

A Didactic Explanation of Standards-based Grading

Bo Zhang

University of Wisconsin – Milwaukee

Abstract

While standards-based grading (SBG) has been implemented in school districts across the country, it faces two major challenges. As research on standards-based grading is limited, many practices under SBG are not substantiated by empirical evidence. Meanwhile, not all teachers are well prepared to implement SBG in the classroom. The lack of training and professional development on SBG has caused anxiety and fear among teachers and confusion among students.

The goal of this didactic guide is two-folded: to serve as an instructional aid for classroom teachers to help them more prepared for conducting SBG, and to serve as a resource reference for educational researchers to conduct more empirical research on SBG. This writing focuses on how to conduct standards-based grading with the traditional point-based grading in comparison. This guide starts with the rationales for standards-based grading, then moves to its major components, and ends with a discussion of major challenges in using standards-based grading.

Standards-based Assessment

Standards-based assessment has been adopted by most states in the United States (Guskey et al., 2011). By 2015, 44 states have joined the Common Core Standards (National Governors Association Center for Best Practices, & Council of Chief State School Officers, 2010). Meanwhile, standards-based grading is also getting more and more momentum in many school districts. However, research on standards-based grading has not caught up with this level of popularity of use. The amount of research on this topic is still limited (Selbach-Allen, et al., 2020; Brookhart, et al., 2016), which leaves many practices under SBG not substantiated by empirical evidence (Hany et al., 2016). More importantly, not all teachers are well prepared to implement SBG in their classroom (Song, et al., 2022). The lack of training and professional development on SBG has caused anxiety and fear among teachers (Zusho et al., 2023). Without systematic training, teachers may conduct SBG very differently (Selbach-Allen, et al., 2020). This inconsistency in practice in turn can create anxiety and lack of trust among students (Pollio & Hochbein, 2015). The goal of this didactic guide is two-folded: to serve as an instructional aid for classroom teachers to help them more prepared for conducting SBG, and to serve as a resource reference for educational researchers to conduct more empirical research on SBG.

Standards-based grading can be loosely defined as grading students' proficiency by "clearly specified learning goals and performance standards" (Tomlinson & McTighe, 2006). In other words, under SBG, students will be graded on specific content or skill standards that they have been taught (O'Connor, 2017). Other terms that have been used for SBG include specifications grading (Nilson & Stanny, 2015) and mastery grading (Fernandez, 2021). While different, these two names actually capture the important nature of SBG of being specific and mastery-oriented. While a good classroom grading system should be based on course objectives, what separates SBG from the traditional point-based grading (PBG) system, is that course objectives must be clearly defined under SBG. This point will be illustrated repeatedly in this guide. Note that while SBG

has been used as the short form of standards-based grading (Iamarino, 2014), PBG for point-based grading is less so in the literature but more for the convenience of the current discussion.

The popularity of SBG may be attributed to the interaction of two forces that have shaped the education landscape in the last four decades, namely, standards-based education and formative assessment. While standards movement has a tradition in the education landscape, it was accelerated by two major events: the publishing of *A Nation at Risk* in the 1980s and the adoption of Common Core Standards in the 2010s. Partially based on international test results, *A Nation at Risk* (United States National Commission on Excellence in Education, 1983) called the nation to the attention that American students underperformed against other nations and the excellence of education should be improved. Afterwards, the No Child Left Behind legislation (NCLB) mandated that all students be proficient in reading, mathematics and science. Through this process of seeking excellence and reaching proficiency, various standards were developed and used to measure student achievement. This standards movement witnessed its epitome in the creation of the Common Core Standards (National Governors Association Center for Best Practices, & Council of Chief State School Officers, 2010). These standards specify what students should achieve in language arts, mathematics, and science. It is believed that Common Core standards have helped make educational standards and learning goals more consistent, transparent and rigorous for classroom instruction (National Governors Association Center for Best Practices, & Council of Chief State School Officers, 2010). Thanks to the Common Core Standards and the Next Generation Science Standards (National Research Council, 2013), teachers more than ever before are aware of the ubiquitous roles that educational standards play in their teaching.

Formative assessment, the other force, has long been promoted by educational research community (Black & William, 1998). However, the increasing use of standardized testing for summative purposes under NCLB has overshadowed the effort on the formative front. Still, the formative nature of classroom assessment requires a viable grading system to fulfill at least two roles: accurately measure the current achievement level and promptly update progress for future learning. The traditional point-based grading shows its disadvantage on the second role, which will be elaborated later. The convergence of these two forces requires a grading system that can track students' progress towards clearly defined learning goals, a role supposedly fulfilled by SBG.

Standards-based grading should not be confused with standardized testing such as state tests and college admission tests (e.g., ACT and SAT). SBG is the formative assessment of student learning while learning is in progress in classroom. It is mainly criterion-referenced in that the performance of students should not be compared to each other but rather to the aligned educational standards. Standardized testing, on the other hand, aims to provide testing in a standardized manner so that test content, test administration and test scoring will remain the same for all students. It is generally more summative than formative and often results in norm-referenced interpretations, such as comparing a student's performance to their peers in terms of percentile rank.

Why Standards-based Grading?

Almost all educators believe that grades should have meaning (Anderson, 2018; Scriffiny, 2008; Wiggins, 1996). It is not that grades based on the point-based system do not have meaning but that their meaning can be hard to interpret. For instance, a student received a B for a cumulative score of 80% over a semester. This grade gives an impressionistic evaluation that this student is on track but fails to convey exactly what this student can or can't do, hence it is not helpful for future learning. SBG aims to make the meaning of grades more explicit by incorporating educational standards and learning goals into the grading metric. For example, if a score report

says a student received a B for having met some learning goals but not others, it communicates clearly what the student has achieved and more importantly, what more can be done.

SBG can also facilitate the formative assessment of students. SBG provides a profile of performance on multiple learning goals. A student can evaluate the performance profile across goals to find the weaker areas to work on. Teachers can evaluate the same performance profiles but across students in the class to find the weaker goals to be reinforced. SBG also has the mechanism to incorporate the multiple assessments of the same learning goal by new assessments or retakes into the grades, which encourages teachers to provide multiple opportunities and students to take advantage of them.

SBG makes the standards or learning goals more explicit to teachers than PBG. By nature, SBG is based on clearly defined standards. While those standards should have also driven PBG, they are more implicit, and teachers can conduct PBG without using them. SBG is based on the direct alignment among three elements: learning goals, classroom instruction, and grading. In that sense, one may argue that SBG can help improve classroom instruction. Marzano (2003) has shown that instruction with clear targets can increase achievement by up to 20%. There are also other indications that SBG may boost student overall achievement (Craig, 2011). Pollio and Hochbein (2015) found that standards-based math grades were a stronger predictor of performance on a high school state assessment than traditional grading practices. This positive effect of standards-based grading on student achievement is not without controversy. Using NAEP data from 1990 to 2017, Song et al (2022) found no significant improvement of student achievement on mathematics or reading, which echoes the findings from Loveless (2016). Townsley and Varga (2018) also showed that SBG does not help with college readiness.

SBG can make the communication of student performance with students and parents easier. Similar to PBG, SBG can report one overall proficiency level score. More importantly, teachers can provide a performance chart for each student, which describes what a student has achieved in terms of relevant standards (see examples below).

Components of Standards-based Grading

There are three major components under standards-based grading: educational standards, assessment tasks, and scoring rubrics. By nature, the first component of SBG is educational standards. SBG starts with clearly phrased standards. While teachers may create their own standards, the priority goes to standards developed by states and school districts under the Common Core Standards. Educational standards such as the Common Core ones are general statements about the content and skills that students are expected to master in a specific period. These standards are usually defined by the school year. In other words, they describe what students should be able to perform at the end of a grade year. Due to its general nature, it is not easy to organize instruction or assessments directly by these standards. Instead, learning goals are derived from them to describe more specifically what students should be able to achieve in a relatively shorter period of time, such as a class for a unit. Learning goals may also be called learning targets, learning objectives, lesson objectives, instructional targets, course objectives, or even educational standards (McMillan, 2017).

Here is an example of one educational standards and four corresponding learning goals for 3rd grade writing (Common Core State Standards Initiative, 2024).

Educational Standard:

1. Write opinion pieces on topics or texts, supporting a point of view with reasons.

Learning Goals:

- a. Introduce the topic or text they are writing about, state an opinion, and create an organizational structure that lists reasons.
- b. Provide reasons that support the opinion.
- c. Use linking words and phrases (e.g., because, therefore, since, for example) to connect opinion and reasons.
- d. Provide a concluding statement or section.

These goals are still labelled as standards in the original document, but it is clear they are more specific, hence more helpful for classroom instruction than the overall standard. For consistency and clarity, this writing uses the term “learning goal” to refer to specific standards.

Educational standards and learning goals can come in layers, as illustrated in the following example. Layer 1 is one Common Core standard for the 4th grade (Common Core State Standards Initiative, 2024). Layer 2 has two sub-standards, each carrying multiple learning goals at the third layer.

Common Core Standard: Use knowledge of language and its conventions when writing, speaking, reading, or listening.

Sub-standard a: Choose words and phrases to convey ideas precisely

Learning Goal 1: Identify figurative language

Learning Goal 2: Apply figurative language to writing/speech

Sub-standard b: Choose punctuation for effect.

Learning Goal 1: Explain the purpose of various punctuation marks

Learning Goal 2: Demonstrate/identify correct use of various punctuation marks.

To assess standards and learning goals like the ones above, assessment tasks aligned with these goals and targets need to be developed. Such test development process usually starts with a test blueprint or test specifications, which specifies what item measure what learning target. Table 1 is an example of a test blueprint adapted from Smarter Balanced Assessment Consortium (2024, p. 98). It assesses the content area of concepts and procedures for the 6th grade math.

Table 1*Example of Test Blueprint*

Assessment Targets	DOK	No. of Items	Total No. of Items
E. Apply and extend previous understandings of arithmetic to algebraic expressions.	1	5-6	16-19
F. Reason about and solve one-variable equations and inequalities.	1, 2		
A. Understand ratio concepts and use ratio reasoning to solve problems.	1, 2	3-4	
G. Represent and analyze quantitative relationships between dependent and independent variables.	2	2	
B. Apply and extend previous understandings of multiplication and division to divide fractions by fractions.	1, 2		
D. Apply and extend previous understandings of numbers to the system of rational numbers.	1, 2	2	
H. Solve real-world and mathematical problems involving area, surface area, and volume.	2		
I. Develop understanding of statistical variability.	1, 2	4-5	
J. Summarize and describe distributions.	1, 2		

Note. Adapted from Smarter Balanced Assessment Consortium (2024, p. 98)

Column 1 in Table 1 lists the learning targets to be assessed by this test. Column 2 specifies the complexity of the learning target by the level of the depth of knowledge (DOK) (Webb, 1999). Column 3 lists the number of items desired for each learning target. Note that the number of items is relatively small. One reason is this is a computer adaptive test where test items are supposed to be tailored to the ability level of each student. A test blueprint like this is essential for standards-based grading in that it not only helps teachers construct high-quality test items but also guides them to grade each learning goal by its corresponding items.

Once test items are developed, the next step is to prepare scoring rubrics. Scoring rubrics are not unique to SBG. Under PBG, teachers also use scoring rubrics. However, there are several distinct features in using rubrics under SBG. First, rubric use is more prevalent under SBG in that there can be three types of rubrics. First, each standard and learning goal carries a rubric, which may be labelled as standard rubric. This rubric describes the levels of achievement for a standard or goal (e.g., exceeding standard, meeting standard, and not meeting standard), regardless of the nature of assessment tasks. That is to say, even if a standard is assessed solely by objectively scored items like multiple-choice items, it still needs a rubric. Moreover, a standard rubric may be used multiple times. When a standard is re-assessed by a new task or a retake, its rubric can be applied again to update the performance level. Second, a rubric is also required to combine the grades from multiple standards to derive one overall grade like a semester grade. This kind of rubric can be labelled as overall grade rubric. Third, for the subjectively scored items such as essay questions or performance-based tasks, scoring rubrics need to be developed. In this case, similar to the ones

used under PBG, their main goal is to reduce the subjectivity in scoring by assigning values to the distinct levels of observable behaviors. This kind of rubric may be labeled as subjective task rubric.

The following examples show these three types of rubrics under SGB. Example 1 in Table 2 is a rubric for a learning goal (adapted from Common Goal Systems, 2024). It is used to designate a performance level by combining information from all assessment tasks related to the learning goal that students can multiply multi-digit numbers.

Table 2

Example of Learning Goal Rubric

Performance Level	Criterion
4	Students have multiplied multi-digit numbers using multiple strategies and can generate multi-digit multiplication equations from word problems.
3	Students have multiplied multi-digit numbers using strategies demonstrated by the teacher.
2	Students have multiplied a single-digit number with a multi-digit number. Students have also multiplied multi-digit numbers with teacher or peer prompting.
1	Students have multiplied single digit numbers, but don't know how to approach multiplying multi-digit numbers

Note. Adapted from Common Goal Systems, 2024.

One advantage of this rubric lies in its clarity in distinguishing the performance levels, which helps teachers use it. For instance, when a new assessment task shows a student is able to independently do multi-digit number multiplication, the performance level can be updated to 3.

In practice, the above numeric values can be replaced by more substantive descriptors such as beginning (Level 1), progressing (Level 2), proficient (Level 3), and exceptional (Level 4) (Guskey & Bailey, 2001). Teachers can also combine Levels 1 and 2 to form a three-point rubric – not meeting standard (Levels 1 and 2), meeting standard (Level 3) and exceeding standard (Level 4). Furthermore, teachers can even form a binary rubric, such as combining Levels 1 and 2 as not meeting the standard and Levels 3 and 4 as meeting the standard. While proficiency labels may seem routine, research has shown the language used to label the levels matters in terms of the interpretation of SBG scores by students and their motivation to move to the next level (O'Donnell & Sireci, 2022).

To report an overall grade for a period of instruction such as a semester, scores on multiple learning goals/standards need to be combined. It is worth noting that school districts may choose not to report such an overall grade. The next example is a rubric for assigning a summative grade (adapted from Nanini, 2024). This will combine the scores on multiple learning goals to create one grade. This example takes a holistic approach to assign an impressionistic overall score. The numeric values can also take the descriptive labels as discussed in Example 1, or traditional letter grades.

Table 3*Example of Summative Score Rubric*

Performance Level	Criterion
4	Exceeds Grade Level Standards: Demonstrates understanding beyond grade level standards consistently and independently.
3	Meets Grade Level Standards: Demonstrates understanding of grade level standards consistently and independently.
2	Partially Meets Grade Level Standards: Demonstrates partial understanding of grade level standards with or without support.
1	Does Not Meet Grade Level Standards: Demonstrates minimal (or no) understanding and does not meet grade level standards even with support.

Note. Adapted from Nanini, 2024

Many other ways can be used to derive such a summative grade. They are similar to how teachers compute the final grade based on multiple assessments in a semester under the point-based system. The only difference is scores from multiple standards rather than from different tests will be combined. One such method is the median method. Scores from each standard or goal will be rank ordered and the median of those scores will be the overall grade. For instance, if a student's scores on five learning goals are 1 1 2 4 4, the overall grade will be 2, or partially meets grade level standards, if using the labels in Example 2. Note that this calculation assumes all the goals are equally important. If not, weights can be assigned to standards. For instance, if the last standard is twice as important as the other four standards, its grade should be counted twice. One may also be interested in computing the mean of the above five scores, which turns out to be 2.4. As it is not easy to assign a grade based on the decimal point, the median is generally preferred for its convenience and its robustness to extreme scores (O'connor, 2017).

While the reporting of the final grade under SBG may sound similar to that under PBG, especially when letter grades like A and B are used, the difference is, for each grade under SBG, teachers will be able to provide more informative feedback for future learning. Teachers can give a profile of student performance on a set of clearly defined learning goals (Guskey et al., 2011). That way, compared to point-based grading, students will have a better idea on where they stand, and more importantly, on what to work on next. Here is a simplified example of a profile of student achievement on multiple goals.

Table 4*Profile of Learning Targets*

	Learning Goal 1	Learning Goal 2	Learning Goal 3	Learning Goal 4
Student 1	4	4	1	1
Student 2	2	2	1	2
Student 3	4	3	1	4

From this example, one can see Student 1 needs more work on Learning Goals 3 and 4. Meanwhile, the class overall needs more work on Learning Goal 3. For more detailed examples of student profiles on learning goals, see Owens (2024).

Finally, Example 3 in Table 5 is a task-based rubric (Delaware Department of Instruction, 2020), which can be used for both SBG and PBG. It helps assign grades for subjectively-graded tasks such as a writing task in this case. This rubric is for Grade 2 opinion writing. The columns in the table represent the achievement levels. The rows represent the content areas. Note that the content areas are weighted to account for their differential contribution to the task. The relevant educational standards, such as 3W1a and 3W1c, can be found on the Common Core Standards website (Common Core State Standards Initiative, 2024 and some of them were given in the above examples on learning goals. The quality of a rubric like this depends on the clarity of the description in each cell, such as how it reflects the actual achievement level and how easily teachers can use it.

Table 5

Example of Rubric for an Assessment Task

Score of 4 Above Grade Level	Score of 3 At Grade Level	Score of 2 Approaching Grade Level	Score of 1 Below Grade Level	
The writing – <ul style="list-style-type: none"> ▪ introduces the topic or text being written about (3W1a) ▪ skillfully states an opinion (3W1a) ▪ creates an organizational structure that lists reasons (3W1a) ▪ uses linking words and phrases to connect opinion and reasons (3W1c) ▪ skillfully provides a concluding statement or section (3W1d) 	The writing – <ul style="list-style-type: none"> ▪ introduces the topic or book being written about (2W1) ▪ states an opinion (2W1) ▪ uses linking words to connect opinions and reasons (2W1) ▪ provides a concluding statement or section(2W1) 	The writing – <ul style="list-style-type: none"> ▪ attempts to introduce the topic or book being written about ▪ attempts to state an opinion ▪ attempts to use linking words to connect opinions and reasons ▪ attempts to provide a concluding statement or section 	The writing – <ul style="list-style-type: none"> ▪ makes little or no attempt to introduce the topic or book being written about ▪ makes little or no attempt to state an opinion ▪ makes little or no attempt to use linking words to connect opinions and reasons ▪ makes little or no attempt to provide a concluding statement or section 	Organization/Purpose $2 \times \frac{\quad}{\quad} = \frac{\quad}{\quad}$
The writing – <ul style="list-style-type: none"> ▪ skillfully provides reasons to support opinion (3W1b) ▪ uses information from experiences or print and digital sources (3W8) 	The writing – <ul style="list-style-type: none"> ▪ provides reasons to support the opinion (2W1) ▪ uses information from experiences or provided sources (2W8) 	The writing – <ul style="list-style-type: none"> ▪ attempts to provide reasons to support the opinion ▪ attempts to use information from experiences or provided sources 	The writing – <ul style="list-style-type: none"> ▪ makes little or no attempt to provide reasons to support the opinion ▪ makes little or no attempt to use information from experiences or provided sources 	Evidence/Elaboration $2 \times \frac{\quad}{\quad} = \frac{\quad}{\quad}$

<p>The writing –</p> <ul style="list-style-type: none"> ▪ demonstrates a command of grade-level appropriate standard English grammar, usage, and conventions (3L1-2)* ▪ has errors that do not interfere with understanding (3L1-2)* 	<p>The writing –</p> <ul style="list-style-type: none"> ▪ demonstrates a command of grade-level appropriate standard English grammar, usage, and conventions (2L1-2)* ▪ has errors that do not interfere with understanding (2L1-2)* 	<p>The writing –</p> <ul style="list-style-type: none"> ▪ attempts to demonstrate a command of grade-level appropriate standard English grammar, usage, and conventions ▪ has errors that may interfere with understanding 	<p>The writing –</p> <ul style="list-style-type: none"> ▪ makes little or no attempt to demonstrate a command of grade-level appropriate standard English grammar, usage, and conventions ▪ has errors that interfere with understanding 	<p>Language/Conventions</p> $\frac{1}{1} \times \frac{1}{1} = 1$
--	--	--	--	--

Note. From Delaware Department of Instruction, 2024

While SBG makes updating student achievement on a learning goal easier than PGB, how to update student performance on a standard once a new measurement is taken is still up for debate. One can use the most recent assessment. One can also average the two most recent assessments. Which one works better probably depends on the nature of the assessments. If the most recent assessment is comprehensive enough to have all aspects of a standard covered, it may be used as the sole indicator. But if the previous assessment still provides unique information for the standard, it should be included too. Generally speaking, using more assessments will improve the reliability of the assigned score as long tests tend to be more reliable. More technically, Marzano (2006) proposed a power law formula for teachers to calculate one summary score based on the results from multiple assessments. This method takes three components into account: total number of assessments, the score on each assessment, and the order of the assessments. Hooper and Cowell (2014) proposed the so-called “history adjusted true score”, which relies on the two most recent scores. While these two methods are more complicated, they can be programmed into a grading system.

It is also important to point out what is not part of SBG. All non-learning goal related factors, such as effort, participation, and behavioral issues, are explicitly excluded. They should be excluded in PBG as well. This does not imply these factors are not important for student learning. It is just that separate measures should be used to report them, such as by the process marks in the Kentucky example in Guskey et al. (2011).

One issue under SBG is homework. Should they be included? There are strong rationales for not grading homework. One can argue homework should be used for practice under the formative assessment paradigm (Stiggins et al., 1989). One can also argue that it is hard to know how much help a student has received for doing homework, thus homework may not measure what the student knows. On the other hand, one main drawback of not using homework for grading is the loss of information on learning goals. Homework provides constant measurement of learning goals than the less frequent exams or tests. Grading homework can provide frequent and prompt feedback to students, which is conducive to learning (Hattie & Timperley, 2007). The other problem of not including homework in grades is the possible adverse effect on students’ motivation to do homework, thus reducing its formative practice effect. More research is needed on how to best incorporate homework into SBG.

Standards-based Grade Reporting vs. Point-based Grade Reporting

Tables 6 and 7 are examples of grade books for PBG and SBG, which were adapted from Scriffiny (2008).

Table 6*Example of PBG Grade Book*

Name	Homework Average	Quiz 1	Chapter 1 Test
John	90	65	70
Bill	50	75	78

Note. Adapted from Scriffiny (2008).

Table 7*Example of SBG Grade Book*

Name	Goal 1: Write an Alternate Ending of a Story	Goal 2: Identify the Elements of a Story	Goal 3: Compare and Contrast Two Stories
John	Partially proficient	Proficient	Partially proficient
Bill	Proficient	Proficient	Partially proficient

Note. Adapted from Scriffiny (2008).

As shown in the above two tables, SBG score reporting is different from that of PBG. The first difference is what scores are reported. PGB reports scores based on assessment units, such as Quiz 1 or Chapter 1 Test, whereas SBG reports scores on standards or learning goals. Accordingly, scores bear different meaning. Under PBG, the exact meaning of a test score can be vague. In the case of 70% on Chapter 1 Test for John, does it mean he knows 70% of each learning goal or total mastery of 70% learning goals but nothing on the other 30%? SBG, on the other hand, provides more specific information on each learning goal.

Another major difference between SBG and PBG is the clarity of the learning goals on which the assigned scores are based. To practice SBG, learning goals must be clearly defined, such as “Identify the elements of a story” in Table 5. Otherwise, the scoring rubrics will be too vague. That is to say, SBG would not work well without them. While PBG should be based on clearly defined learning goals as well, it does not depend on it. In a point-based system, teachers can add up the correct responses to calculate a total score or a percentage. SBG grading can also be seen as more formative than PBG in that each assessment becomes a means to updating the progress towards meeting a standard. PBG, on the other hand, is more static in that a score is assigned to each assignment and it is harder to update student achievement by multiple assessments.

Not everything is different under SBG and PBG. Both grading systems aim to provide accurate information about the current achievement level. The assessment tasks can also be the same, such as tests, projects, or assignments. Both systems rely on scoring rubrics for subjectively scored tasks, but the role of rubrics is quite different under SBG and PGB, as discussed above.

Challenges in Using SBG

The challenges of using SBG lie in both the assessment itself and its administration. For the assessment itself, the biggest challenge should be the collection of the large amount of assessment information, as outlined above. SBG involves developing all the necessary standards, learning

goals, assessment tasks, and scoring rubrics. Take scoring rubrics alone, rubric is needed for each learning goal, for each reporting period (if an overall score is desired) and for each subjectively graded assessment task. This requires collective effort and support from schools and school districts, as the current teaching materials may not be aligned with SGB (Bay-Williams et al., 2016).

The second challenge in assessment lies in the processing of the large amount of information once collected. Under standards-based grading, teachers need to match assessment tasks with learning goals. They also need to provide a score for each learning goal. That is more work than PBG where teachers can calculate a total score for each assessment. Professional development needs to be provided to improve teachers' competence in SBG grading. Otherwise, they may still conduct PBG in the name of SBG.

These two assessment challenges in turn create challenges in the SBG implementation (Peters et al., 2017). The first is how to support teachers for SBG grading. The lack of training has caused anxiety and fear among teachers and uneven use of SBG in classroom (Zusho et al., 2023). Without systematic training, teachers tend to conduct SBG in their own way (Selbach-Allen, et al., 2020). This inconsistency can create anxiety among students (Pollio & Hochbein, 2015). Stornaiuolo et al. (2023) introduced a specific yet flexible system that school districts can use to support teachers for standards-based instruction. Efforts have also been made to improve teacher's knowledge about SBG, such as by using the ESBW short scale (Müser et al., 2023).

Implementing retakes of assessment can also be challenging. Consistent with the principles of formative assessment, retakes help SBG monitor students' continuous progress towards the mastery of learning goals. Grading the retake helps capture the current achievement level after remedial work has been done. The challenge, though, is how to manage them. Grading retakes adds to the already high demand of time for teachers under SBG. There are concerns that students may take advantage of retakes differently. For instance, at the college level, female students are more willing to retake than male students (Lewis, 2020). One more equitable practice of reassessment is to provide second chance exams to all students instead of letting students self-choose (Fernandez, 2021). On the concern that retakes may inflate grades, students who regularly miss classes or assignments may still not do well even with retake opportunities (Selbach-Allen et al., 2020).

Finally, SBG also poses some technical challenges in test construction and score reporting. Under SBG, the score for an educational standard is based on the items assessing that standard only. As the number of such items is generally small, how to ensure appropriate test validity can be challenging. In other words, SBG requires constructing high-quality test items. It also requires finding the best way to report those scores, such as the techniques developed for subscale score reporting (Stone et al, 2009).

Concluding Remarks

In alignment with the overall standards-based movement in education, standards-based grading has gained momentum in schools. To do it correctly is important not only for assigning fair grades but also for ensuring the overall quality of standards-based education. SBG starts with clearly defined educational standards and learning goals. Teachers conduct classroom instruction based on those goals and standards. Accordingly, students will be evaluated by standards-based grading. In that sense, SBG, along with standard development and standards-based instruction, is an essential component of standards-based education.

Standards-based grading is more challenging than point-based grading. It requires more work from school districts, schools, and teachers. Districts need to develop appropriate educational standards and learning goals. They may also need to provide high-quality assessment tasks and scoring rubrics. By choosing standards-based grading over point-based grading, teachers need to collect and process more assessment information. School districts and teachers also need to find appropriate ways to communicate to students (Scarlett, 2018) and parents (Peters, et.al., 2017).

Standard-based grading brings promises and challenges to student assessment. Its success depends on the support from the educational research community. More research should be devoted not only to providing empirical evidence for the best practices under SBG, such as how to handle assignments, but also to finding innovative practices with the change of overall student learning landscape, such as how to integrate AI into SBG. More research is also urgently needed on how to provide effective professional development to teachers so that they will not only be proficient in using SBG for their class but also able to adapt it to meet the needs of individual students.

References

- Anderson, L. W. (2018). A critique of grading: Policies, practices, and technical matters. *Education policy analysis archives*, 26, 49.
- Bay-Williams, J., Duffett, A., & Griffith, D. (2016). Common core math in the K–8 classroom: Results from a national teacher survey. Thomas B. Fordham Institute.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5, 7–74.
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., & Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research*, 86(4), 803–848.
- Common Core State Standards Initiative (2024, October 10). *Common Core State Standards for English Language Arts & Literacy in History/Social Studies, Science, and Technical Subjects*. Retrieved from https://learning.ccsso.org/wp-content/uploads/2022/11/ELA_Standards1.pdf
- Common Goal Systems. (2024, October 10). *What Is Standards-based Grading?* Retrieved from <https://www.teacherease.com/standards-based-grading.aspx>
- Craig T. A. (2011). *Effects of standards-based report cards on student learning*. (Doctoral dissertation). Retrieved from <https://repository.library.northeastern.edu/files/neu:1127>
- Delaware Department of Instruction. (2024, October 10). *Opinion Writing Rubric*. Retrieved from <https://www.doe.k12.de.us/site/handlers/filedownload.ashx?moduleinstanceid=6016&dataid=15719&FileName=Gr2OpinionR.pdf>
- Fernandez, O. E. (2021). Second chance grading: An equitable, meaningful, and easy-to-implement grading system that synergizes the research on testing for learning, mastery grading, and growth mindsets. *Primus*, 31(8), 855–868.
- Guskey, T. R., & Bailey, J. M. (2001). *Developing grading and reporting systems for student learning*. Corwin Press.
- Guskey, T. R., Swan, G. M., & Jung, L. A. (2011). Grades that mean something: Kentucky develops standards-based report cards. *Phi Delta Kappan*, 93(2), 52–57.
- Hany, K., Proctor, M., Wollenweber, J., & Al-Bataineh, A. (2016). Teacher perception of standards-based grading: Implication and effectiveness. *Journal of Teaching and Education*, 5(1), 749–764.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81–112.
- Hooper, J., & Cowell, R. (2014). Standards-based grading: History adjusted true score. *Educational Assessment*, 19(1), 58–76.
- Iamarino, D. L. (2014). The benefits of standards-based grading: A critical evaluation of modern grading practices. *Current Issues in Education*, 17(2).
- Lewis, D. (2020). Gender effects on re-assessment attempts in a standards-based grading implementation. *Primus*, 30(5), 539–551.

- Loveless, T. (2016). The 2016 Brown Center report on American education: How well are American students learning? Part I: Reading and math in the Common Core era. Brookings Institution.
- Marzano, R. J. (2003) *What Works in Schools: Translating Research into Action*. Alexandria, VA: ASCD.
- Marzano, R. J. (2006). *Classroom assessment & grading that works*. Alexandria, VA: Association for Supervision and Curriculum Development.
- McMillan, J. H. (2017). *Classroom assessment: Principles and practice that enhance student learning and motivation*. Pearson.
- Müser, S., Fleischer, J., Kunina-Habenicht, O., & Leutner, D. (2023). The ESBW short scale: A test for assessing teachers' standards-based educational knowledge. *European Journal of Psychological Assessment*.
- Nanini, K. (2024, October 10). *Standards-based Grades*. Retrieved from <https://youngteacherlove.com/walking-through-standards-based-grading/>
- National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Common Core State Standards*. Retrieved from <http://www.corestandards.org>
- National Research Council. (2013). *Next generation science standards: For states, by states*.
- Nilson, L. B., & Stanny, C. J. (2015). *Specifications grading: Restoring rigor, motivating students, and saving faculty time*. Routledge.
- O'Connor, K. (2017). *How to grade for learning: Linking grades to standards*. Corwin Press.
- O'Donnell, F., & Sireci, S. G. (2022). Language matters: Teacher and parent perceptions of achievement labels from educational tests. *Educational assessment*, 27(1), 1–26.
- Owens, K. (2024, October 10). *A Beginner's Guide to Standards Based Grading*. Retrieved from <https://blogs.ams.org/matheducation/2015/11/20/a-beginners-guide-to-standards-based-grading/>
- Peters, R., Kruse, J., Buckmiller, T., & Townsley, M. (2017). "It's Just Not Fair!" Making Sense of Secondary Students' Resistance to a Standards-Based Grading. *American Secondary Education*, 45(3), 9.
- Pollio, M., & Hochbein, C. (2015). The association between standards-based grading and standardized test scores as an element of a high school model reform. *Teachers College Record*, 117(11), 1–28.
- Scarlett, M. H. (2018). "Why did I get a C?": Communicating student performance using standards-based grading. *InSight: A Journal of Scholarly Teaching*, 13, 59–75.
- Scriffiny, P. L. (2008). Seven reasons for standards-based grading. *Educational Leadership*, 66(2), 70–74.
- Selbach-Allen, M. E., Greenwald, S. J., Ksir, A. E., & Thomley, J. E. (2020). Raising the bar with standards-based grading. *Primus*, 30(8-10), 1110–1126.

- Smarter Balanced Assessment Consortium. (2024, October 10). 2017–18 Summative Technical Report. Retrieved from <https://portal.smarterbalanced.org/library/en/2017-18-summative-assessment-technical-report.pdf>
- Song, M., Garet, M. S., Yang, R., & Atchison, D. (2022). Did states' adoption of more rigorous standards lead to improved student achievement? Evidence from a comparative interrupted time series study of standards-based reform. *American Educational Research Journal*, 59(3), 610–647.
- 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.
- Stone, C. A., Ye, F., Zhu, X., & Lane, S. (2009). Providing subscale scores for diagnostic information: A case study when the test is essentially unidimensional. *Applied Measurement in Education*, 23(1), 63–86.
- Stornaiuolo, A., Desimone, L., & Polikoff, M. (2023). “The good struggle” of flexible specificity: Districts balancing specific guidance with autonomy to support standards-based instruction. *American Educational Research Journal*, 60(3), 521–561.
- Tomlinson, C. A., & McTighe, J. (2006). *Integrating differentiated instruction & understanding by design: Connecting content and kids*. ASCD.
- Townsley, M., & Varga, M. (2018). Getting High School Students Ready for College: A Quantitative Study of Standards-based Grading Practices. *Journal of Research in Education*, 28(1), 92–112.
- United States National Commission on Excellence in Education. (1983). A nation at risk: The imperative for educational reform. United States Department of Education. Washington, D.C.
- Webb, N. L. (1999). *Alignment of Science and Mathematics Standards and Assessments in Four States*. Research Monograph No. 18.
- Wiggins, G. (1996). Honesty and fairness: Toward better grading and reporting. In T. R. Guskey (Ed.), *ASCD yearbook, 1996: Communicating student learning* (pp. 141–177). Alexandria, VA: Association for Supervision and Curriculum Development.
- Zusho, A., Kumar, R., & Bondie, R. S. (2023). Transforming fear into rigor, love, freedom, and joy: A new paradigm of standards-based reform. *Educational Psychologist*, 58(4), 244–260.