

# Standards-Based Grading: Theory and Practice

*Reach Institute for Instructional Leadership*  
*Oakland, California*

Submitted by Kate Sugarman  
July, 2015

## **Abstract**

ARISE High School is a small public charter school in Oakland's Fruitvale neighborhood. In 2013, the school leaders (including myself) made the decision to transition from a holistic grading system to a standards-based grading system. In the subsequent years, the teachers and leaders worked to develop a codified set of learning targets that defined what students should know and be able to do in each course. During this time, it became apparent that teachers had different levels of understanding of the theory of standards-based grading, and had different levels of technical competence in implementing this system in their classroom. The literature indicates that standards-based grading is more beneficial to student learning than traditional grading systems, and that it provides more insight into student learning and gives feedback to teachers about their instruction. In order to address the diversity of teacher understanding and expertise in implementing standards-based grading, I designed an intervention that focused on three group workshops followed by individual coaching sessions. The goal of this intervention was to improve the ability of the Social Science department to use standards-based grading in their classrooms by deepening their conceptual understanding and by improving their technical skills. In order to evaluate the effectiveness of the intervention, I examined scripted notes from the workshops and coaching sessions, written feedback from the participants, survey results, and the artifacts created by the department. The data revealed that such an intervention is effective in shifting teacher thinking and improving teacher skills given the right amount of time and deliberate facilitation. Teachers reported a level of technical competence, and they created artifacts that reflected this thinking. They responded well to the deliberate facilitation, but observed how time-consuming the work was, and how little capacity they had to focus on this work during the allotted time. Although the intervention did increase teacher investment through the collaborative meaning-making process,

the intervention was only effective at beginning the process, and did not ultimately go far enough in order to truly shift teacher practice.

### **Context and Problem of Practice**

The transition to standards-based grading has allowed ARISE High School many opportunities to examine the expectations that teachers, students, and families have for the type and quality of student work in each discipline. Standards-based grading is a system where students are evaluated on their mastery of each standard, rather than on a holistic or points-based system. According to Marzano (1996), standards are needed because traditional forms of grading are inconsistent and subjective. O'Connor (2009) states that standards based grading allows for communication between students and teachers based on language, rather than numbers, which opens up the conversation to be one centered on learning rather than on compliance or completion. Scriffiny (2008) argues that a system based on points can obscure student learning rather than illuminate it. ARISE teachers and leaders wanted to have grades that accurately reflected student learning, and allowed teachers to know exactly where their students were succeeding and struggling. The goal of the standards-based grading system is to prioritize learning over compliance. "Standards-based grading prioritizes the final result of learning instead of a summation of grades earned during the learning process" (Iamarino, 2014).

Because standards-based grading is a relatively new innovation, none of the teachers at ARISE experienced such a system when they themselves were students in school. Because teacher expectations are heavily shaped by their own experiences as students, it has been difficult for some teachers to understand and implement standards-based grading. In addition, implementing standards-based grading takes a considerable amount of technical expertise. Marzano (1996) states that devising appropriate standards is a technically demanding process that must be differentiated to each local context. Through a process of breaking down Common

Core and Next Generation Science Standards, many departments have developed a set of learning targets that all current teachers understand and value. However, most departments have yet to create performance assessments that are codified, and that are aligned to the learning targets. This means that teachers are developing their own means of assessing the learning targets without sharing a common vision of what mastery of the learning target looks like. “In order to successfully operate and maintain a standards-based grading system, there must exist an exceptionally clear model of criteria by which to evaluate students’ academic advancements within that system. Attention to criteria, after all, is what separates standards-based grading from other forms of grading; criteria are what help teachers determine whether or not standards have been met” (Iamarino, 2014). By having clear criteria for success, teachers and students build shared understanding of what it actually means to master a learning target, and have a clear roadmap for how to improve when the standard is not yet met. At ARISE, this shared understanding does not yet exist.

Although teachers believe in the standards-based grading system in theory, implementing it “requires sophistication and commitment of teachers above and beyond the expectations of many schools” (Iamarino, 2014). As a result, although the expected learning outcomes are articulated, ARISE students, parents and teachers are not clear on what type and quality of work is expected in order to show mastery of these outcomes. ARISE Teachers are inconsistent in the number and type of assignments they give, the way they assess those assignments against the rubrics, and the level of critical thinking their performance tasks are designed to measure. When interviewed, teachers agreed that the idea of standard based grading makes sense, and that student performance as measured through this system generally correlates to their internal sense of where students are at with the course. Teachers also agreed that the system was inconsistently implemented, due to lack of familiarity with how to use it, and with inadequate coaching for individual teachers just learning the system. Teachers also felt that the system operated differently for content-based standards versus skill-based standards, since the skill-based standards are recursive and offer more opportunities for mastery than the

content-based standards do. Students and families are unsure of the exact expectations for their performance, and as such are not fully aware of how they are graded, or where they need to improve. When interviewed, students asserted that they liked the system because it told them what they were expected to learn and do. They appreciated that the system gave them multiple opportunities for improvement. However, they also felt that the system was inconsistently implemented from subject to subject and teacher to teacher. They also felt that they did not always understand why they were given the evaluations they were given, because criteria for success was not always clear. Although no parents at ARISE were interviewed, research shows that parents are often divided on the value of standards-based grading. In one study, parents overwhelmingly preferred the standards-based report card to the traditional report card (Swan, 2014). However, in a different study, parents resisted the reform because they felt students would lose motivation in a system that provides multiple opportunities for success and does not track compliance for things like homework (Spencer, 2012). **My problem of practice is that a lack of codified assessments with corresponding criteria for success and a lack of understanding of the theory of standards based grading leads teachers to implement standards based grading inconsistently, which does not lead to improved learning outcomes for all students.** This action research project will investigate how surfacing beliefs about the purpose of grades, collaborating to develop assessments and criteria for success, and individualized feedback and coaching can move teacher practice to implement standards-based grading more consistently.

## **Review of Relevant Literature**

### ***Introduction to the Literature***

The movement for standards-based grading has been going on for the past 20 years, and is gathering steam with the emergence of Common Core standards, and the need to determine student proficiency in higher-order thinking skills. ARISE has shifted to standards-based grading from a more holistic grading practice, which is different from the shift from more traditional grading practices. This shift was made without enough research into rationale, best practices, and effective teacher learning. This literature review attempts to backfill some of that missing knowledge by uncovering the core principles of standards-based grading, exploring the obstacles to implementation, and understanding effective practices for teacher learning in this area.

### ***History of Grading***

Although assessment of students has always been part of the Western education tradition, it is only in the past 150 years that summative grades have come into common use (Guskey, 1994). In the early days of American education, students were in multi-age one-room schoolhouses, and made progress individually. Students rarely advanced beyond elementary education, and those that did often pursued their education independently, or with private tutors (Guskey, 1994, Kirschenbaum, 1971). In the early 1900s when education became compulsory, and high school attendance dramatically increased, a practice of creating grades based on percentages of assignments emerged (Guskey, 1994). It is unclear exactly why this practice developed, but it likely was a response to the need to find an efficient way to handle the high volume of students and work that high school teachers were faced with (Guskey, 1994).

As early as 1912, researchers published a study showing the arbitrariness of teacher grading practices. The study revealed that teachers scored the same assignment completely differently depending on what the teacher valued, and that

often neatness and handwriting mattered more than the content of the assignment (Starch, D., and E. C. Elliott, 1912). However, this research had little impact on grading practices. In 1918, broader grading categories such as Excellent, Fair, and Poor, and A,B,C,D,F were developed (Johnson, 1918, Rugg, 1918).

In the 1930's grading on a curve became more commonplace in an attempt to separate students by some sort of performance indicators (Guskey, 1994). It is at this point that grading became a tool to "sort" students in relation to one another. At this time, there was a belief that since intelligence was believed to be distributed along a curve, then it made sense that academic achievement should correspond (Guskey, 1994). The use of grades as a means to determine which students had "aptitude" for academics led to the development of tracking, vocational programs, gifted programs, and other sorting mechanisms (Guskey, 1994). Not surprisingly, these sorting processes often intersected with race and class in a way that privileged the wealthy white students. Interestingly, as this practice was becoming common, there was also a movement in the opposite direction. Some schools abolished grades, others reverted to individual narratives, others used a system of "of either pass or fail, and still others used a "mastery" system to determine when students could move to the next level (Guskey, 1994). From the beginning of the grading system that is currently most prevalent, there have been critics such as Guskey and O'Connor who see the system as too arbitrary, too subjective, and as having a negative affect on students.

Currently, the majority of schools use letter or number grades in what is termed a "hodge-podge" fashion. Grades are derived from non-standard elements such as percentages of homework, academic achievement, participation, behavior, and effort. This "hodge-podge" method results in a lack of clarity of what grades truly mean, and provide no clear measure of student learning (Bowers, 2010).

In addition to the system of letter grades or percentages being arbitrary in itself, several scholars have revealed that there are many other factors that create grading bias in teachers. (Rauschenberg, 2014). Gender, ethnicity, effort, and personal relationship all play factors when teachers assign grades. Because teachers have wide latitude in assigning grades, these biases can have significant

impacts on various groups of students. Teachers need to become aware of these biases in order to learn to recognize patterns in their own grading behavior (Conroy, 2001).

In the past 20 years, more people have been advocating for a grading system that is more reflective of student learning, connected to concrete standards, and less subjective (Marzano, 1996). With the development of state standards, and now the Common Core, there is a growing understanding that if these standards are to have value, they must be the criteria against which classroom achievement is measured.

### ***Impacts of Grading on Students***

Beginning in the 1950's, researchers began to study the impact of grades as motivational tools. It soon emerged that letter or number grades absent any concrete feedback did very little to motivate students (Page, 1958). Further studies have shown that although some high achieving students are motivated to continue receiving high grades and to avoid poor grades, low achieving students are not at all motivated by poor grades, and actually become less motivated by the constant experience of failure. Students considered "at-risk" such as students with learning disabilities were seen to be especially vulnerable to this kind of negative motivation (Selby and Murphy, 1992).

Corresponding to the decreased motivation created by low grades is a lowered graduation rate. Teacher-assigned grades in 7<sup>th</sup> and 8<sup>th</sup> grade were seen to be the strongest predictor of high school graduation rates (Bowers, 2010). Although it is recognized that teacher-assigned grades are not accurate measures of academic achievement, due to the "hodge-podge" nature of the grades, Bowers concludes that the reason grades are predictive of graduation rates is because they correlate to students' ability to navigate the institution of school. Many researchers have argued that the institution of school is often unfriendly or hostile to non-white students of low economic status (Delpit, 2006), and that the majority white, middle-class, female teaching workforce is often ill-equipped to teach low-income students of color. As race, class and ethnicity have already been shown to have negative impacts on student grades due to teacher bias, the significance of lowered grades on



low-income populations of color cannot be overstated. The graduation rate gap between low-income students of color and middle class white students is significant in Oakland, with white students graduating at a rate of 74.8% compared to 56.1% of Latinos and 54.2% of African Americans in 2009 (Oakland Youth Indicator Report, 2011). Although the gap in graduation rates is attributable to many factors, one clear contributor is a lack of passing grades for African American and Latino students. Lowered grades, therefore, have a very strong impact on graduation rates in Oakland.

### ***Arguments for Standards-Based Grading***

The traditional system of grading is ineffective in serving as an accurate measure of student learning, which is arguably the most important function of grades and evaluation in general. Standards-based grading has emerged as an alternative to traditional grading, and a means to communicate a more authentic picture of student achievement. There are many inconsistencies in curriculum and grading, and a standards-based approach attempts to address both of those issues (Marzano, 1996).

The movement to develop state and Common Core standards is driven by the idea that there should be shared expectations for student learning. Marzano (1996) further argues that each school needs to develop its own set of standards derived from the state in order to have local control and understanding of what the standards mean. Once the standards are clear, it is essential to develop clear criteria of what mastering those standards means. Having clear criteria that is understood by teachers, students, and parents, is the rock upon which standards-based grading is built (Iamarino, 2014). This criteria is what distinguishes standards-based grading from more arbitrary grading practices.

In addition to clear criteria for success, teachers must have curriculum and assessment that is aligned to these criteria (Spencer, 2012). Once these things are in place, then teachers can provide specific feedback to students about their relationship between their work level and the desired outcome. This is the space in which the communication between teacher and student becomes rich and develops

greater learning. Having this feedback drive the assessment process transforms the communication of grades to being one focused on language rather than numbers, which opens communication rather than shut it down (O'Connor, 2009). This feedback process changes the nature of the conversation between students and teachers to being about quality or work rather than compliance (Spencer, 2012). In other words, the focus shifts to learning rather than doing.

In his book "A Repair Kit for Grading," Ken O'Connor (2009) explains that traditional grading practices are "broken" in that they are either arbitrary or possibly unfair, and that they signify something other than actual learning. He offers standards-based grading as a solution to many of these problems, because this type of assessment is based solely on student attainment of learning objectives in relation to an agreed-upon criteria and centers on communication between teacher and student. He posits that "the primary purpose [of grades] to be communication about achievement," and that grades that do not do that are "broken." He suggests creating alternate types of feedback mechanisms to communicate information about things such as work completion, participation, and effort in order to preserve the single purpose of standards-based grades.

### ***Teacher Perceptions of Purpose of Grading***

One of the key factors in successfully implementing standards-based grading is teacher perceptions and beliefs about grading. Fairness is held as a core tenet of education, in that all students must be measured in the same way at the same time. However, since we know that not all students are the same, the basic idea of uniformity being equivalent to fairness is inherently inequitable (O'Connor, 2009). Yet deep-rooted ideas about fairness and equality persist among teachers that are in conflict with this understanding of equity (O'Connor, 2009). Teacher preparation programs rarely every focus on best practices in grading and assessment, and so teachers often enter the classroom with ideas about grading that reflect their own experiences as students (Guskey, 2009). Teachers commonly say that it is important to them that grades be "fair," but ideas about what make grades "fair" vary widely (Brookhart, 2011). Interestingly, teachers often assume that their ideas

about grading practices are commonly held, even when they are not (Brookhart, 2011). In a study conducted by Tierney, Simon, and Charland (2011), teachers turned out to have limited understanding of the principles underlying their grading system. They were uncertain about the primary purpose of grades, what they were supposed to communicate, and how they were supposed to be determined. Guskey (2009) further explains that teachers have different ideas about whether to privilege product, process, or product in their grades. Many teachers try to do all three at once, contributing to the “hodge-podge” effect. Guskey argues that schools must determine which of these three elements the grading process should reflect, and then design alternate systems to measure the others. Brookhart says that any school interested in implementing standards-based grading must first have a conversation about what meaning grades should convey.

Without an understanding of meaning, teachers, students, and parents will be caught up in the details of the system and miss the important point (Brookhart, 2011). Many teachers, students, and parents believe that grades should serve as extrinsic motivators. Some of the beliefs teachers hold about grades are that grades should reward effort and punish misbehavior, that grades should serve as an incentive or punishment for students to do work, and that grades should mirror the ‘real world consequences’ for failure to meet deadlines or comply with rules (Tierney et al, 2011). However, much research indicates that relying on grades as extrinsic motivators will only work with a small segment of students, and can often have a damaging impact on the students who struggle most (O’Connor, 2009). Additionally, using grades as an extrinsic reward is unlikely to actually shift patterns of behavior. If the true goal of education is to develop life-long learners, then the primary reward should be intrinsic: the satisfaction of learning something well (O’Connor, 2009).

In addition to regarding grades as motivators, many teachers believe that grades should teach students responsibility and accountability. Secondary teachers specifically feel responsibility to prepare students for college or work, and this often leads to the consideration of non-academic factors when determining grades (Guskey, 2009). It is also possible that secondary teachers use grades more as a

means to control student behavior, since more traditional behavior management systems are often not present in high school (Guskey, 2004).

These beliefs are incompatible with the idea behind standards-based grading that grades should solely communicate academic achievement. Tierney et al (2011) explain, “despite the appeal of real-world philosophy, it can lead teachers to interpret assessment policies in such a way that their practices do not effectively support student learning.” This illustrates that when teachers are not completely clear on the philosophy of the grading practice, or if their own beliefs outweigh their belief in the standards-based philosophy, they will take actions that are counter to the goals of standards-based grading. Guskey (2011) asserts that efforts to implement standards-based grading systems that do not take in to consideration elements of personal responsibility that teachers value are unlikely to be successful unless there is an alternative way to report this information.

### ***Standards-Based Grading Implementation***

Most schools or districts do not start out as standards-based grading institutions. Most schools begin with traditional grading, and then begin a process of conversion. The available literature is mostly concerned with schools or districts that are converting from traditional grading to standards-based grading. Sometimes the change is driven by teacher interest, and sometimes by district decision (Brookhart, 2011). The results of these conversions illustrate the promise and pitfalls of such work. The most significant obstacle to grading reform is teacher’s beliefs about grading, and the status quo (Guskey 2009, Scriffny, 2008). In addition, some schools face opposition from parents, who either find standards-based report cards confusing, or who want to have non-academic factors such as effort and participation factor into their children’s grades (Spencer, 2012). The desires of students in these matters are rarely found in the literature, so it is difficult to know how students’ opinions contribute to the reform efforts.

In the cases of most reform efforts, change can be described as new ideas taking root in one pocket of the school and then slowly spreading (McMunn, 2003).

When teachers lead the change process, it is usually because one or a few teachers become dissatisfied with traditional grading practices and receive permission from the administration to experiment with standards-based grading. In one case where a teacher has positive results from switching to standards-based grades, the practice spread to a few like-minded teachers who formed groups to study the practice. From there, with the administration's support, that group of teachers began to encourage more teachers to adopt this practice. Once a critical number of teachers began using standards-based grading, a 'tipping point' was reached (Townesley, 2013). The discrepancy between the two types of grading became obvious, and the district needed to make a choice about what method to use. The district began a process of studying research on standards-based grading, sharing research with the teachers, and holding community meetings with stakeholders such as students and parents to discuss the pros and cons of converting to standards-based grading (Townesley, 2013). Finally, teachers were polled on their agreement with and readiness for implementation. Based on the results, a 2-year implementation plan was developed in order to eventually move all teachers to standards-based grading (Townesley, 2013). At the end of that period, a large number of students reported an improved understanding of their learning and a shift in their thinking about the meaning of grades. The entire transition process took place over the course of many years, but resulted in a significant level of buy-in from all stakeholders (Townesley, 2013).

In a case where change was initiated by the district, the conversion process was quite different (McMunn, 2003). The district developed core principles of grading, and then developed grading guidelines. A corresponding toolkit was developed and field-tested with teachers at different schools throughout the district for 3 years. A cohesive professional development plan was then designed to train teachers in the purpose and processes of standards-based grading. Teachers participated in the professional development, and although many of them rated the PD highly and found it helpful, the transition from strategies learned to classroom practice was slow and inconsistent. The district's approach was a combination of "pressure and support," but the pressure approach was not particularly effective in

leading to change in teacher practice (McMunn, 2003). Furthermore, the study concluded that it is not realistic for teachers to make the switch without considerable expertise and support surrounding them, and extensive time to make the transition (McMunn, 2003). The evaluators of this process determined that the following elements of professional development needed more research as to how districts “can” and “should”:

1. Offer year-long, job-embedded courses on classroom assessment,
2. Organize and support teacher "learning teams" who engage in self-study,
3. Provide quality assessment materials to teachers along with opportunities to talk with other teachers about student results on those assessments,
4. Mandate teacher participation in the initiatives involving assessment,
5. Offer individualized feedback to teachers on how they can improve their use of assessment,
6. Look more closely at the link between good assessment and the instructional strategies that teachers use to support the assessment,

The researchers additionally noted that #5, individualized coaching and support, was probably the most valuable and most likely to drive change in classroom practice, but is the hardest to implement, especially across a district (McMunn, 2003).

### ***Effective Intervention Methods***

My intervention will focus on understanding the purpose for assessment and clarifying the criteria of success for learning targets through core assessments. In order to do so, I will focus on improving the practices of the teachers in the Social Studies department by what the research indicates are appropriate teacher learning processes: 1. Engaging the teachers, in a process to build a shared understanding of the grading system; 2. Calibrating their collaborative grading of core assessments; and 3. Providing individual feedback through coaching.

### ***Building Shared Understanding***

Susan Brookhart (2011) argues that any school wishing to improve its grading practices must begin by building consensus around what grades should communicate. Building consensus comes from sharing a common language and

constructing shared meaning around a concept. According to John Kotter (1996), a shared vision is essential to any change process, and that vision must be clearly and repeatedly articulated. ARISE has an articulated vision of grading and assessment, but teachers have not been engaged enough in making meaning of that vision, debating the vision, or understanding the reasons behind the vision. Dufour (1998) argues that co-creating a vision, although inefficient, is far more likely to lead to teacher investment. Novice teachers, especially, build knowledge and skills best in collaborative groups, and those groups are most effective when the context is shared (Herrera, 2014). I hope that by examining the articulated vision, exploring research, and bringing clarity to issues that are unclear, the Social Studies department will develop a greater and more authentic understanding of the goals and rationale of standards-based grading.

### ***Teacher-Centered Collaboration***

In addition to understanding the theories behind standards-based grading (the “why”), teachers need to understand the criteria of success for each of the standards they teach; the “what” and “how” of what they assess (Marzano, 1996). Currently, teachers are developing their expectations in isolation, and sometimes do not have a clear idea of what they expect from the work until the students turn it in. Although rubrics for many assessments already exist, the teachers do not have a shared understanding of the expected quality of student work for each level of the rubric. Developing exemplars and criteria of success is a process that is best done collaboratively, as it will increase teacher investment and be more likely to have a long-lasting impact on teacher practice (Choi, 2013). Creating a culture of collaboration in which teachers can freely and safely debate and discuss their ideas is a key, if elusive, aspect of effective change. Literature suggests that there is no foolproof way to create this culture, but that key practices are deprivatizing practice, addressing conflict openly and constructively, clarifying roles, and building consensus (Choi, 2013).

### ***Individualized Feedback***

In the context of this research, coaching is defined as the process of improving teacher practice through the process of observation and reflection. In order for teaching to improve, the expected outcome must be clear, teachers must be motivated to improve, and teachers must have the necessary skills and knowledge to implement new or improved practices (Herrera, 2014). The transfer of skills and knowledge from the department workshop setting to the classroom is greatly facilitated through the process of individualized feedback, and through the process of individual reflection (Herrera, 2014). Teacher's belief in the practices they are being asked to adopt will greatly increase as they experience greater efficacy, and short coaching cycles can greatly improve that process (Herrera, 2014). As a result, I conclude that department collaboration is not enough to ensure that teachers transfer their knowledge to the classroom, and so I plan to follow up each collaborative session with an individual coaching cycle.

### ***Conclusion***

Standards-based grading is a practice that has deep implications for student learning as well as teacher practice. When teachers, students, and families have a shared understanding of the goals and purposes of assessing for learning, the results can be transformational for students. Grades should be true reflections of student learning, and should be tools to help students guide their own learning process. In order for ARISE to improve its implementation of standards-based grading, confidence must be built in the teachers. If teachers improve their understanding of the ideas behind standards-based grading, collaborate to bring clarity to their expectations for student work, and receive technical support to improve their ability to assess for learning, they will see the impact in their own classrooms and thereby increase their engagement and sense of efficacy.