administrators in developing their current grading practices.

Factor 2: Herman Boone

“I don’t scratch my head unless it itches and I don’t dance unless I hear some music.”

Herman Boone, “Remember the Titans” (2000).

Herman Boone, from “Remember the Titans”, is a football coach who demands high standards and discipline. In the film, Herman Boone’s actions are intentional with specific outcomes or purposes in mind. In terms of grading, Herman Boone represents a participant perspective whose grading practices are intentionally designed to provide feedback and improve student outcomes.



**Table 5.** Distinguishing statements for factor 2, Herman Boone.

Statement number	Statement	Endorsement
36	I often reflect on my grading practices and how I can better use data from a variety of assessments to help improve my instruction.	+
41	I have to keep my students' needs foremost in my considerations when grading and designing assessments.	+
38	I would say compared to the rest of the teachers in my school, I have a unique grading practice.	+
45	Anything to make me think about grading practices was through informal professional conversations.	+
1	Grading practices are important measures of student achievement.	-

Notes: Endorsement indicated as positive (+), neutral (^), or negative (-).

**Figure 2** depicts the factor array for Herman Boone, which had three statistically loading participants, accounted for 26.2% of the variance, and had an eigenvalue of 1.82. **Table 5** lists the distinguishing statements for Herman Boone. The Herman Boone participants ranged in number of years teaching from six years to over 25 years of experience, with two of the participants having more than 16 years of experience (67%). Herman Boone included two male and one female participants. One of the Herman Boone participants was from a career-technical high school and the remaining were from traditional high schools. Herman Boone represented the mathematics, science, and social studies content areas.

### 3.5. The general viewpoint of Herman Boone

Herman Boone is reflective and purposeful about their grading practices. Herman Boone seeks out ways to improve grading practices in order to provide feedback about student learning and guide future instruction. As a result, Herman Boone believes their grading practices are unique compared to their colleagues. Herman Boone does not cite any formal training in regard to grading students but does have informal conversations about grading practices. Over the course of Herman Boone's career, their grading practices have evolved. Herman Boone views grading as an informative practice for instructional purposes, not as a means to measure student achievement.

Factor 3: John Keating

“Now we all have a great need for acceptance, but you must trust that your beliefs are unique, your own, even though others may think them odd or unpopular, even though the herd may go, (imitating a goat) that's baaaaad. Robert Frost said, ‘Two roads diverged in the wood and I, I took the one less traveled by, and that has made all the difference.’”

John Keating, “Dead Poets Society” (1989).

In “Dead Poets Society”, John Keating inspires students to think critically, challenge conventions, and embrace individuality. He represents a pedagogical approach that values creativity, passion, and the pursuit of knowledge for its own sake. For this study, John Keating represents a perspective that has chosen to explore a different path for grading by experimenting with alternative grading practices and using grading to promote student ownership of learning.

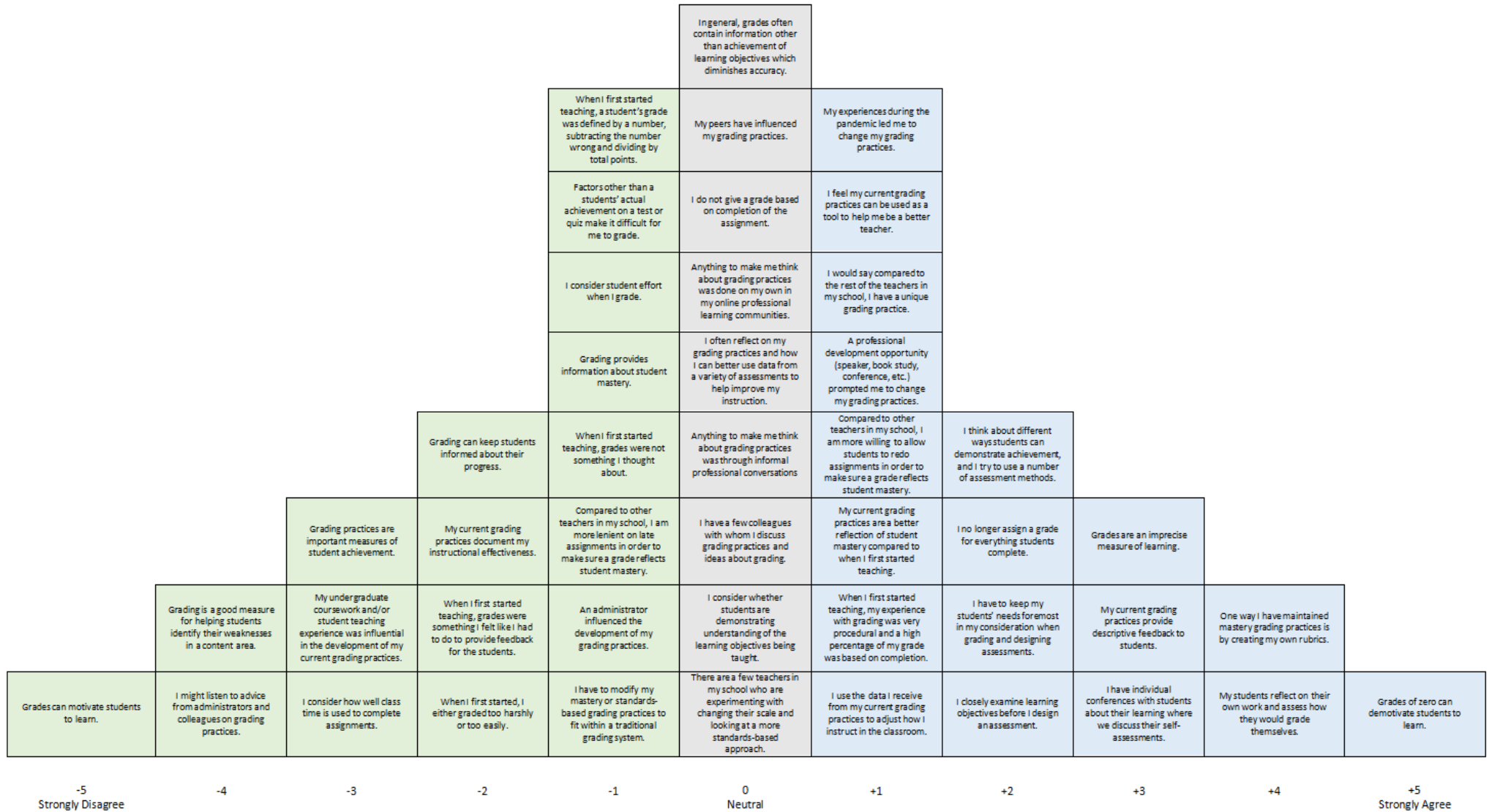


Figure 3. Model sort for teachers who loaded significantly on factor 3: John Keating.

**Table 6.** Distinguishing statements for factor 3, John Keating.

Statement number	Statement	Endorsement
7	Grades of zero can demotivate students to learn.	+
34	My students reflect on their own work and assess they would grade themselves.	+
33	One way I have maintained mastery grading practices is by creating my own rubrics.	+
35	I have individual conferences with students about their learning where we discuss their self-assessments.	+
19	My current grading practices provide descriptive feedback to students.	+
15	My current grading practices are a better reflection of student mastery compared to when I first started teaching.	+
38	I would say compared to the rest of the teachers in my school, I have a unique grading practice.	+
40	I might listen to advice from administrators and colleagues on grading practices.	-
2	Grading is a good method for helping students identify their weaknesses in a content area.	-
6	Grades can motivate students to learn.	-

Notes: Endorsement indicated as positive (+), neutral (^), or negative (-).

**Figure 3** depicts the factor array for John Keating, which had four statistically loading participants, accounted for 34.5% of the variance, and had an eigenvalue of 2.00. **Table 6** lists the distinguishing statements for John Keating. The John Keating participants ranged in number of years teaching from one year to 25 years, with three of the participants having more than 11 years of experience (75%). John Keating were all female participants. All but one John Keating reported having their master’s degree. One of the John Keating participants reported being from a career-technical high school and the remaining were from traditional high schools. John Keating included two math teachers, one career-technical teacher, and one social studies teacher.

### 3.6. The general viewpoint of John Keating

John Keating has been innovative in the development of grading practices. John Keating utilizes descriptive feedback, teacher-generated rubrics, and individual student conferences to provide students with information about their learning. John Keating does not conform to traditional grading practices, but rather utilizes alternative grading practices to fit within the context of a traditional grading system. John Keating does not view grades as a motivator for students. Instead, John Keating engages students in the learning process through self-reflection and self-assessment. Through this process of self-assessment, John Keating is able to promote student ownership of learning and help students identify individual strengths and areas for growth.

Factor 4: Ms. Valerie Frizzle

“Take chances, make mistakes, get messy.”

Ms. Valerie Frizzle, “The Magic School Bus” (1994).

Ms. Valerie Frizzle from “The Magic School Bus” embodies experiential learning, curiosity, and a sense of adventure in education. In this study, Ms. Valerie Frizzle translates to a perspective that has identified the problems with traditional grading, but is apprehensive about experimenting with alternative grading practices. These participants need to embrace their inner Ms. Frizzle and experiment, but something is holding them back from doing so.



**Table 7.** Distinguishing statements for factor 4, Ms. Valerie Frizzle.

Statement number	Statement	Endorsement
8	Grades are an imprecise measure of learning.	+
9	Factors other than a student’s actual achievement on a test or quiz make it difficult for me to grade.	+
5	I consider student effort when I grade.	+
10	In general, grades often contain information other than achievement of learning objectives which diminishes accuracy.	+
15	My current grading practices are a better reflection of student mastery compared to when I first started teaching.	^

Notes: Endorsement indicated as positive (+), neutral (^), or negative (-).

**Figure 4** depicts the factor array for Ms. Valerie Frizzle, which had two statistically loading participants, accounted for 39.5% of the variance, and an eigenvalue of 1.19. **Table 7** lists the distinguishing statements for Ms. Valerie Frizzle. Both Ms. Valerie Frizzle participants have been teaching between 11 and 20 years and indicated they were female. Both also reported being English language arts teachers and having their master’s degree. Ms. Valerie Frizzle teaches at a career-technical school and a traditional high school.

### 3.7. The general viewpoint of Ms. Valerie Frizzle

Ms. Valerie Frizzle recognizes traditional grading practices are imprecise measures of student learning but has not taken as big of a leap into alternative grading as Mr. Miyagi, Herman Boone, and John Keating. Ms. Valerie Frizzle has identified a variety of non-academic factors that influence a student’s grade that diminishes the ability of a traditional grade to measure achievement or master of learning objectives. Ms. Valerie Frizzle does not provide completion grades and attempts to make certain each learning activity that is to be graded demonstrates student mastery as opposed to any non-academic factors. Ms. Valerie Frizzle has identified the problem with traditional grading but has not entirely embraced alternative grading practices as a means to improve assessment of student learning.

Factor 5: Dr. Sean Maguire

“You’ll never have that kind of relationship in a world where you’re afraid to take the first step because all you see is every negative thing 10 miles down the road.”

Dr. Sean Maguire, “Good Will Hunting” (1997).

In the film “Good Will Hunting”, Dr. Sean Maguire is a teacher and mentor who prioritizes emotional intelligence, empathy, the impact of personal experiences, and the understanding of individual needs. However, in the film, Dr. Maguire is able to mentor others, but is unable to overcome his own personal tragedies and experiences keeping him from moving forward in his life. For this study, Dr. Maguire represents a perspective that is confident in their current grading practices and does not see the need to apply alternatives to their assessments of student learning. Like the character in the film, these participants are able to mentor others, but don’t focus on their professional growth.

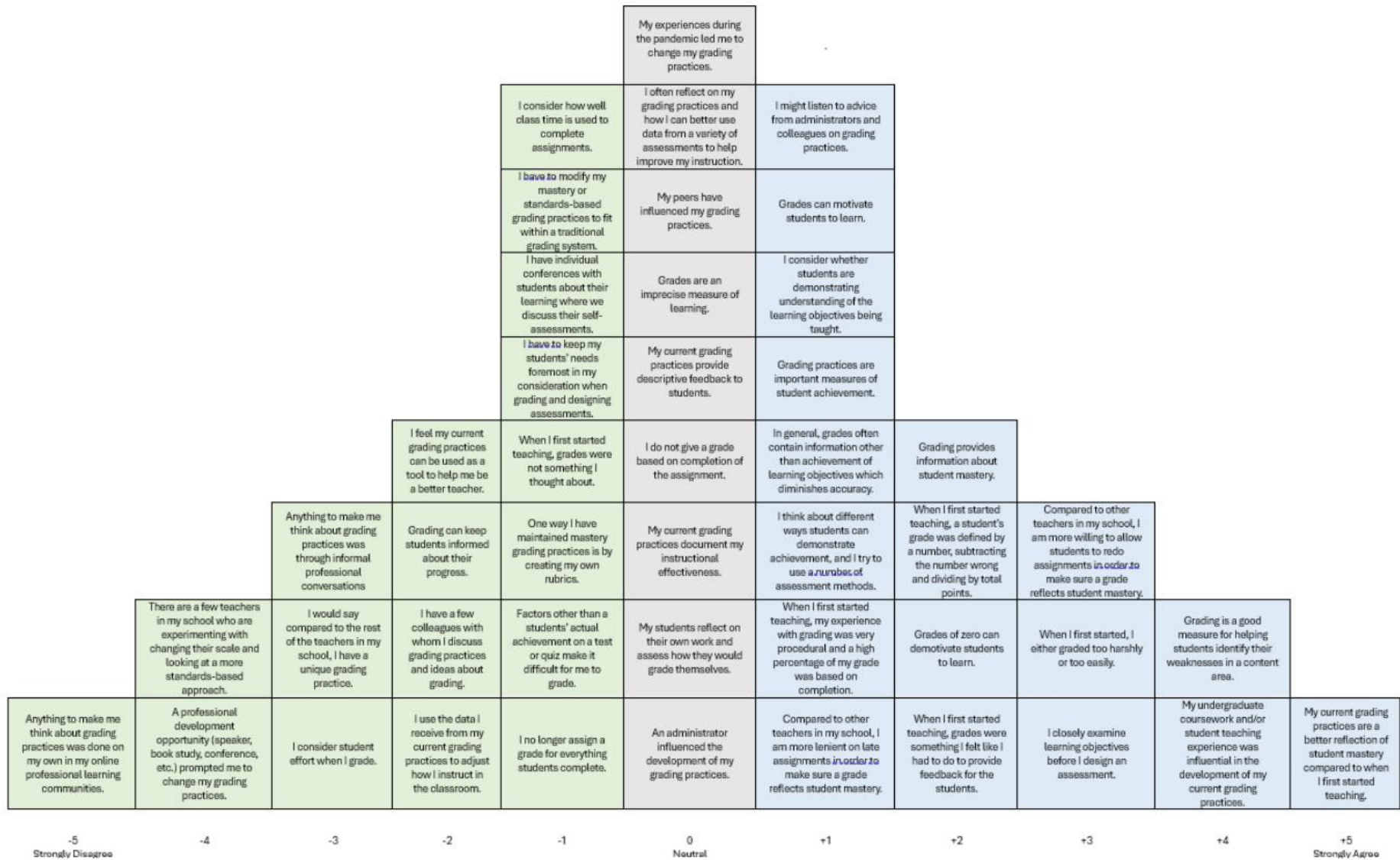


Figure 5. Model sort for teachers who loaded significantly on factor 5: Dr. Sean Maguire.

**Table 8.** Distinguishing statements for factor 5, Dr. Sean Maguire.

Statement number	Statement	Endorsement
43	My undergraduate coursework and/or student teaching experience was influential in the development of my current grading practices.	+
41	I have to keep my students' needs foremost in my considerations when grading and designing assessments.	-
2	Grading is a good method for helping students identify their weaknesses in a content area.	+
45	Anything to make me think about grading practices was through informal professional conversations.	-
18	I use the data I receive from my current grading practices to adjust how I instruct in the classroom.	-
5	I consider student effort when I grade.	-
23	A professional development opportunity (speaker, book study, conference, etc.) prompted me to change my grading practices.	-

Notes: Endorsement indicated as positive (+), neutral (^), or negative (-).

**Figure 5** depicts the factor array for Dr. Sean Maguire, which had three statistically loading participants, accounted for 43.7% of the variance, and had an eigenvalue of 1.02. **Table 8** lists the distinguishing statements for Dr. Sean Maguire. The Dr. Sean Maguire participants ranged in number of years teaching from one year to 25 years, with two of the participants having more than 16 years of experience (67%). Dr. Sean Maguire included one female and two male participants, all having earned a master's degree. Only one of the Dr. Sean Maguire participants was from a career-technical high school and the remainder were from traditional high schools. Dr. Sean Maguire included two career-technical teachers, and one English language arts teacher.

### 3.8. The general viewpoint of Dr. Sean Maguire

Dr. Sean Maguire is confident in their grading practices. The Q-sorts from Dr. Sean Maguire indicate very little interest in learning about alternative grading practices that may enhance their ability to assess student learning. In addition to a negative view toward professional development and conversations about grading practices, Dr. Sean Maguire focuses on the teacher's use of grades, rather than purposes to inform students about their learning or take ownership of their learning. Despite the indication of adopting alternative grading practices as a participant in this study, the viewpoints of Dr. Sean Maguire appear more aligned with teachers using traditional grading practices and overlook the student benefits to alternative grading practices.

Non-factor loading participants: Professor Minerva McGonagall

“Five points ... will be awarded to each of you. For sheer, dumb luck.”

Professor Minerva McGonagall, “Harry Potter and the Sorcerer's Stone” (2001).

Professor Minerva McGonagall, from “Harry Potter and the Sorcerer's Stone”, was used to identify that group of participants that did not load on any of the five factors. This character was selected because of the above quote attributed to her in the film, rather than any personal characteristics, since it is more coincidence or dumb luck that these participants did not load on any of the factors.

**Table 9** displays the correlations between Q-sorts of Professor Minerva McGonagall. While there is some correlation between Q-sorts, Professor Minerva

McGonagall did not load significantly on any of the five factors and a six-factor model did not adhere to the acceptable eigenvalue cutoff of 1.0. Professor Minerva McGonagall included five participants who ranged in number of years teaching from one year to over 25 years, with two of the participants having more than 16 years of experience (67%). Professor Minerva McGonagall included one male and four female participants, all having earned a master’s degree, but one. All of the Professor Minerva McGonagall participants were from traditional high schools and included three math teachers, one fine arts teacher, and one English language arts teacher.

**Table 9.** Correlation between professor Minerva McGonagall.

	G4C9	G5Q5	IZAH	UUPF	Y4EE
G4C9	1.00	0.61	-0.02	0.23	0.37
G5Q5	-	1.00	0.21	0.12	0.24
IZAH	-	-	1.00	0.02	-0.15
UUPF	-	-	-	1.00	0.36
Y4EE	-	-	-	-	1.00

### 3.9. Additional analysis across variables

A Chi-square test is a statistical test commonly used to determine if there’s a significant association between two categorical variables [58]. For this study, a Chi-square test was used to compare the participants’ demographics across the five factors identified during the factor analysis. Analysis of the Chi-square test indicated little evidence to suggest a significant association between the five factors and the categorical variables. The closest association was between participants’ gender and factor. **Table 10** displays the observed frequencies of gender and factor.

**Table 10.** Observed frequencies for gender and factor.

Gender	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	No factor
Male	0	2	0	0	2	4
Female	7	1	4	2	1	1

The difference between the observed counts and the counts that would be expected if there were no association between the variables ( $X^2$ ) was 11.07. A higher value indicates a greater difference between the expected and observed data. The  $p$ -value indicates the probability of association between the variables. A common threshold for significance is 0.05. In this case, the  $p$ -value was calculated to be 0.05, suggesting that the observed association between gender and factor could reasonably occur by chance.

### 3.10. Post Q-sort reflection analysis

Following the Q-sort, participants were asked two open-ended questions. The first question asked how the participants’ graded practices reflect student-centered feedback. The second question asked participants how they know a student has shown growth. The open-ended questions provided participants with the opportunity to share greater details, in their own words, about their grading practices.

Participants’ answers were coded using words or phrases that summarize their responses. The first question was categorized based on the method used for feedback to students about their learning. The second question was categorized based on the activities the participants identified as measurements of student growth. **Table 11** reflects the coding for the first question.

**Table 11.** Distribution of coding for participants’ responses to open-ended question 1.

Code	Count
Rubrics	6
Student reflection activities	5
Student and teacher conference	7
Peer editing or revisions	1
Written feedback from teacher	9

**Table 12** reflects the coding for the second question.

**Table 12.** Distribution of coding for participants’ responses to open-ended question 2.

Code	Count
Formative assessments	3
Revisions or re-takes	4
Student self-tracking	4
Summative assessments	13

### 3.11. Post Q-sort viewpoints of Mr. Miyagi

Mr. Miyagi provides student-centered feedback in a variety of formats. Students receive feedback from Mr. Miyagi primarily through rubrics, student and teacher conferences, and written feedback. Participants 5TMU, 2JZ6, 76EA, and QXV0 all describe how they use rubrics and conferencing with students as a means for both the teacher and student to evaluate assignments. Participant 2JZ6 best captures the manner by which Mr. Miyagi provides student-centered feedback by stating, “When students write in my AP Language class, they have access to the mastery-oriented checklist and rubric from the start of the project. As students write, we revisit the rubric regularly. I conference with students daily and give them feedback throughout the process.”

Mr. Miyagi also measures student growth in a variety of ways. Formative and summative assessments are the most common means by which Mr. Miyagi examines student growth. Participants 5TMU, IIQ9, A3IT, and 21D8 identified formative and summative assessments as ways they examine student growth. In addition, Mr. Miyagi has students track their own growth. Participant QXV0 explains how Mr. Miyagi uses rubrics and conferencing to track students’ growth:

Evidence of growth is shown by an initial analysis of the rubric where both the student and I evaluate the work independently. We then compare our rubrics and identify areas for discussion. At our next conferences, I would expect the student to show growth by advancing across the rubric to the areas that indicate mastery.

### **3.12. Post Q-sort viewpoints of Herman Boone**

Herman Boone provides student-centered feedback through the use of rubrics, reflection activities, and student and teacher conferences. Students receive feedback from Herman Boone primarily through conferences. Participants UTQ7 and LK7Y use conferencing in a manner that provides student-centered feedback and encourages student reflection about their learning. Participant GVCT indicated they use a reflective approach to encourage students to set goals and assess their achievement. Participant GVCT described their method for student-centered feedback by stating, “My students fill out a Google form at the beginning of the year to set goals, and they reevaluate themselves every quarter and at the end of the year.”

Herman Boone analyzes student growth through summative assessments and revisions. Herman Boone indicated summative assessment is used to measure student growth from one assessment to the next. In addition, Herman Boone allows students to revise or retake assessments in order to show growth and demonstrate mastery of content.

### **3.13. Post Q-sort viewpoints of John Keating**

John Keating primarily provides student-centered feedback through written feedback. Participants VV9Z, MCTG, and 1MRD all indicated written feedback was the primary manner through which they provided students with feedback. Participant VV9Z wrote, “I leave comments on every assignment with feedback that I think will benefit the students on the next assignment.” Participants MCTG, 1MRD, and 325T all shared their use of reflection activities and student and teacher conferencing as additional means of student-centered feedback. Participants MCTG and 1MRD use these methods to provide further depth to their written feedback. Participant 1MRD writes, “I use a self-evaluation tool that students edit and I add additional comments to. I also hold conferences with students throughout the year to discuss their thoughts about mastery and offer feedback.”

Student growth is monitored by John Keating through summative assessment. Students also track their growth and have revision opportunities from John Keating, but the primary method for observing student growth is summative assessment. Participant 1MRD shared a unique perspective that captured the essence of why some teachers prefer alternative grading practices over traditional grading practices. Participant 1MRD writes:

“By using rubrics and a self-assessment tool with my students I know more about their abilities than I have ever known with a traditional grading system. These tools now give me the student’s perspective in addition to my own in terms of what needs to be improved upon. I can see growth in terms of what I’d like to see the student do better with and what they would like to improve on individually.”

### **3.14. Post Q-sort viewpoints of Ms. Valerie Frizzle**

Ms. Valerie Frizzle indicated written feedback as the only manner in which feedback is provided to students. Written feedback is primarily provided in response to writing assessments or as a written narrative at the end of the grading period. Participant T94W wrote, “Since I teach English Language Arts, I am able to provide

more detailed feedback when assessing students' writing. I try to identify a couple of areas that a student can focus on for improvement.”

Student growth is measured by Ms. Valerie Frizzle through summative assessments. Participants TL45 and T94W indicated summative assessments provide them with the ability to determine if a student is making growth and mastering content. Participant TL45 captured these sentiments by writing, “From assessment to assessment, I can track the improvement the student has made in the main categories I am assessing.”

### **3.15. Post Q-sort viewpoints of Dr. Sean Maguire**

Rubrics, reflection activities, and written feedback are all means through which Dr. Sean Maguire provides student-centered feedback. Dr. Sean Maguire did not provide great detail or depth into the manner in which feedback is provided to students. Participant BNTZ gave the most details in regard to feedback, specifically the use of written feedback. Participant BNTZ wrote, “I have a running list of comments that I use, and many of them have links to video lessons/explanations that help the student see where they need to improve. It's easy to copy and paste that comment into the Doc (document). Then, I let the students revise their work for my comments.”

Summative assessment was the primary method by which Dr. Sean Maguire measured student growth. Participant RJE7 specifically stated using a pre-test and post-test to assess student growth. Participant BNTZ also shared the use of revisions as a means to assess student growth. Participant BNTZ wrote, “Their revisions help me see if they now understand the concepts.”

### **3.16. Post Q-sort viewpoints of Professor Minerva McGonagall**

While Professor Minerva McGonagall did not load significantly on any of the five factors, there were some commonalities among the post Q-Sort responses for this group of participants. Professor Minerva McGonagall uses rubrics and reflection activities as primary means of providing students with feedback. Participants IZAH, G4C9, G5Q5, and Y4EE mentioned their use of rubrics and student reflection activities. Participant IZAH writes, “I provide a 1–5 scale rubric in 5 areas, based on a 25-point value. 1 = little to no effort, 3 = meeting criteria 5 = exceeding expectation. Students get to grade themselves and then I get to grade them separately.”

The use of assessment was identified by Professor Minerva McGonagall as the primary means of measuring student growth. Professor Minerva McGonagall compares scores on summative and formative assessments as documentation of student growth. Professor Minerva McGonagall uses assessment to determine student progress toward learning objectives and goals. Participant G5Q5 best captures these sentiments by writing:

I know if a student has shown growth through formative and summative assessments. Formatively, as students are working toward a goal, I can see through their work and through conversation with them if they are growing/striving toward meeting their goal. Summative assessments should be data that represents their meeting of the goal.

## **4. Discussion**

While some educators and educational institutions have experimented with diverse grading approaches, there is a scarcity of research exploring the underlying motivations of teachers for adopting these methods, the sustainability of these grading practices, and effective support strategies for the implementation of alternative grading practices. This mixed-method study aimed to describe current grading practices in teachers' classrooms and investigate the developmental trajectory of these practices throughout teachers' careers. Furthermore, the research sought to identify the various influences, particularly leadership, on teachers' grading practices. The findings from this study may lead to enhancing grading practices with a focus on student learning, contribute to the effective implementation and continuation of alternative grading practices, and expand upon existing research supporting alternative grading approaches.

### **4.1. Interpretation of findings**

Participants in this study have dedicated themselves to developing innovative and thoughtful methods of providing students with feedback that is both meaningful and centered on their individual growth and academic achievements. The emphasis on growth and mastery shifts the focus away from arbitrary percentages, letters, and non-academic components in a student's grade. Instead, the participants in this study prioritize the development of students' skills, knowledge, and understanding, ensuring that assessment truly reflects their academic achievements. This commendable approach aims to enhance the learning experience and foster a growth mindset among students.

The primary methods employed by the participants to facilitate student-centered feedback include rubrics, conferences, and self-reflection activities. Rubrics serve as a valuable tool, offering clear and detailed criteria against which students' work is assessed [59,60]. This allows students to understand specific expectations and improvement areas, promoting clarity and direction in their learning [60]. Conferences provide an opportunity for students to engage in meaningful conversations with their educators [61]. These discussions focus on students' strengths, areas for growth, and strategies for improvement [61]. Through these conferences, students gain valuable insights into their performance and develop a deeper understanding of their learning journey. Self-reflection activities empower students to take ownership of their learning by reflecting on their progress, identifying areas for improvement, and setting personal goals [62]. This practice encourages students to become active participants in their education, fostering self-awareness and critical thinking skills.

By embracing a student-centered approach to feedback, the participants can provide students with a learning experience within a traditional grading system that nurtures growth, empowers individuals, and paves the way for academic success. This approach recognizes that feedback is not merely a means to evaluate progress, but also an opportunity to engage in meaningful dialogue, unlock potential, and inspire intellectual curiosity. By tailoring feedback to students' individual learning styles, strengths, and areas for improvement, teachers can create a safe and supportive space where students feel comfortable taking risks, asking questions, and

pushing the boundaries of their understanding [63]. This student-centered approach to feedback fosters a sense of ownership and responsibility in students, encouraging them to become active participants in their learning journey. It empowers students to reflect on their progress, identify their strengths, and develop strategies to overcome challenges [64]. Through this transformative educational experience, students not only acquire knowledge but also develop the critical thinking skills, problem-solving abilities, and resilience necessary to thrive in their personal, academic, and professional endeavors.

Teachers are a critical component in the successful implementation of educational reforms. The current study found that participants' social interactions with peers, professional development, conversations with colleagues, and professional experiences played a role in the development of their grading practices. These findings support previous research that suggests teachers who participate in social networks characterized by robust connections, expertise sharing, and meaningful interactions are better equipped to maintain innovative practices even in the face of challenging circumstances or widespread school-wide changes [48]. The intricate interplay of interpersonal relationships, cultural norms, and power dynamics within social networks could potentially drive the continued prevalence of traditional grading practices. Social networks could also be used to drive reform.

It is important for school leaders to be aware of the influence that social networks can have on teachers' grading practices. Initially, school leaders should identify teachers who are already experimenting with alternative grading practices. Providing these teachers with exposure to professional development related to alternative grading would help strengthen the teachers' grading practices and focus on student growth. Since teachers often lack knowledge of effective grading practices and rely on their own experiences as students and professionals to develop grading practices, school leaders should facilitate opportunities for teachers to learn from each other and share best practices associated with grading for student growth and mastery. The creation of alternative grading pilot programs could build further interest among teachers and lead to the collaborative development of reformed grading policies and practices. School leaders can play a critical role in developing a culture of innovation and change that supports teachers in adopting more effective grading practices.

The results from the current study also indicate a lack of pre-service training related to grading. All but one of the factor groups, Dr. Sean Maguire, indicated their undergraduate coursework and student teaching experience had done very little to prepare them for grading student work. As a result, teacher preparation programs should look to develop coursework and experiences that expose pre-service teachers to grading practices focused on student-centered feedback, growth, and mastery.

## **4.2. Limitations**

This study encountered several limitations, including a limited sample size of Ohio high school teachers employing alternative grading practices. The absence of a national database of such educators necessitated reliance on peer recommendations for participant identification. The data collection was based on participants' recollections, which may have introduced biased or incomplete accounts. Furthermore, the study's

scope was restricted by the exclusive collection of teacher perspectives, without incorporating input from administrators or students. The researchers' involvement in grading reform efforts posed a potential source of bias. Additionally, the QMethod Software's lack of interactive features limited participant support to a help box. Although a training video was provided, there was no assurance that all participants viewed it, potentially affecting data reliability.

### **4.3. Future directions**

Due to the intricate and entrenched nature of traditional grading practices, further research is imperative to support the adoption of alternative grading practices. Teachers might be hesitant to modify their grading methods due to uncertainty regarding effective implementation. Moreover, grading practices are deeply rooted in school culture, making educators reluctant to alter long-standing norms.

Future research endeavors should expand the scope of this study. Exploring the development of grading practices among elementary and middle school teachers, where alternative grading methods are often more prevalent, could provide a more comprehensive view of grading practices across various educational levels. Another avenue of research would be to examine the influence of technological tools such as learning management systems, digital portfolios, and online gradebooks on the evolution of grading practices, specifically, how these technological advancements have influenced the development of teachers' grading practices. Given that assessment of learning outcomes is a vital component of grading, it would also be valuable to evaluate student achievement and engagement before and after implementing alternative grading practices to assess their effectiveness. Finally, further investigation of how parents perceive and respond to teachers experimenting with alternative grading practices would further aid in the successful implementation of alternative grading policies.

Additional research methods could be utilized to better understand the development of teacher grading practices and the use of alternative grading methods. A qualitative approach focused on interviews with teachers who have experimented with alternative grading practices could provide a deeper understanding of their individual motivations, challenges, and successes in implementing alternative grading practices. Observation of teacher interactions and meetings could provide greater context to support the impact of social network theory and collective leadership.

To bridge the existing knowledge gap, this study aimed to delve deeper into the development and sustainability of alternative grading practices among teachers. By expanding upon the current study's research questions, researchers would better understand how grading practices are established, why teachers modify their grading practices, and how educators maintain alternative grading practices within conventional grading policies.

Specifically, further investigation should include:

- The influence of teacher preparation programs, mentoring, and current professional development on grading practices.
- The impact of external mandates, such as standardized testing and accountability measures, on changes in grading practices.

- Strategies teachers employ to overcome resistance to alternative grading practices from colleagues, administrators, or parents.
- Identification of specific factors or individuals who have been instrumental in the dissemination of alternative grading practices within and across school districts.

By addressing these specific areas, additional research would provide a more comprehensive analysis of the complexities surrounding the establishment and sustainability of alternative grading practices.

## **5. Conclusion**

Serving in various roles as an educator for nearly 20 years has made it clear there are flaws with traditional grading practices. While educators have experimented with alternative grading practices, there has been limited study of how to build consensus around the need to develop improved grading practices and systematic reform of how student learning is evaluated. The ability for school leaders and teachers to identify like-minded individuals is important in creating sustained change to grading practices. This study explored the establishment, change, sustainability, and spread of grading practices among Ohio high school teachers. Findings reveal that teachers primarily rely on personal experience and limited training to develop grading practices. While some teachers engage in informal discussions with colleagues, the lack of structured professional development and teacher preparation related to grading is prevalent.

Teachers' motivations to change grading practices stem from a desire to provide more meaningful feedback and accurately reflect student learning. However, the constraints of traditional grading systems, such as the emphasis on numerical grades and external pressures, often hinder the implementation of alternative approaches. Despite these challenges, teachers have implemented strategies like rubrics, conferences, and self-reflection to provide more formative, student-centered feedback.

The complexity and established nature of grading practices can further influence teachers' willingness to adopt alternative approaches. Teachers may be hesitant to change their grading practices because they are not sure how to do so effectively. Additionally, grading practices are often deeply ingrained in school culture, and teachers may be reluctant to change something that has been in place for many years. Given these factors, teachers' social networks could play a significant role in shaping their grading practices. Teachers surrounded by colleagues open to change are more likely to experiment with different grading approaches. Conversely, teachers surrounded by colleagues who are resistant to change are less likely to change their own grading practices. This study highlights the need for increased support and professional development for teachers in developing effective grading practices. By fostering a culture of collaboration and innovation, schools can create environments where teachers feel empowered to implement alternative grading approaches that prioritize student learning and growth.

**Author contributions:** Conceptualization, SV and KHL; methodology, SV and KHL; software, SV and KHL; validation, SV and KHL; formal analysis, SV and KHL; investigation, SV; resources, SV; data curation, SV; writing—original draft preparation, SV; writing—review and editing, SV and KHL; visualization, SV;

supervision, KHL; project administration, SV and KHL. All authors have read and agreed to the published version of the manuscript.

**Institutional review board statement:** The study was conducted in accordance with the Declaration of Helsinki, and approved by the YSU Institutional Review Board at Youngstown State University (Protocol code: 2024-116; Date of approval: 11 April 2024).

**Informed consent statement:** Informed consent was obtained from all subjects involved in the study per the YSU IRB specifications.

**Conflict of interest:** The authors declare no conflict of interest.

## References

1. Durm MW. An A is not an A is not an A: A History of Grading. *The Educational Forum*. 1993; 57(3): 294-297. doi: 10.1080/00131729309335429
2. Schneider J, Hutt E. Making the grade: a history of the A–F marking scheme. *Journal of Curriculum Studies*. 2013; 46(2): 201-224. doi: 10.1080/00220272.2013.790480
3. Guskey TR. The case against percentage grades. *Educational Leadership*. 2013; 71(1): 68-72.
4. Brookhart SM, Guskey TR, Bowers AJ, et al. A Century of Grading Research. *Review of Educational Research*. 2016; 86(4): 803-848. doi: 10.3102/0034654316672069
5. Starch D, Elliott EC. Reliability of the Grading of High-School Work in English. *The School Review*. 1912; 20(7): 442-457. doi: 10.1086/435971
6. Brimi HM. Reliability of grading high school work in English. *Practical Assessment, Research, and Evaluation*. 2011; 16(1): 1-17.
7. Malouff JM, Thorsteinsson EB. Bias in grading: A meta-analysis of experimental research findings. *Australian Journal of Education*. 2016; 60(3): 245-256. doi: 10.1177/000494411666461
8. Hu S. Beyond grade inflation: Grading problems in higher education. *ASHE Higher Education Report*. 2005; 30(6): 1-90.
9. Quinn DM. How to reduce racial bias in grading: New research supports a simple, low-cost teaching tool. *Education Next*. 2021; 21(1): 72-78.
10. Hinnerich BT, Höglin E, Johannesson M. Discrimination against students with foreign backgrounds: evidence from grading in Swedish public high schools. *Education Economics*. 2014; 23(6): 660-676. doi: 10.1080/09645292.2014.899562
11. Bygren M. Biased grades? Changes in grading after a blinding of examinations reform. *Assessment & Evaluation in Higher Education*. 2019; 45(2): 292-303. doi: 10.1080/02602938.2019.1638885
12. Angelo C, Reis AB. Gender gaps in different grading systems. *Education Economics*. 2020; 29(1): 105-119. doi: 10.1080/09645292.2020.1853681
13. Coates J, Draves WA. Smart boys, bad grades. *Learning Resources Network*. 2006.
14. Tomlinson CA. Teach up for equity and excellence. *Educational Leadership*. 2023; 80(8): 28-34.
15. Doyle L, Easterbrook MJ, Harris PR. Roles of socioeconomic status, ethnicity and teacher beliefs in academic grading. *British Journal of Educational Psychology*. 2022; 93(1): 91-112. doi: 10.1111/bjep.12541
16. Carifio J, Carey T. A Critical Examination of Current Minimum Grading Policy Recommendations. *The High School Journal*. 2009; 93(1): 23-37. doi: 10.1353/hsj.0.0039
17. Cross LH, Frary RB. Hodgepodge grading; Endorsed by students and teachers alike. *National Council on Measurement in Education*; 1996. doi: 10.1207/s15324818ame1201\_4
18. Erickson JA. How grading reform changed our school. *Educational Leadership*. 2011; 69(3): 66-70.
19. Guskey TR, Link LJ. Exploring the factors teachers consider in determining students' grades. *Assessment in Education: Principles, Policy & Practice*. 2018; 26(3): 303-320. doi: 10.1080/0969594x.2018.1555515
20. O'Conner K, Wormelli R. Reporting student learning. *Educational Leadership*. 2011; 69(3): 40-44.
21. Westphal A, Lazarides R, Vock M. Are some students graded more appropriately than others? Student characteristics as moderators of the relationships between teacher-assigned grades and test scores in mathematics. *British Journal of*

- Educational Psychology. 2020; 91(3): 865-881. doi: 10.1111/bjep.12397
22. Chen PP, Bonner SM. Teachers' Beliefs About Grading Practices and a Constructivist Approach to Teaching. *Educational Assessment*. 2016; 22(1): 18-34. doi: 10.1080/10627197.2016.1271703
  23. Guskey TR. Bound by tradition: Teachers' views of crucial grading and reporting issues. In: *Proceedings of the Annual Meeting of the American Educational Research Association*; 2009; San Francisco, CA, United States.
  24. Percell JC. Lessons from Alternative Grading: Essential Qualities of Teacher Feedback. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*. 2017; 90(4): 111-115. doi: 10.1080/00098655.2017.1304067
  25. Frankin A, Buckmiller T, Kruse J. Vocal and Vehement: Understanding Parents' Aversion to Standards-Based Grading. *International Journal of Social Science Studies*. 2016; 4(11). doi: 10.11114/ijsss.v4i11.1923
  26. Sackstein S. Earning good grades versus learning. *Education Week*. 2018.
  27. Townsley M, Buckmiller T. What does the research say about standards-based grading? A research primer. *Faculty Publications*; 2026.
  28. Feldman J. Beyond standards-based grading: Why equity must be part of grading reform. *Phi Delta Kappan*. 2019; 100(8): 52-55. doi: 10.1177/0031721719846890
  29. Moore C. *Alternatives to traditional grading*. Oakland University; 2023.
  30. Ajawd MI, Bilo S. Seen and unseen effects of COVID-19 school disruptions. *Brookings*; 2022.
  31. Patrinos HA, Vegas E, Carter-Rau R. COVID-19 school closures fueled big learning losses, especially for the disadvantaged. *World Bank*; 2022.
  32. Sawchuk S. Grading students during the coronavirus crisis: What's the right call? *Education Week*. 2020.
  33. Strauss V, George D. Teachers second-guess letter grades as they search for a fairer way. *The Washington Post*, 28 February 2022.
  34. Sultana F, Bari R, Munir S. Impact of school closures due to COVID-19 on education in low- and middle-income countries. *Journal of Global Health Reports*. 2022; 6. doi: 10.29392/001c.36902
  35. Kohn A. The accidental benefits of COVID-19. *Education Week*. 2020; 40(1): 20.
  36. Townsley M. Grading Principles in Pandemic-Era Learning: Recommendations and Implications for Secondary School Leaders. *Journal of School Administration Research and Development*. 2020; 5(S1): 8-14. doi: 10.32674/jsard.v5is1.2760
  37. McMillan JH. Secondary Teachers' Classroom Assessment and Grading Practices. *Educational Measurement: Issues and Practice*. 2001; 20(1): 20-32. doi: 10.1111/j.1745-3992.2001.tb00055.x
  38. Link L. Teachers' perceptions of grading practices: How pre-service training makes a difference. *Journal of Research in Education*. 2018; 28(1): 62-91.
  39. Grainger PR, Adie L. Title: How do Preservice Teacher Education students Move From Novice to Expert Assessors? *Australian Journal of Teacher Education*. 2014; 39(7). doi: 10.14221/ajte.2014v39n7.9
  40. Bergman DJ. Pre-service teachers' pliable perceptions of ethical practices in student evaluation. *Issues in Teacher Education*, 27(3). 2018; 5-23.
  41. Grogan M, Shakeshaft C. A new way. In: *The Jossey-Bass reader on educational leadership*. Wiley; 2013. pp. 111-130.
  42. Louis KS, Leithwood K, Wahlstrom KL, Anderson SE. *Learning from leadership: Investigating the links to improved student learning*. The Wallace Foundation; 2010.
  43. Hallinger P, Heck RH. Collaborative leadership and school improvement: understanding the impact on school capacity and student learning. *School Leadership & Management*. 2010; 30(2): 95-110. doi: 10.1080/13632431003663214
  44. Gano-Phillips S, Barnett RW, Kelsch A, et al. Rethinking the Role of Leadership in General Education Reform. *The Journal of General Education*. 2011; 60(2): 65-83. doi: 10.5325/jgeneeduc.60.2.0065
  45. Daly AJ, Moolenaar NM, Bolivar JM, et al. Relationships in reform: the role of teachers' social networks. *Journal of Educational Administration*. 2010; 48(3): 359-391. doi: 10.1108/09578231011041062
  46. Nimmon L, Artino AR, Varpio L. Social Network Theory in Interprofessional Education: Revealing Hidden Power. *Journal of Graduate Medical Education*. 2019; 11(3): 247-250. doi: 10.4300/jgme-d-19-00253.1
  47. Dokuka S, Valeeva D, Yudkevich M. How academic achievement spreads: The role of distinct social networks in academic performance diffusion. *PLOS ONE*. 2020; 15(7): e0236737. doi: 10.1371/journal.pone.0236737
  48. Coburn CE, Russell JL, Kaufman JH, et al. Supporting Sustainability: Teachers' Advice Networks and Ambitious Instructional Reform. *American Journal of Education*. 2012; 119(1): 137-182. doi: 10.1086/667699
  49. Datnow A. Teacher Agency in Educational Reform: Lessons from Social Networks Research. *American Journal of*

- Education. 2012; 119(1): 193-201. doi: 10.1086/667708
50. Grunspan DZ, Wiggins BL, Goodreau SM. Understanding Classrooms through Social Network Analysis: A Primer for Social Network Analysis in Education Research. Dolan E, ed. *CBE—Life Sciences Education*. 2014; 13(2): 167-178. doi: 10.1187/cbe.13-08-0162
  51. Desing R, Kajfez R. How to Use Q Methodology in Engineering Education Research. In: *Proceedings of the 2020 ASEE Virtual Annual Conference Content Access Proceedings*. doi: 10.18260/1-2--34737
  52. Coogan J, Herrington N. Q methodology: An overview. *Research in Secondary Teacher Education*. 2011; 1(2): 24-28.
  53. Irie K. Q methodology for post-social-turn research in SLA. *Studies in Second Language Learning and Teaching*. 2014; 4(1): 13-32. doi: 10.14746/ssllt.2014.4.1.2
  54. Lundberg A, de Leeuw R, Aliani R. Using Q methodology: Sorting out subjectivity in educational research. *Educational Research Review*. 2020; 31: 100361. doi: 10.1016/j.edurev.2020.100361
  55. Valenta AL, Wigger U. Q-methodology: Definition and Application in Health Care Informatics. *Journal of the American Medical Informatics Association*. 1997; 4(6): 501-510. doi: 10.1136/jamia.1997.0040501
  56. Damio SM. Q methodology: An overview and steps to implementation. *Asian Journal of University Education*. 2016; 12(1): 105-122.
  57. Watts S, Stenner P. *Doing q methodological research: Theory, method & interpretation*. Sage; 2012.
  58. Franke TM, Ho T, Christie CA. The Chi-Square Test. *American Journal of Evaluation*. 2011; 33(3): 448-458. doi: 10.1177/1098214011426594
  59. Mueller J. The authentic assessment toolbox: Enhancing student learning through online faculty development. *Journal of Online Learning and Teaching*. 2005; 1(1): 1-7.
  60. Wolf K, Stevens E. The role of rubrics in advancing and assessing student learning. *The Journal of Effective Teaching*. 2007; 7(1): 3-14.
  61. Shum DL. Empower students through individual conferences. *ASCD*. 2019; 14(22).
  62. Lektorich B. Empowering students: How to foster self-directed learning. *Education World*. 2023.
  63. Stenger M. 5 research-based tips for providing students with meaningful feedback. *Edutopia*. 2014.
  64. Hattie J, Timperley H. The power of feedback. *Review of Educational Research*. 2007; 77(1): 81-112.

## **Appendix**

Concourse statements for final Q-set:

- 1) Grading practices are important measures of student achievement.
- 2) Grading is a good method for helping students identify their weaknesses in a content area.
- 3) Grading can keep students informed about their progress.
- 4) Grading provides information about student mastery.
- 5) I consider student effort when I grade.
- 6) Grades can motivate students to learn.
- 7) Grades of zero can demotivate students to learn.
- 8) Grades are an imprecise measure of learning.
- 9) Factors other than a students' actual achievement on a test or quiz make it difficult for me to grade.
- 10) In general, grades often contain information other than achievement of learning objectives which diminishes accuracy.
- 11) When I first started teaching, my experience with grading was very procedural and a high percentage of my grade was completion-based.
- 12) When I first started I either graded too harshly or too easily.
- 13) When I first started teaching, grades were something I felt like I had to do to provide feedback for the students.
- 14) When I first started teaching, grades were not something I thought about.
- 15) My current grading practices are a better reflection of student mastery compared to when I first started teaching.
- 16) When I first started teaching, a student's grade was defined by a number, subtracting the number wrong and dividing by total points.
- 17) I feel my current grading practices can be used as a tool to help me be a better teacher.
- 18) I use the data I receive from my current grading practices to adjust how I instruct in the classroom.
- 19) My current grading practices provide descriptive feedback to students.
- 20) My current grading practices document my instructional effectiveness.
- 21) My experiences during the pandemic led me to change my grading practices.
- 22) My peers have influenced my grading practices.
- 23) A professional development opportunity (speaker, book study, conference, etc.) prompted me to change my grading practices.
- 24) I have to modify my mastery or standards-based practices to fit within a traditional grading system.
- 25) I no longer assign a grade for everything students complete.
- 26) I do not give a grade based on completion of the assignment.
- 27) Compared to other teachers in my school, I am more willing to allow students to redo assignments in order to make sure a grade reflects student mastery.
- 28) Compared to other teachers in my school, I am more lenient on late assignments in order to make sure a grade reflects student mastery.
- 29) I closely examine learning objectives before I design an assessment.
- 30) I think about different ways students can demonstrate achievement, and I try to use a number of different assessment methods.
- 31) I consider whether students are demonstrating understanding of the learning objectives being taught.
- 32) I consider how well class time is used to complete assignments.
- 33) One way I have maintained mastery grading practices is by creating my own rubrics.
- 34) My students reflect on their own work and assess they would grade themselves.
- 35) I have individual conferences with students about their learning where we discuss their self-assessments.
- 36) I often reflect on my grading practices and how I can better use data from a variety of assessments to help improve

my instruction.

- 37) I have a few colleagues with whom I discuss grading practices and discuss ideas about grading.
- 38) I would say compared to the rest of the teachers in my school, I have a unique grading practice.
- 39) There are a few teachers in my school who are experimenting with changing their scale and looking at a more standards-based approach.
- 40) I might listen to advice from administrators and colleagues on grading practices.
- 41) I have to keep my students' needs foremost in my considerations when grading and designing assessments.
- 42) An administrator influenced the development of my grading practices.
- 43) My undergraduate coursework and/or student teaching experience was influential in the development of my current grading practices.
- 44) Anything to make me think about grading practices was done on my own in my online professional learning communities.
- 45) Anything to make me think about grading practices was through informal professional conversations.