

Teacher Grading Decisions: Influences, Rationale, and Practices

Author

JOSHUA P. KUNNATH, Ed.D., is an educational researcher for California State University, Fresno and a high school English Teacher at Highland High School in Bakersfield, California.

Abstract

This mixed-methods study applied a decision-making theoretical framework to an investigation of teacher grading in a large urban school district in California. A sample of 251 high school teachers of core subjects were surveyed, and 15 teachers participated in four focus group interviews in order to provide data on the influences, rationale, and practices of the grading decision-making process. Findings identified three important influences on grading decisions and two components of teachers' rationales for grading. The findings also indicated that teachers' grading practices included factors such as students' effort and ability level as well as their achievement. The discussion, presented in the context of relevant research, describes multiple sources of subjectivity that contributed to inaccurate and inflated grades.

Key words: *grading, grades, decision making, report cards, secondary teachers, grading reform*

Student report card grades are an integral part of K-12 education, and grading is an important teacher responsibility, but experts are in agreement that grades largely fail to accomplish their main purpose of communicating student academic achievement (Cizek, Fitzgerald, & Rachor, 1995; Guskey, 2015; McMillan, 2001; Stiggins, Frisbie, & Griswold, 1989). Instead, because student grades are often created from an imprecise amalgam of factors that includes achievement, ability, effort, and behavior (Cross & Frary, 1999; Randall & Engelhard, 2010; Reeves, 2011), exactly what they communicate is rarely clear (Allen, 2005; Marzano, 2000). Failure to use grades for their primary purpose leaves educational stakeholders largely uninformed about

students' true level of learning. This may lead to such negative consequences as passing students with deficient knowledge and skills, misinforming parents of student learning, and misrepresenting student merit to colleges and universities (Allen, 2005; Reeves, 2011). Although a substantial amount of literature has documented the discrepancy between expert recommendations and teacher grading practices and has called for reform (e.g., Cizek et al., 1995; Cross & Frary, 1999; Randall & Engelhard, 2010), grading practices have remained largely unchanged in schools throughout the country (Reeves, 2011).

Teacher Grading Decision-Making

Grading is an especially important area for teacher decision-making (Allen, 2005; Sadler, 2009), and requires an effective teacher decision-making process (Cauley & McMillan, 2000; McMillan, 2003). The literature on teacher decision-making about grading is sparse, but it does include some promising initial findings. Frary, Cross, and Weber (1993) and McMillan (2001), for instance, found that teachers' grading practices often differ from their own grading beliefs. Randall and Engelhard (2010) examined teachers' grading decisions in cases of borderline grades (e.g. grades at the border of an A and B, B and C, etc.) and found teachers made grading decisions based more on student ability, behavior, and effort than on achievement. The authors claimed that teachers regularly rely on a combination of factors to make grading decisions such as personal philosophy, college classes and professional development, school or district policy, and perceived consequences, but they provided no systematic way of interpreting the process.

In an article advocating for the use of grades as valid measures of student academic achievement, Allen (2005) suggested that a solution to enabling teachers to create valid grades is helping them to make good grading decisions. He explained that teacher education programs may accomplish this by challenging future teachers' previously held knowledge and beliefs about grades, by presenting strong measurement theory, by providing grading practice opportunities before student teaching, and by facilitating students' creation and implementation of a valid individual evaluation and grading plan.

McMillan and Nash (2000) and McMillan (2003) have likely shed the most light on teacher decision-making about grading. McMillan and Nash (2000) conducted a qualitative study of 24 elementary and secondary English and mathematics teachers that explored teacher grading and assessment decision-making. Findings showed that teachers' idiosyncratic grading practices were determined by a decision-making process that largely balanced teachers' personal learning and beliefs about the topic, with classroom reality and external pressures. McMillan and Nash (2000) explained that the tension created in balancing these three influences leads to varying grading practices within and across classrooms. McMillan (2003) further analyzed

these findings in a follow-up article and created a refined grading decision-making model.

Considering the current state of teacher grading practices, the importance of grading decision-making, and the dearth of research on the latter topic, further research on teacher grading decision-making seems warranted. At the conclusion of his study on secondary teachers' grading and assessment practices, McMillan (2001) echoed this sentiment: "There is a need to more fully understand how teachers make decisions about their grading practices. What rationales are used? What factors are considered? Why are some factors included and others not included?" (p. 30). More recently, at the conclusion of their study on the factors and assessment methods involved in the grading decision-making of Chinese secondary teachers, Cheng and Sun (2015) reiterated the need for research, specifically research using qualitative methods, to explain grading decision-making.

Theoretical Framework

This study adapted the theoretical framework from McMillan's (2003) teachers' classroom assessment decision-making model, which is displayed in Figure 1. The first domain (of three) in the framework, referred to in the current study as *Grading Influences*, is composed of three distinct components. The first is *teacher knowledge, beliefs, expectations, and values*, which entails elements of teacher cognition arising from teacher education classes, teaching experience, and other life experiences that may apply to the classroom. The second component, *external factors*, consists of those elements outside of the classroom that influence grading, such as parents, administrators, and state testing. The third component, *classroom realities*, consists of elements within the classroom that affect grading, such as heterogeneity of students, motivation, and behavior.

The second domain in the framework is *Teacher Decision-making Rationale*, which is the reasoning that teachers employ for their grading practices. The third domain, *Grading Practices*, is composed of the particular practices that teachers use to create student report card grades (McMillan, 2003). The framework suggests that the complex interaction of influences in Domain 1 often leads to compromises in teacher knowledge, beliefs, expectations, and values that impact grading rationale (Domain 2). This may lead to a hodgepodge of grading practices (Domain 3) that results in inaccurate student report card grades.

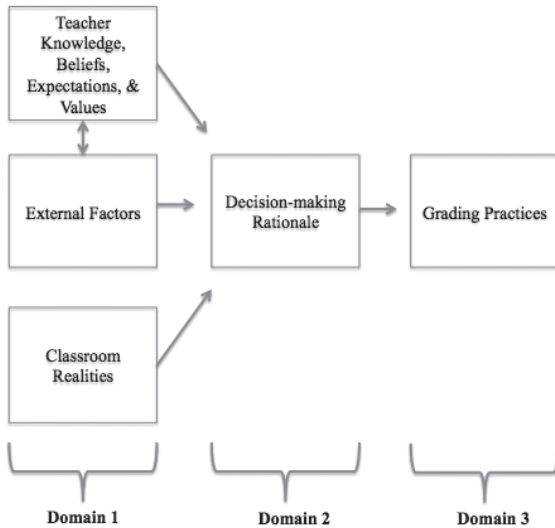


Figure 1. Teacher grading decision-making theoretical framework.

Adapted from "Understanding and Improving Teachers' Classroom Assessment Decision Making: Implications for Theory and Practice," by J. H. McMillan, 2003, *Educational Measurement: Issues and Practice*, 22(4), p. 36. Copyright 2005 by John Wiley and Sons. Adapted with permission.

Purpose of the Study

The purpose of this study was to better understand the grading decision-making process that high school teachers use to create student report card grades in order to facilitate future work in grading reform. Guided by the theoretical framework, this study posed three research questions:

1. What influences affect teacher grading practices?
2. What grading rationale do teachers use to produce student report card grades?
3. What grading practices do teachers use to create student report card grades?

Method

Research Design

The complexity of the teacher grading decision-making process (Wilen et al., 2008) necessitated the use of both quantitative and qualitative data sources to understand the intricacies of the three domains of the theoretical framework. A convergent mixed methods design was used to address the three research questions, as both quantitative and qualitative methods were deemed equally important, and thus equally prioritized (Creswell & Plano Clark, 2011). Quantitative methods were used to provide an overview of grading influences and teacher practices while qualitative methods were used for

triangulation and to provide a detailed understanding of teachers' grading rationale. The two strands were kept separate until data interpretation, because the nature of both data sets was deemed best understood independently to ensure the essence of each was uninfluenced by the complimentary strand (Creswell & Plano Clark, 2011).

Participants

Teachers from a large urban school district in California participated in the study. The district, selected for convenience purposes, contained close to 20 comprehensive high schools. A student body of over 37,000 students was predominantly Latino (63%), and 64% qualified for free or reduced price lunch (FRPL). The district employed approximately 1,500 teachers, with an average teaching experience of 15 years. The survey sample was obtained through comprehensive sampling of all regular education teachers of English, foreign language, mathematics, science, and social studies, which led to invitations being sent to 915 teachers. A total of 251 teachers from all but one school in the district completed the quantitative portion of the survey for a response rate of 27%. The sample consisted of mostly veteran teachers, as 86% taught 7 years or more, and just 2% taught fewer than 3 years. Of five academic subjects, English teachers comprised the largest number of respondents (37%), followed by teachers of science (22%), mathematics (20%), social studies (18%), and foreign language (4%). The vast majority of teachers taught 50% or more college preparatory classes in their schedule (86%). A total of 121 survey respondents that completed the quantitative items also included qualitative data.

Purposive sampling was used to select focus group school sites from low-, mid-, and high-poverty schools (as measured by the proportion of students who qualified for FRPL). School poverty level was considered because of the wide range of school FRPL rates within the district (from 17.4% to 89.2%) and the reported influence of school poverty level on teacher grading (Cauley & McMillan, 2000). Snowball sampling was used to select participants, as principals at each site were asked to recruit five to eight teachers based on two criteria: (a) they taught regular education and (b) they taught English, foreign language, mathematics, science, or social studies. These procedures produced four school sites and a total of fifteen teachers in four different focus groups, with each focus group consisting of two to five teachers. The FRPL rates of participating schools ranged from 49.6% to 89.5%.

Data Collection

An original 35-item web-based questionnaire was developed, pilot tested, and employed to collect quantitative and qualitative data on grading influences and grading practices. The instrument used survey elements from the grading studies of Cross and Frary (1999), Frary et al. (1993), McMillan (2001), and McMillan and Lawson (2001). The core survey items (30 items)

used a five-point Likert-type scale, ranging from “(1) not at all” to “(5) completely.” Grading practices were measured with 17 items using the question stem “To what extent were your final second semester student grades in your primary teaching assignment based on _____?”

Grading influences were measured with 13 items using the question stem “How influential is _____ on your grading practices?” These items contained a Likert-type scale that ranged from “(1) not at all” to “(5) extremely.” Five additional items collected background information on participants. The survey collected qualitative data by providing participants with an optional open-ended comment box following each close-ended response item on the 17 grading practices and 13 grading influences.

The primary source of qualitative data about influences on grading was semi-structured focus group interviews. The interview protocol, which was piloted at a local high school, was consisted of an introduction to the study and five guiding questions aligned to the theoretical framework and research questions. The guiding questions asked teachers about grading challenges; grading procedures; and the influence of their own education philosophy and beliefs, external factors, and classroom realities on their grading. Each focus group interview, ranging from 43 to 57 minutes in length, was digitally recorded and transcribed by the researcher within 48 hours of its completion.

Data Analysis

Quantitative data were analyzed with descriptive statistics, focusing on 17 teacher grading practices and 13 grading influences. Qualitative data were analyzed with a constant comparative analysis method (Corbin & Strauss, 2008). Transcribed focus group data were combined with the survey open-response data in NVivo 11.0 (citation?) before undergoing three rounds of coding. Analysis commenced with open coding to generate a rough understanding of the data and to create data categories. Axial coding followed to reconfigure and reorganize emergent categories (Saldana, 2012). Reorganization was conducted around one category designated as the *core phenomenon*, which represented a central concept found in the data (Corbin & Strauss, 2008). Finally, a selective coding process used refined categories and the theoretical framework to interpret emergent themes (Corbin & Strauss, 2008).

Findings

Grading Influences

In the survey data, the two influences on teacher grading practices with the highest mean scores, measuring just under *very influential*, were *desire to promote student understanding* ($M = 3.95$) and *philosophy of teaching and learning* ($M = 3.93$). Teachers also rated three factors between *somewhat* and *very influential*: *desire to accommodate student individual differences and needs* ($M = 3.35$), *desire for student success* ($M = 3.31$), and *student motivation and engagement* ($M = 3.13$). Complete results are displayed in Table 1.

Table 1.
Survey Results of 13 Influences on Teacher Grading Practices

Grading Influence	Percentages					M	SD
	(1) Not at all	(2) Slightly	(3) Somewhat	(4) Very	(5) Extremely		
Desire for student success	10	18	24	28	20	3.31	1.25
Philosophy of teaching and learning	2	6	20	41	32	3.93	.94
Desire to promote student understanding	2	5	19	42	31	3.95	.95
Desire to accommodate student individual differences and needs	4	14	37	31	13	3.35	1.01
Student motivation and engagement	10	18	35	25	12	3.13	1.15
State standardized testing	61	18	14	5	2	1.70	1.03
Formal or informal school or district policies	34	34	23	7	2	2.10	1.02
School administrators	52	27	13	6	2	1.79	1.00
Parents	68	22	9	1	0	1.43	.70
Student absenteeism	26	25	33	12	4	2.43	1.12
Student disruptive behavior	58	31	10	1	1	1.55	.75
Differing student ability levels in a class	20	32	37	8	3	2.41	.99
Student disruptive and/or non-supportive home environments	50	28	19	2	1	1.75	.88

Analysis of the qualitative data from open-ended survey responses and focus groups identified three themes to explain grading influences. Table 2 provides examples of significant teacher statements and formulated meanings.

Theme 1: Concern about external perceptions. Teachers expressed concern about judgment from other teachers and administrators on their professional efficacy based on the grades they assign. Participants described a

desire to use and be perceived as using sound grading practices, which were viewed as fair and consistent practices that resulted in a correlation between student learning and the resulting grade. One participant questioned his own practices and expressed his concern about others' practices: "I always wonder, like what is everybody else doing, and am I playing the same game as everybody else? Or am I just being obstinate?" Another teacher questioned the grading practices of teachers who assign many high grades by asking, "Have you really challenged those kids . . . or did you just pencil whip it for your pure ego so you can go to the office and say, 'Hey, these kids got A's and B's—look how wonderful I am.'" This teacher acknowledged that teachers might assign high student grades because of the positive perception they receive by doing so—especially by administrators.

Table 2.
Examples of Significant Statements and Formulated Meanings for Emergent Themes Addressing Research Question 1

Significant Statement	Formulated Meaning
Theme 1: Concern about external perceptions	
<ul style="list-style-type: none"> • "Even my own external perception [is a factor in grading]. I don't want to be the teacher who all their kids have C's." • "There is an ego thing for a teacher that you know, have you really challenged those kids...and does the grade show that?" 	<ul style="list-style-type: none"> • Being known as an incompetent grader is a concern. • Teachers may inflate grades because of the positive perception that may result.
Theme 2: Desire to facilitate student success	
<ul style="list-style-type: none"> • "My students know if there is something going on at home that is impacting their ability to complete work, they just need to let me know and I will make accommodations for them." • "I will sometimes give students a better grade than they have earned on a particular assignment because I know (or think) that the student is capable/knowledgeable but is perhaps not applying themselves." 	<ul style="list-style-type: none"> • If students experience a personal hardship, the teacher will work to improve the grade. • Grades are inflated if the teacher feels the student has the knowledge or skills, but has not put forth the effort to provide evidence
Theme 3: Administrator pressure on assigning low grades	
<ul style="list-style-type: none"> • "I certainly don't want to be the teacher who is in the office for getting spoken to about why my gradebook—why kids have bad grades." • "I've never been told by the administration to change a grade—never—in 30 years...I am aware however of some teachers...who had a great deal of pressure because their high-level senior kids were getting bad grades, and there was pressure to have grades changed." 	<ul style="list-style-type: none"> • A concern exists about administrators perceiving a problem in teaching practices because of excessive student D's and F's. • Not everyone experiences direct pressure by administrators about grades, but teachers are aware of it happening on campus.

Theme 2: Desire to facilitate student success. Teachers described their use of grades to facilitate student success when students experience exceptional circumstances, such as a death in the family, illness, learning disabilities, problems at home, or an uncharacteristically poor test performance. Teachers explained they address these factors by excusing work or test grades, providing additional time, or relying more on nonachievement factors. One teacher explained, “There’s a kid who this year in my class had to leave the country—his father passed away. He came back, he was behind, and I just dismissed the work during that time.” Teachers also explained using grades to aid the success of graduating seniors. These teachers felt pressure to do everything within reason to help students to pass the class so they could graduate, and this influence was especially strong with students who were on the border of passing the class. A senior government teacher described the pressure that she experienced at the end of the school year: “As I sit there and create those grades, you know, mainly in May . . . it’s not that easy when you’re making that decision right there—are you going to graduate or not.”

Finally, teachers explained that sometimes it is simply the desire for their students to obtain high grades that influences them to use lenient practices. This includes teachers providing test or assignment retake opportunities, manipulating grade categories, and increasing the use of nonachievement factors. One teacher explained her belief that students need multiple opportunities to show their learning on assessments to earn high grades: “There is the idea of giving the kids the second chance and third chance and believing that they can do it.” Another teacher provided her rationale for using nonachievement factors to create a report card grade: “I don’t think that if they just don’t understand it they should be completely penalized.”

Theme 3: Administrator pressure on assigning low grades. Teachers explained a small, yet existent pressure from administrators on the assigning of D’s and F’s. Some participants explained this pressure as an unwritten rule that they were not to assign too many failing grades. Although no one defined “too many” in this context, teachers did explain that once a teacher broke this threshold it brought negative results. Multiple participants explained that it was common knowledge that assigning too many F’s would likely result in a request to discuss the grades with an administrator in the office. A veteran teacher explained, “If you’re giving too many F’s, what happens? ‘Hey, can you come up to the principal’s office today after school?’ . . . or you find that your schedule’s been changed because there’s just too many F’s happening.”

While not all participants experienced direct grading pressure from an administrator, most knew of someone who had, and teachers explained that this knowledge also added pressure on their practices. Before assigning report card grades, teachers described addressing low student grades to avoid administrator attention. These actions included re-grading student work or pressuring students to complete missing work to bring up their grades. One participant explained both of these phenomena: “Now I do know that people

in my department have had that conversation, but I personally have never had that conversation. And I will at times on a quarter grade . . . [when] 50% of my kids are failing . . . I'm pushing them trying to get them to do things."

Grading Rationale

Two themes were drawn from the qualitative data to explain grading rationale. Table 3 provides examples of significant teacher statements and formulated meanings.

Theme 1: Balance between grading rigor and student promotion. Teachers described the importance of creating student grades largely determined by assessments and accurate in displaying student learning, but they also explained their obligation and desire to help students be successful in class and school. When explaining this balance, teachers often described the difficulty of accomplishing both objectives. An English teacher gave the example of considering students with low grades: "I see them working all of the time, so I don't know—do I give them the F? Do I give them another opportunity? For me, that's the problem." Another teacher expressed a similar concern about grading senior students by using the term "flexible"—a term often used by participants to describe grading that is heavily reliant on nonachievement factors. The teacher questioned, "How flexible do I need to be to make sure that we hit a reasonable graduation rate? How accountable do I hold the kid?"

Table 3.

Examples of Significant Statements and Formulated Meanings for the Emergent Themes Addressing Research Question 2

Significant Statement	Formulated Meaning
Theme 1: Balance between grading rigor and student promotion	
<ul style="list-style-type: none"> • "They have to earn their F." • "Am I gonna say you don't even know how to do tangent, sin, cosine, you're in physics—I'm gonna give you a C because you are a senior? Or am I gonna say when you walk out of physics you gotta know something?" 	<ul style="list-style-type: none"> • Although a student may get an F, it is only if a student puts forth little-to-no effort • It is challenging to make grading decisions for students with little background knowledge and few pre-requisite skills
Theme 2: Consideration of class level	
<ul style="list-style-type: none"> • "If you're teaching a GATE or an AP class, you have a much higher yoke on your back to pass those kids, versus a Gen kid can get a D and be OK." • "College prep kids are a little bit different. There is greater flexibility in how you grade them." 	<ul style="list-style-type: none"> • There are higher expectations and pressures to pass higher-ability students in advanced classes as compared to lower-level classes • There is less pressure grading CP students—they can be graded with more nonachievement factors.

Theme 2: Consideration of class level. Teachers described a strong consideration of class level when creating grades. Class level refers to curriculum rigor, and in the district of study it consisted of four levels (listed from lowest to highest rigor): general, college preparatory (CP), Gifted and Talented Education (GATE)/honors, and Advanced Placement (AP). Teachers explained their belief that advanced classes, specifically GATE/honors and AP, should utilize rigorous grading practices and produce report card grades derived mostly from achievement factors with little-to-no inclusion of nonachievement factors. They reasoned that doing so set high expectations commensurate with high-ability students and prepared students for the grading practices utilized by college professors. Conversely, teachers explained that general and CP classes need more “flexibility” and greater use of nonachievement factors to create report card grades. These beliefs applied to the grading of individual assignments and assessments and in creating report card grades. One teacher explained, “The expectations in the honors class—they are graded differently on individual assignments [from CP students]. They are graded much more rigorously.” Similarly a math teacher described how she weighs report card grades differently by class level, as her test-to-practice grading ratio (i.e., the ratio of the grade devoted to test grades compared to that devoted to practice grades such as homework, classwork, etc.) is higher in advanced classes and lower in lower-level classes. The teacher explained,

In a foundation [class], it might be 50-50: tests [to] practice. And then algebra, geometry: 60-40, advanced algebra: 70-30, math analysis: 80-20. So as they move up in difficulty, we put more emphasis on the assessment component because, again, we’re preparing them for college, where a lot of times it is only assessment.

Grading Practices

Of the 17 different grading practices presented to teachers in the survey, two were recommended by grading experts (Cross & Frary, 1999; McMillan, 2001), *student academic achievement* and *specific learning targets mastered by students*. These two practices were rated highest by teachers in this study ($M = 3.97$ and $M = 3.85$, respectively). The next two highest-scoring practices were both nonachievement factors: *student ability level* ($M = 3.58$) and *student effort* ($M = 3.06$). Teachers reported using all four of these practices between *largely* and *somewhat* when creating student report card grades. Table 4 displays the complete survey ratings.

Table 4.
Survey Results of Teacher use of 17 Grading Practices

Grading Practice	Percentages					M	SD
	(1) Not at all	(2) Slightly	(3) Somewhat	(4) Largely	(5) Completely		
Student ability level	4	7	25	57	8	3.58	.87
Student academic achievement	1	3	14	63	19	3.97	.73
Student disruptive behavior/conduct	75	19	4	2	0	1.33	.66
Student effort	6	25	31	32	6	3.06	1.01
Student participation and/or paying attention	23	31	27	17	2	2.46	1.09
Student improvement of performance	13	25	41	20	1	2.71	.96
Grade distributions of other teachers	91	5	2	2	1	1.16	.57
Student performance of other students in class	66	22	7	4	1	1.51	.84
Student performance compared to students from previous years	83	9	5	3	1	1.28	.72
Specific learning objectives mastered by students	2	4	17	62	15	3.85	.79
Formal or informal school or district policy or grading distributions	77	7	8	5	4	1.51	1.06
Student effort, improvement, behavior, and/or other nontest indicators for borderline grades	18	37	32	12	2	2.44	.97
Student completion of homework	37	36	23	4	1	1.95	.90
Quality of student completed homework	18	25	35	20	2	2.64	1.06
Inclusion of zeros from incomplete assignments or assessments	11	16	36	18	18	3.17	1.23
Student extra credit for academic performance	46	47	6	1	0	1.61	.63
Student extra credit for nonacademic performance	88	10	1	1	0	1.15	.45

The analysis of qualitative data about grading practices produced two themes to explain grading practices. Those themes are explained below, and Table 5 provides additional examples of significant teacher statements and formulated meanings.

Theme 1: Importance of ability and effort. Although teachers explained that assessment is an essential part of student grades, nonachievement factors were also described as important, especially ability and effort. Teachers explained that considering student ability in grading means more stringent practices for higher-ability students and more lenient practices for lower-ability students. Teachers described stringent practices as those most reliant on achievement factors and lenient practices as those most reliant on nonachievement factors to create report card grades. One participant expressed his reasoning for using ability in grading: “Because I want my students to be successful in life, I show them an accurate—rather than inflated—reflection of their abilities.” To this teacher, it was important to avoid grade inflation, presumably by avoiding the use of nonachievement factors in grading; however, the teacher was unaware that ability is, in fact, a nonachievement factor.

Alternatively, many teachers recognized effort as a subjective, yet necessary practice to motivate students and to reward their work. For these teachers, use of effort varied widely, as some described the factor as a major component of their report card grades, others described limited use, and a minority completely avoided the use of effort because they felt it led to inflated grades. An English teacher explained her inclusion of effort in a grade to reward student work completion: “Everybody’s working, nobody just slacks off and does nothing . . . I personally want to make sure I honor that and respect that as part of my grade.”

Theme 2: Dichotomy of grading: Assessment and work categories. Weighted grading, a method popularized with the advent of computerized gradebooks, allows teachers to assign each articulated grading category a percentage of the total report card grade (e.g., tests: 50%, homework: 30%, quizzes: 20%). Participants described two types of categories that were most important: assessment and work. While some teachers exclusively used these two categories to create grades, others used three or four categories. One participant explained her particular use of weighted categories: “My grading system is based solely on whether or not the students complete their homework and how they perform on their exams.” This use of two grading categories was the simplest type of weighted grading that was described and less common than the use of three or more categories.

Assessment categories were almost always the single greatest weighted category, but designated weights widely ranged from 30% to 90% of the total grade. Teachers of advanced classes tended to place more weight on assessment categories than lower-level classes. A math teacher explained this phenomenon: “In algebra, it’s 60% assessment, 40% other—which would

be everything else, and advanced algebra/GATE, it's 70% assessment, 30% other." When creating work categories, some teachers made this a single category comprised of everything that was not an assessment, while most teachers created multiple nonassessment categories (e.g., classwork, homework, and quizzes). A teacher described his use of assessment and work categories: "15 to 20% range for homework, 60 to 70% on the testing, and then I have daily warm-ups, quizzes—stuff like that—maybe in the 5 to 10% range."

Table 5.
Examples of Significant Statements and Formulated Meanings for the Emergent Themes Addressing Research Question 3

Significant Statement	Formulated Meaning
Theme 1: Importance of ability and effort	
<ul style="list-style-type: none"> • "If they bust it and try, they won't fail." • "Sometimes you have students that make the effort, but are unable to grasp the concept. This student should not be penalized." 	<ul style="list-style-type: none"> • If students put in effort, they will earn at least a D. • If they put in effort, student grades shouldn't be reduced because they don't understand a topic.
Theme 2: Dichotomy of grading: Assessment and work categories	
<ul style="list-style-type: none"> • "For 10th, 11th, and 12th grade, we've all in our—at PLC levels—have decided to do 60% summative assessments . . . 30% for classwork, practice, anything else. And 10% is reading." • "For me, it's 30, 35, 35[%]...they have labs, they have assignments, and they have tests, so it's pretty evenly spread." 	<ul style="list-style-type: none"> • Teachers use common weighted grading across classes: 60% tests, 30% work, 10% reading. • Weighted grading consists of 30% for labs, 35% for work, and 35% for tests.

Discussion

Grading Influences

Three influences, philosophy of teaching and learning, concern about external perceptions, and administrator pressure on assigning low grades, were found to strongly affect teacher grades. The importance of philosophy of teaching and learning, which McMillan (2003) classified as a component of teacher knowledge, beliefs, expectations, and values, is consistent with the findings of Frary et al. (1993) and Zoeckler (2007), who added that the influence leads to subjective practices that differ grading expert recommendations. However, despite being the second-strongest influence measured by the survey, teachers rarely used the term *philosophy* in focus groups. Instead, teachers discussed four different influences that together were interpreted as key components of their philosophy of teaching and grading: desire to promote student understanding, student motivation and engagement, desire for student success, and desire to accommodate student differences.

Desire to promote student understanding was the only strong influence leading to objective grading practices. Although McMillan (2003) found this influence to be an emergent theme explaining teachers' use of assessment to determine student learning, in this study the theme displays teachers' intent to use grades to communicate student achievement, which is measurement experts' single recommended use of grades (e.g., Brookhart, 2009; Guskey, 2015; Sadler, 2009). Teachers also indicated student motivation and engagement had a strong influence on their grading, as they largely believed grades to be an important form of extrinsic motivation needed for students to complete assignments. This finding, consistent with previous studies (McMillan, 2003; Stiggins et al., 1989), displays teachers' use of grading to address student behaviors such as classwork and homework completion, both subjective practices contributing to grading inaccuracy.

Teachers' desire for student success was another strong influence, similar to findings by Cizek et al. (1995) and McMillan (2001) that largely revealed teachers' willingness to use nonachievement factors to raise a student's report card grade if the teacher determines that the student is somehow deserving of this additional support. Brookhart (1993) and Cheng and Sun (2015) explained that teachers espousing this belief often consider the future consequences of student grades, and Brookhart further expounded that these teachers play the *advocate role* when grading instead of the *judge role*. However, it is the latter role, Brookhart explained, that is required for objective practices. Finally, teachers' desire to accommodate student differences was also found to be an important influence on grading, supporting findings of previous studies that teachers consider morality, fairness, and equity an important part of grading (Manke & Loyd, 1991; Zoekler, 2007). The current study found that teachers actively sought to identify issues of inequity in the classroom and used grades as a tool to address them. Although done with the best of intentions, practices resulting from this influence are highly subjective, leading to further grade inflation for students with the greatest needs.

Teachers' concern about external perceptions was an unexpected influence, as it was absent from previous grading literature. Teachers explained they recognize and care that administrators and other teachers judge their professional efficacy based upon the grades they assign, knowing that high grades (A's, B's, and to a lesser extent C's) generally reflect highly on teachers and low grades (D's and F's) often reflect poorly. At the same time, teachers expressed the desire to be seen as an effective grader that uses valid practices. Implications of this influence likely depend on the school culture of teaching and grading. If schools narrowly emphasize high student grades, then this influence would likely contribute to subjective and inflated grades. However, if schools instead emphasize accurate measurement of student learning, then the influence would likely contribute to objective and accurate grades.

Finally, administrator pressure on assigning low grades was an important external influence on grading, as participants often described administrators as a small, yet constant consideration when determining report card grades. This finding is consistent with the suggestions of Rauschenberg (2014), who claimed that administrator pressure might explain his finding of differential grading across grade levels in North Carolina high schools. Teachers' concerns for this pressure, no matter how small, can lead to higher student grades. Although administrators likely have good intentions in addressing concerns about low student grades, the influence is an additional contribution to the subjectivity of report card grades and grade inflation.

Grading Rationale

The rationale that teachers use when creating report card grades is notoriously difficult to determine, as described by McMillan (2003); however, this study produced two themes to address this domain of the theoretical framework. First, teachers often consider the balance between grading rigor and student promotion, one that requires ensuring rigorous grading practices *and* helping students to succeed in their class. Bishop (1992) and Brookhart (1993) described this phenomenon and explained that although teachers often attempt to play the role of judge and advocate in the grading process, they cannot do both effectively. The current study's findings that teachers attempt such an untenable balance likely leads to cognitive dissonance within the grader. This can result in inconsistent grading practices, as alternating the roles of judge and advocate also means alternating objective and subjective grading practices. This phenomenon is similar to the tension described by McMillan (2003) when teachers' beliefs and values clash with external factors in grading decision-making.

Second, teachers' consideration of class level was found to be an important component of grading rationale. This differential objectivity likely results in highly accurate grades in advanced classes and highly inaccurate grades in lower-level classes. Similarly, McMillan (2001) found class ability level to be positively related to teacher use of achievement factors and negatively related to nonachievement factors. In the current study, teachers described "flexibility" in grading lower-level classes, a concept that seemed to allow teachers to establish desired grade distributions or assign students deserving grades by maximizing the use of nonachievement factors. Alternatively, teachers' reasoning for emphasizing objective practices in advanced classes seems to result from a confounding of grading rigor with curriculum rigor. Overall, the subjectivity of lower-level class grades results in students, parents, and other interested parties with a misunderstanding of student achievement, while the same stakeholders can gain a clear understanding of student achievement in advanced-level classes. This situation leads to implications of inequity of grading practices across class levels.

Grading Practices

Like Randall and Engelhard (2010) and McMillan (2001), both the quantitative and qualitative results of this study show that student academic achievement is the largest factor that teachers use when creating report card grades. Additionally, teachers indicated they largely consider specific learning targets mastered by students when grading, further emphasizing the importance of academic achievement in grading. However, multiple influences and elements of rationale also cause nonachievement factors to play a key role in teacher grading practices.

Like previous studies finding that teachers use nonachievement factors to create student grades (Cheng & Sun, 2015; Cizek et al., 1995; Cross & Frary, 1999; McMillan, 2001), this study found teachers' grading practices to consist of student effort and ability level in addition to achievement. Teachers described considering student effort to ensure low-achieving and low-ability students would not fail. Additionally, effort was factored into grades when teachers graded classwork and homework for completion, although teachers seemed largely unaware of the role of effort in this practice. Ability was found to affect student grades differently within the same classroom, depending on the student. When teachers considered ability, grades of low-achieving, low-ability students became more subjective and less accurate, as teachers looked to increase grades with other nonachievement factors, such as effort. Grades of high-achieving, high-ability students became more objective and accurate, as teachers sought to maximize the weight of assessments to increase rigor, something discussed in grading rationale. Interestingly, ability was the only factor that teachers failed to recognize as a subjective grading practice. Because of a strong teacher belief that consideration of student ability level is important to ensure equity and fairness in the classroom, teachers largely failed to question the objectivity of the practice. In fact, even the most objective graders often included ability level as a factor in grading.

Finally, teachers' use of weighted categories, mainly assessment and work, helped to explain their use of gradebooks to create report card grades. Although the category weight differed across classrooms, teachers displayed their intentionality in using both objective and subjective grading practices, as the assessment category predominantly consists of objective practices and the work category consists of subjective practices. These findings are similar to those of McMillan (2003), who found that teachers use practices such as tests, quizzes, homework, and participation to create grades, and Zoeckler (2007), who found widespread use of teacher weighted grades.

Limitations

Several limitations existed in this study, and a few elements in particular limited the generalizability of results. First is the use of convenience sampling

to select the district of study, although the district demographics do parallel those of California. Purposive sampling to select focus groups may also limit generalizability, and although there are many benefits of using a single school district for a study setting, this may also limit the generalizability of results. Other limitations include a survey response rate that was lower than desired (despite numerous actions to maximize responses). Also, the absence of first-hand observations of teacher decision-making inherently reduces the potential for description of the process, as social desirability may have played a role in teachers' survey and focus group responses.

Conclusions

The teacher grading decision-making process is complex, but successful grading reform efforts will likely benefit from an understanding of the grading influences, rationale, and practices that teachers employ in the grading process. This study found teacher grading to be strongly influenced by teachers' philosophy of teaching and learning, their concern for external perceptions, and administrator pressure on assigning low grades. Teachers' philosophy of teaching and learning, found to consist of four smaller components, was interpreted as a subjective influence that adds to grading inaccuracy. The impact of teachers' concern for external perceptions on grading is unclear, and this influence deserves further research to understand its implications. Administrator pressure on assigning low grades is seen as an influence leading to grade inflation, but additional research may help to understand the extent of its impact and occurrence.

Teachers' rationale for their grading practices was found to include a conscious balance between maintaining grading rigor and ensuring student promotion, which likely adds to grading inconsistency within classrooms, and a consideration of class level, which results in greater grading subjectivity in lower-level classes. Like so many earlier studies, teachers considered student academic achievement to be the most important grading factor, but they also commonly utilized effort and ability nonachievement factors. Finally, teachers explained they consider two major categories in their gradebooks: assessment and work.

References

- Allen, J. (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 78(5), 218-223.
- Bishop, J. H. (1992). Why U.S. students need incentives to learn. *Educational Leadership*, 49(6), 15-18.
- Brookhart, S. M. (1993). Teachers' grading practices: Meanings and values. *Journal of Educational Measurement*, 30(2), 123-142.
- Brookhart, S. M. (2009). *Grading* (2nd ed.). Upper Saddle River, NJ: Pearson.

- Cauley, K., & McMillan, J. H. (2000). *Do teachers grade students differently in high poverty middle schools?* Paper presented at the American Educational Research Association, New Orleans, LA.
- Cheng, L., & Sun, Y. (2015). Teachers' grading decision making: Multiple influencing factors and methods. *Language Assessment Quarterly*, 12(2), 213-233.
- Cizek, G. J., Fitzgerald, S. M., & Rachor, R. E. (1995). Teachers' assessment practices: Preparation, isolation, and the kitchen sink. *Educational Assessment*, 3(2), 159-179.
- Corbin, J., & Strauss, A. (2008). *Basic of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Los Angeles, CA: Sage.
- Creswell, J. W., & Plano Clark, V. W. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.
- Cross, L. H., & Frary, R. B. (1999). Hodgepodge grading: Endorsed by students and teachers alike. *Applied Measurement in Education*, 12(1), 53-72.
- Frary, R. B., Cross, L. H., & Weber, L. J. (1993). Testing and grading practices and opinions of secondary teachers of academic subjects: Implications for instructional measurement. *Educational Measurement: Issues and Practice*, 12(3), 23-30.
- Guskey, T. R. (2015). *On your mark: Challenging the conventions of grading and reporting*. Bloomington, IN: Solution Tree Press.
- Manke, M. P., & Loyd, B. H. (1991, April). *A study of teachers' understanding of their grading practices*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.
- Marzano, R. J. (2000). *Transforming classroom grading*. Alexandria, VA: ASCD.
- McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice*, 20(1), 20-32.
- McMillan, J. H. (2003). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice. *Educational Measurement: Issues and Practice*, 22(4), 34-43.
- McMillan, J. H., & Lawson, S. R. (2001). *Secondary science teachers' classroom assessment and grading practices*. Metropolitan Education Research Consortium, Richmond, VA. (ERIC Document Reproduction Service N. Ed 450 158)
- McMillan, J. H., & Nash, S. (2000, April). *Teacher classroom assessment and grading practice decision making*. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA.
- Randall, J., & Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education*, 26, 1372-1380.
- Rauschenberg, S. (2014). How consistent are course grades? An examination of differential grading. *Education Policy Analysis Archives*, 22(92), 1-38.
- Reeves, D. B. (2011). *Elements of grading: A guide to effective practice*. Bloomington, IN: Solution Tree Press.
- Sadler, D. R. (2009). Grade integrity and the representation of academic achievement. *Studies in Higher Education*, 35(7), 807-826.
- Saldana, J. (2012). *The coding manual for qualitative researchers*. Los Angeles, CA: Sage.
- Stiggins, R. J., Frisbie, D. A., & Griswold, P. A. (1989). Inside high school grading practices: Building a research agenda. *Educational Measurement: Issues and Practice*, 8(2), 5-14.
- Wilen, W., Hutchinson, J., & Ishler, M. (2008). *Dynamics of effective secondary teaching* (6th ed.). Boston, MA: Pearson.
- Zoeckler, L. (2007). Moral aspects of grading: A study of high school English teachers' perceptions. *American Secondary Education*, 35(2), 83-102.

APPENDIX A: FOCUS GROUP INTERVIEW PROTOCOL

Date:	Time:
Length of Activity:	Site:
Participants (gender & subject taught):	

Introductory Statement

Good afternoon, and thank you for coming. My name is Josh Kunnath. I teach at Highland High School, and I'm also a doctoral candidate at Fresno State. Does anyone mind if I record the meeting for data collection purposes? This focus group is part of the data collection methods of grading study. You may remember a survey on grading practices sent out a couple of months ago. That was also part of my study—it was the first phase. This is the second and final phase, and two other schools were also selected to participate in focus groups.

The purpose of this focus group is to better understand the decision-making process that teachers use to create student report card grades. Like the survey, the focus is on regular education teachers of at least one class of English, foreign languages, mathematics, science, and social studies. Like the survey, your participation is voluntary. Your participation in this focus group will represent your informed consent.

I have a group of questions that I'd like to ask all of you, and I'd like to encourage you to respond openly and honestly. All responses are completely anonymous and confidential, and all data collected will only be used for the purposes of this independent study or subsequent independent studies. I will be recording this session so that I can create a transcript of the session. I will also be taking notes during the session.

Your participation is greatly appreciated, and the successful completion of this study is attributed in part to the time and input that you have provided. For this I thank you.

Focus Group Questions

- 1) What are some challenges you face in creating student report card grades?
- 2) How do you create student report card grades?
- 3) How do your own knowledge, beliefs, and/or values (personal or professional) influence the decisions you make in creating report card grades?
- 4) How do external factors (i.e., state testing, district policy, parents, administrators) influence the decisions you make in creating report card grades?
- 5) How do classroom realities (i.e., social promotion, student absenteeism, disruptive behavior, and differing student ability within classes) influence the decisions you make in creating report card grades?

Reproduced with permission of
copyright owner. Further
reproduction prohibited without
permission.