Special Issue: Exploring Current & Future Grading Practices in K-12 Schools
Best Practice Article

JSARD
open access | est. 2016
Volume 11, Issue S1 (2026), pp. 121-130
Journal of School Administration Research and Development
ISSN: 2470-8496 Print/2470-850X Online
ojed.org/JSARD

Boolean Logic Gates: A New Method of Grade Calculation for Standards-Based Grading

Shawn Carney
Fargo Public Schools, USA
Teresa Shume
North Dakota State University, USA

ABSTRACT

Standards-based grading (SBG) has emerged as a reform movement aimed at improving the accuracy, equity, and transparency of grading systems. SBG involves grading at two levels: the overall course or subject area, and individual standards. This article examines multiple methods for calculating grades for individual standards in SBG systems, including logic rules, reliance on most recent evidence, various types of means, power law projections, and frequency-based approaches such as mode and bi-modal matrices. Each method is analyzed based on four criteria: quantity of assessment evidence, recency of assessment evidence, how easily and transparently students can track their own progress, and whether students who learn content more slowly are disadvantaged by the grading method. A novel method, Boolean logic gates, is proposed for determining a student's level of proficiency on an individual standard. By systematically incorporating conditional logic, Boolean logic gates are a promising and innovative grade calculation method that has the potential to provide educators with nuanced and actionable representations of student learning.

***Keywords*:** grading, standards-based grading, classroom assessment

Many schools today are seeking to reform their grading systems by shifting away from traditional grading and implementing standards-based grading (SBG) (Peters & Buckmiller, 2014). The central purpose of grades is to communicate the extent to which students have mastered the learning targets; thus, accuracy is a key tenet that underpins effective grading systems (Brookhart et al., 2016). Traditional grades are typically based on calculating the mean of an accumulation of points earned through various assessment methods (e.g., homework, quizzes, tests, projects), as well as compliance with various classroom policies and expectations (Guskey & Jung, 2012). Traditional grading has been criticized as invalid and unreliable for measuring and reporting students' academic achievement, and has been called "a hodgepodge of attitude, effort, and achievement" (Brookhart, 1991, p. 36). SBG, in contrast, reflects students' understanding based on the standards being taught. SBG approaches are not obscured by non-academic factors such as class participation, effort, or meeting deadlines (Townsley, 2019). Moreover, homework and other formative assessments intended to support students on their pathway of learning are also excluded from SBG (Brookhart et al., 2020). Consequently, SBG systems provide a more accurate, equitable, and actionable representation of what a student knows (Guskey et al., 2024).

While a variety of approaches to SBG exist (Link & Guskey, 2022), Townsley and Wilcox (2024) have identified four core principles that distinguish SBG from traditional grading approaches. For the purposes of this article, SBG will be defined as grading systems that adhere to the following key tenets provided by Townsley and Wilcox (2024, p. 76):

1. Grading should communicate students' current level of learning based on standards.
2. Teachers should use a finite number of performance categories to assess students' learning.
3. Grades should be based on summative assessments rather than formative assessments or student behaviors.

4. Students should have multiple opportunities to demonstrate their learning.

The definition from Townsley and Wilcox (2024) was chosen for this article because it is current and mostly agreed upon by SBG scholars.

Grading at Two Levels

As school leaders navigate the challenges of establishing SBG systems, crucial decisions must be made to select specific methods for calculating standards-based grades (Heflebower et al., 2018). These decisions typically occur at two levels: conveying students' overall learning in a course or subject area and communicating a summary of student proficiency for each standard. At the overall course level, an informative way to communicate student learning is to present proficiency statements for each individual standard, providing stakeholders with clear messages about what a student knows, understands, and is able to do (Iamarino, 2014). However, schools frequently encounter a pervasive need to report out course grades as single letters due to requirements for college admissions, criteria for scholarship applications, and eligibility for co-curricular activities, among other reasons (O'Connor, 2018). To translate a student's scores on individual standards into a single course letter grade, the use of "logic rules" is regarded as an appropriate and suitable method (O'Connor, 2018).

When using logic rules to determine a course grade, each conditional statement, or "rule," is applied in turn until one is found that corresponds to the student's collection of standard-level grades. Table 1 provides an illustrative example of a set of logic rules that could be applied to determine a course grade. In this example, "4" indicates advanced, "3" indicates proficient, "2" indicates partially proficient, and "1" indicates developing.

Table 1

Logic Rules Example

Course Grade	Logic Rule
A	All standard scores are at the 3 or 4 level
B	One standard score of 2, with no standard scores of 1
C	More than 1 standard score of 2, with no standard score of 1
D	One standard score of 1
F	More than 1 standard score of 1

Methods to determine overall course grades, such as logic rules, are predicated on determining students' scores for each standard, the second major type of decision educators must make when implementing SBG in their school.

The aim of this article is to analyze a variety of grade calculation methods for translating a set of assessment scores pertaining to a particular standard into a single summary score for that standard. The grade calculation methods analyzed in this article include logic rules, most recent score, three types of means (calculated average, average of the three highest scores, and decaying average), the Power Law, and two types of frequencies (mode and bi-modal matrix). Finally, a novel system that uses Boolean logic gates is proposed and analyzed in comparison with the other calculation methods.

Purely professional decision-making based on evidence has been omitted because this article focuses on comparing calculation methods. Professional decision-making, simply put, is not a calculation method. It requires evaluating qualitative properties of assessments that earned quantitative scores, which cannot be numerically represented. Some SBG experts emphasize the value and importance of teachers' professional judgment in determining SBG grades (Guskey & Jung, 2016; Townsley & Wear, 2020). This article introduces a novel calculation method that incorporates professional judgment; however, it does not evaluate professional judgment as a standalone alternative to the calculation methods considered.

APPROACH TO ANALYSIS

Criteria

In this article, we analyzed a variety of grade calculation methods using four criteria derived from research literature and the authors' combined teaching experience of over four decades, consisting of 17 years for one author and 32 for the other. First, we considered the quantity of evidence used to determine the summary grade for each standard because this impacts the grade's validity. Using a single piece of evidence to determine a grade presumes that the assessment used is entirely valid and reliable. While those are attributes to which educators should aspire, it is not feasible to completely avoid threats to validity and reliability in the design of each assessment (O'Connor, 2018). Some SBG scholars highlight the value of collecting a body of evidence to use when determining grades (e.g., Heflebower et al., 2014). However, such bodies of evidence sometimes conflate formative and summative assessments, which is problematic since SBG should be based exclusively on summative assessments (Townsley & Wilcox, 2024). In sum, using more evidence is better than using less evidence, but the quality and salience of the evidence must also be considered.

Second, we examined whether the most recent data were prioritized. Compared to initial attempts, a student's most recent scores are more likely to accurately indicate the level of proficiency attained by the student (Townsley & Wear, 2020). Some scholars recommend that only the most recent performance that occurred at the end of the grading period should be considered (Feldman, 2019). Recentness of the evidence does contribute to its salience. However, there exists a tension between recentness and quantity since relying exclusively on a single piece of evidence, even if most recent, presents limitations described above. Moreover, restricting grade calculation to the single most recent assessment creates a "high-stakes" assessment that can induce anxiety among students (Schimmer, 2016).

Third, we checked if standard summary scores produced through each calculation method incentivize student growth. Fundamental tenets of effective SBG systems include fostering student self-efficacy, contributing to student confidence, and nurturing a growth mindset (Schimmer, 2016). Beyond opportunities to practice multiple formative assessments and to repeat summative assessments, growth is incentivized by transparency when students can easily track their own progress by calculating their ongoing grade as they earn additional summative scores (O'Connor, 2018). Grading practices should convey the message that eventual success is possible, incentivizing rather than penalizing students for growth (Dueck, 2011).

Last, we considered whether students who take different amounts of time to learn the same content are awarded the same summary score for a standard. Students who eventually attain the same level of proficiency should earn the same grade, regardless of whether one student learned the content faster or slower than another (Townsley & Wear, 2020). Students' rates of learning can be influenced by numerous factors, such as learning disabilities, interruptions to formal schooling, and English language learning needs, among many others. Simply put, this is a matter of equity (Feldman, 2023).

GRADE CALCULATION METHODS

What follows is an explanation of each calculation method identified for determining a summary grade for a single standard and an analysis of each method to identify strengths and weaknesses. For added clarification, each method is applied to the following set of example student scores earned on an SBG grade scale in chronological order: 1, 1, 2, 2, 3, each representing a score for a different summative assessment of relatively comparable length and rigor pertaining to the same standard. Depending on specific instructional contexts, these scores could represent assessments of closely related applications of a single standard or a cluster of distinct subskills related to the standard. In this example, 4 indicates advanced, 3 indicates proficient, 2 indicates partially proficient, and 1 indicates developing.

Logic Rules

As described earlier, logic rules aim to identify the most consistent level of achievement through application of a series of conditional statements. For the purpose of example, one way the logic rules listed in Table 1 could be adjusted to produce a summary grade for an individual standard could be as follows: a standard-level score of 3 requires no 1s or 2s, a 2 requires one score of 2 and no scores of 1, and a score of 1 requires any amount of 1s. This rule set is quickly observed as taking the worst score from the data set and assigning it for the standard score. For argument's sake, a more lenient set of rules could be applied: a standard-level score of 3 requiring no more than one 2 and no 1s, a 2 requiring two or more 2s and no more than one 1, and a 1 requiring more than one 1. Using either set of logic rules, our student with hypothetical scores of 1, 1, 2, 2, 3 would earn a 1 as a summary score for the standard, as they have more than one score evaluated at a 1, even though their most recent score was a 3. Beyond these two examples, logic rules can be operationalized in a wide variety of ways.

The transparency of logic rules can be impacted by the timing of when the number of scores that will be used to calculate the grade is determined. Transparency of logic rules is enhanced when the number of assessment scores is set at the beginning of the marking period, yet this can be constricting to teachers when designing instruction and assessment in response to evolving student needs. Logic rules do not evaluate grades based on recency unless recency is explicitly considered; for example, a set of logic rules could be applied to only the five most recent scores. Additionally, when applying logic rules, the quantity of scores below proficient can be more impactful than the quantity of proficient scores. In effect, this causes all assessments to be "high stakes" and penalizes students who take longer to demonstrate proficiency in a standard. Further, a logic rule approach means that growth may not be incentivized, as once a student earns a lower score, depending on the rules presented, they may be locked out of ever attaining a standard-level score of proficient. In our examples presented above, after receiving one or two scores of 1, the student is not incentivized to grow, as they will never be eligible to earn a 3 in this standard.

Most Recent Score

In this straightforward calculation, only the most recent piece of evidence is used to determine the grade. Using our sample score of 1, 1, 2, 2, 3, the student's grade for this standard would be 3 (proficient) because it was their final assessment score.

A number of SBG experts recommend the use of a single score based on the most recent evidence (Feldman, 2019). Simply taking the most recent evidence to be a student's grade, however, relies heavily on the validity of a single assessment. As Marzano (2010, p. 29) points out, "Even if a very good final assessment is designed, it will still contain error. Consequently, it cannot be absolutely trusted as the true score for an individual student." While this method prioritizes recent evidence and avoids penalizing students for learning concepts more slowly, it does not address the value of multiple assessment results forming a body of evidence (Heflebower et al., 2014). Moreover, some students may put less effort into initial assessments, knowing that scores can be later replaced. Additionally, it is possible for students to be awarded a grade that may, in fact, be an exception.

Types of Means

Reflecting a common understanding of the term "average," the calculated mean is determined by taking the sum of the scores for the pieces of evidence and dividing it by the number of pieces of evidence. Using our example scores of 1, 1, 2, 2, 3, the calculation becomes (1+1+2+2+3)/5, yielding a score of 1.8. A second approach to calculating a mean is to truncate the data set to only the three pieces of evidence with the highest scores. For the data set 1, 1, 2, 2, 3, only the 2's and the 3 would determine the student's standard-level grade. This would calculate as (2+2+3)/3 = 2.3. A third type of mean is a decaying average, where scores are weighted progressively, with the most recent score assigned the highest weight. For this example, the most recent evidence is allocated 35% of the total grade, scaling down to 22%, 20%, 13%, and finally only 10% for the first piece of evidence the student provided. For the set of scores 1, 1, 2, 2, 3 earned in this order, a student's grade calculation would appear as (1*10) + (1*13) + (2*20) + (2*22) + (3*35)/100 = 2.12.

Compared to the most recent method, all three of these methods rely on a more robust body of evidence to support a student grade, though only the decaying average prioritizes recent evidence. Regardless of whether an average is weighted or not, however, using a mean calculation penalizes students who do not achieve proficiency on their first assessment and must demonstrate proficiency beyond the proficient level in order to reach a standard-level grade of proficient, which may discourage students (O'Connor, 2018). As Schimmer (2016, p. 61) points out, "The mean can be mean!" Consequently, no method that relies on calculating any type of average can accurately represent a student's standard-level grade at the end of a grading period because of the need for higher scores to compensate for any that fall below the proficient level (O'Connor, 2018). Moreover, means that produce overall scores with decimals echo the illusion of precision perpetuated by traditional grading methods (Guskey, 2024).

Power Law

Identified as one of the main approaches to standard-level grade calculation in SBG literature (Townsley & Wilcox, 2024), the power law approach is a statistical technique that models score changes over time (Marzano, 2006). This approach is not designed to average grades, but rather to predict the next grade a student would earn based on a statistical model. Applying a power law to the data set of 1, 1, 2, 2, 3, the student grade would be projected as 2.57.

The power law approach draws on the full body of evidence available and largely relies on the sequence in which scores are earned. Because such an approach attempts to fit a data set to a linear projection, it creates a near opposite flaw compared to means. Where means can create a deficit score that penalizes students who learn material more slowly, the power law instead penalizes those who demonstrate proficiency early on. Having low scores early on and higher scores later in the grading period will lead to a higher projected grade by adjusting the slope of the projection. This can lead students to intentionally perform below proficient early on in a grading period because it will increase the projected score later in that grading period. Additionally, because the power law is a projection, a student does not necessarily need to ever achieve proficiency, only to show growth that would project them on a course to demonstrate proficiency. Finally, because of the complexity of the power law statistical calculations, this method may not be as transparent or appealing as more easily explained grading practices (Townsley & Wilcox, 2024).

Types of Frequencies

When the mode is used to determine a grade, whichever score a student earns most frequently is the grade they are assigned. In the case of ties, a decision needs to be made to either select the higher mode, lower mode, or calculate a mean average of the two. If the option of calculating a mean average of the two is chosen, all the drawbacks of mean calculations described above permeate this system as well. Consequently, selecting the lower or higher mode should be considered instead. In the example score set of 1, 1, 2, 2, 3, the modes of 1 and 2 are tied.

A second type of mode is the bi-modal matrix, which references both the most common score and the second most common score that a student earns (Reibel et al., 2025). To determine a grade, the primary and secondary modes are determined, and then a matrix table is used to produce a final score (Reibel et al., 2025). Under each grade in the illustrative example in Table 2, corresponding primary and secondary modes appear in order. For example, a grade of 4 is assigned when a student's modes are 4 and 3, 4 and 2, or 3 and 4. For the sample score set of 1, 1, 2, 2, 3, the modes identified are 1 and 2. Deferring to the lower mode, the summary grade for the standard would be 1. Deferring to the higher mode, the summary grade for the standard would be 2.

Table 2

Example of a Bi-modal Matrix

Grade: 4	Grade: 3	Grade: 2	Grade: 1
4-3	4-1	2-3	1-4
4-2	3-2	2-4	1-3
3-4	3-1	2-1	1-2

These two grading approaches do not penalize students whose learning progress takes more time. However, both are limited by the fact that neither evaluates the entire body of evidence provided by a student nor prioritizes the most recent evidence. Further, it is possible for students to be penalized for earning higher grades that lie outside the modal value. For example, our student who earns the scores 1, 1, 2, 2, 3 over a series of assessments will still earn a 1 or 2, even if they were to score a 4 on their final assessment. For the bi-modal matrix, our student would earn a final grade of 1 if rounding down for ties between modal scores. Notably, had the student scored 2 instead of 3 on the final assessment, the primary mode would have been 2, and the secondary mode would have been 1, resulting in a summary score of 2. Ironically, this student's final grade would be higher if they had scored lower on the last assessment. A student may be disincentivized when they realize they have earned so many similar scores that their last assessments will not impact their final grade or when it is more strategic to earn a lower score.

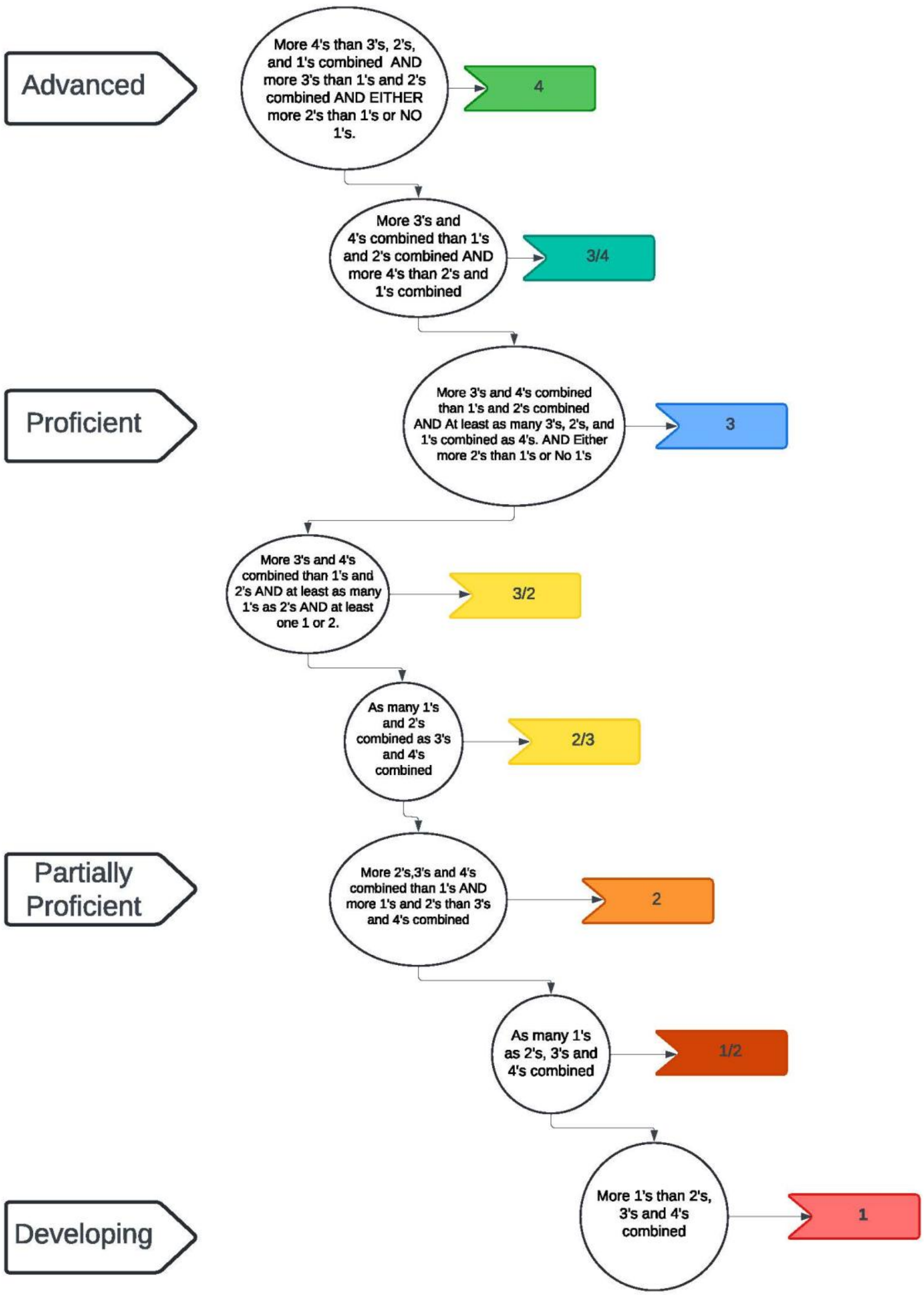
Boolean Logic Gates

Presented below is a novel approach to grade calculation for a single standard based on if-then Boolean logic gates. To interpret a set of scores based on Boolean logic gates, refer to the example provided in Figure 1. If a statement is found to be true, the resulting grade appears in a flag to the right. If the statement is found to be false, progress to the next bubble until a true statement is located. Notably, the Boolean logic gate system identifies instances where the student's grade appears as a split score. In these situations, the teacher should use professional judgment to determine which score best represents the student's proficiency level. Figure 1 offers one example of a Boolean logic gate system; other variants are possible.

Utilizing true/false statements in a similar manner to logic rules used to determine a course grade, Boolean logic gates chain together multiple Boolean statements to analyze a full body of evidence of student learning to determine a grade for a single standard. Compared to logic rules, Boolean logic gates use comparative operators rather than absolutes to evaluate the capacity and frequency of student demonstrations of proficiency. Unlike logic rules, which typically operate purely on the basis of the presence of low scores, Boolean logic gates account for the frequency of higher scores. This allows for bodies of evidence wherein demonstrations of proficiency can have varied levels of success and permits adjustments to the

Figure 1.

Example of a Boolean Logic Gates Grade Calculation



quantity of assessments during a grading period. Additionally, designing grade calculation as an ordered series of gates, rather than a table of separate rules, guides stakeholders to more easily understand the order of progression and what is required to move from one proficiency level to the next, thus promoting greater transparency.

In Boolean logic gate systems, a student's final grade accounts for all data points collected and weighs them against lower-level demonstrations, creating a score that uses more data than modes and avoids the penalizations of means. As pointed out previously, means can create insurmountable obstacles for students to recover from if poor levels of proficiency are demonstrated early on. Boolean logic gates do not require students to show demonstrations above proficient as a mean would in order to account for lower scores; rather, they only require students to demonstrate proficient more often than not. Additionally, modes fail to accurately represent student proficiency because they do not represent the full body of evidence. When a variety of scores are earned in a modal calculation, the least frequent scores become lost, which penalizes students if those scores were above their modal score.

Notably, some variants of Boolean logic gate systems can reveal student growth earlier than other methods. Especially in cases where students have wider distributions of scores, a student who struggles early on may not show even partial proficiency until late in a grading period. For example, our example student who chronologically earns a score set of 1, 1, 2, 2, 3, for a standard will never see their grade projected as higher than 1 if calculated using a modal or bi-modal method. This can be harmful to students who are improving yet do not see their projected grade increase. Within the Boolean logic gate example depicted in Figure 1, this student starts with the same scores, but will see their grade projected as a 1-2 split after the 4th assessment, and 2 after the fifth, encouraging growth and engagement. Finally, unlike all other methods analyzed, Boolean logic gates identify cases with split scores where professional judgment should be applied. Incorporating professional judgment into a calculation method allows for recognition that scores are not simply numbers, but rather are representations of learning (O'Connor, 2018).

As with any grading system, there are trade-offs to consider. For example, like logic rules, Boolean logic gates do not recognize recency of scores. Also, some variants of Boolean logic gates, such as the one depicted in Figure 1, can result in a specific scenario where, regardless of other scores, a student who has more 1s than 2s cannot earn a rating of proficient or advanced. A learner in such a situation could feel discouraged, and student growth might not be incentivized.

COMPARING APPROACHES

To more fully distinguish between the methods analyzed in this article, scores for four sample students are calculated in Table 3 below. Each student earned seven scores pertaining to a particular standard. Like the previous set of sample student scores, it can be assumed these are summative assessments of relatively comparable length and rigor and are listed in chronological order. Also, these scores are based on the same four-level proficiency scale.

Students C and D highlight the value of considering recency in grades. While both students earned the same number of scores of each level of proficiency, they earned those scores in different orders. Student D shows growth while Student C shows that they understood content early on, but their ability to demonstrate a standard diminished over time. Some methods, including Boolean logic gates, did not pick up on this distinction.

Students A and B highlight the value of using all data in grades. While both students demonstrate the same primary and secondary modal scores as well as logic rules scores, the remaining scores show they have quantitatively different understandings of course material. Additionally, these two students demonstrate that recency on its own may not be enough to evaluate a student's level of understanding. Because all assessment scores include potential error, a singular grade may be fallible (Marzano, 2010). In a course where competencies or skills within a standard do not vary greatly, the use of more recent evidence may be helpful. However, in a course where each of these pieces of evidence is a demonstration of content of varied difficulties, recency may not provide a sufficient picture.

Table 3

Comparison of Grade Calculation Methods for a Single Standard

Grading Method	Student A (4,3,2,2,4,3,2)	Student B (2,2,2,1,3,3,4)	Student C (3,2,3,1,1,2,1)	Student D (1,1,3,2,1,2,3)
Logic Rule (original)	2	2	1	1
Logic Rule (lenient)	2	2	1	1
Most Recent	2	4	1	3
Mean	2.9	2.4	1.9	1.9
Mean of Highest 3	3.7	3.3	2.6	2.6
Decaying Average	2.7	3.0	1.5	2.2
Power Law	2.4	2.8	1.2	2.3
Mode	2	2	1	1
Bi-modal Matrix	2	2	1	1
Boolean Logic Gates	3-4	2	2	2

Note. Calculations adhere to the methods described in each corresponding section. When a calculation resulted in a decimal value, the value was rounded off to the nearest tenth.

LIMITATIONS

The Boolean logic gate method is a novel conceptual idea that has not been empirically tested. Pilot testing and research studies need to be designed and implemented before widespread adoption could become feasible. Additionally, like many other SBG calculation methods, the Boolean logic gate method may not be compatible with typical student information systems and electronic gradebooks commonly used in schools today. Further, before widespread implementation could be possible, consideration would need to be given to teacher training, professional learning community calibration, and communicating this calculation method to students and families. Because Boolean logic gate systems can appear complex at first glance, it would be important to ensure all stakeholders understand the model.

CONCLUSION

As schools navigate the process of SBG grading reform, school-based educators need to select grade calculation methods to determine not only overall course grades, but also summary scores on each standard. Trusting the most recent score as the singular piece of evidence relies entirely on the validity of a single assessment. The power law approach generates a projection that diminishes transparency for stakeholders because of its complexity (Townsley & Wilcox, 2024). Means can penalize students for low scores (O'Connor, 2018) and risk perpetuating the illusion of precision perpetuated by traditional grading systems (Guskey, 2024). Approaches based on modes leave out evidence and can discourage students if they determine the remaining assessments will have little effect on their overall grade or when earning a lower grade becomes the more strategic choice.

The newly proposed Boolean logic gates method, on the other hand, draws on advantages of means and modes while avoiding their pitfalls. Like mean calculations, Boolean logic gate grades account for all available evidence. Like mode calculations, Boolean logic gate grades do not provide false precision of decimal values and do not require that students score higher than proficient in order to account for low scores earned early in their learning. Means show growth earlier by having every score contribute to the ongoing grade. Modes, in turn, can be narrow in scope, requiring significant amounts of data to reflect progress in student learning. The Boolean logic gate method strikes a balance between these, showing growth that may incentivize students earlier than modal calculations without trying to interpolate a value between levels of proficiency, as mean calculations do. For these reasons, the Boolean logic gates grade calculation method for individual standards is poised to offer an informative representation of student learning and shows strong potential as a useful tool for school-based educators implementing SBG systems.

School leaders who wish to explore the possibility of adopting Boolean logic gates into their school or district's grading practices could begin by leading a gradebook investigation. Such an undertaking would involve applying a variety of grading methods, including Boolean logic gates, to samples of prior gradebooks and leading a discussion with teachers to examine hypothetical impacts on student grades. Guiding teachers to examine prior gradebooks through the lenses of various grade calculation methods can evoke deeper discussions about the meaning of grades. Schools have complex decisions to make about grade calculation methods. Boolean logic gates, like any grade calculation method, introduce strengths and limitations. School and district administrators can provide valuable leadership when working collaboratively with educators to make context-specific decisions about the trade-offs that best align with schools' needs and priorities.

REFERENCES

Brookhart, S. M. (1991). Grading practices and validity. *Educational Measurement: Issues and Practice, 10*(1), 35–36. https://doi.org/10.1111/j.1745-3992.1991.tb00182.x

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

Brookhart, S. M., Guskey, T. R., McTighe, J., Wiliam, D. (2020). Eight essential principles for improving grades. *Educational Leadership, 78*(1), 10–17.

Dueck, M. (2011). How I broke my rule and learned to give retests. *Educational Leadership, 69*(3), 72–75.

Feldman, J. (2019). Beyond standards-based grading: Why equity must be part of grading reform. *Phi Delta Kappan*, *100*(8), 52-55. https://doi.org/10.1177/0031721719846890

Feldman, J. (2023). *Grading for equity: Why it matters, and how it can transform schools and classrooms* (2nd ed.). Corwin.

Guskey, T. R. (2024). Addressing inconsistencies in grading practices. *Phi Delta Kappan, 105*(8), 52–57.

Guskey, T. R., Frey, N., & Fisher, D. (2024). *Grading with integrity: A research-based approach grounded in honesty, transparency, accuracy, and equity.* Corwin.

Guskey, T. R., & Jung, L. A. (2012). The case against averaging grades. *Principal Leadership, 13*(6), 28–32.

Guskey, T. R., & Jung, L. A. (2016). Grading: Why you should trust your judgment. *Educational Leadership, 73*(7), 50–54.

Heflebower, T., Hoegh, J., & Warrick, P. (2014). *A school leader's guide to standards-based grading*. Solution Tree Press.

Heflebower, T., Hoegh, J., Warrick, P., & Flygare, J. (2018). *Teacher's guide to standards-based learning* (1st ed.). Marzano Resources.

Iamarino, D. L. (2014). The benefits of standards-based grading: A critical evaluation of modern grading practices. *Current Issues in Education*, *17*(2), 1-11.

Link, L., & Guskey, T. R. (2022). Is standards-based grading effective? *Theory into Practice, 61*(4), 406–417. https://doi.org/10.1080/00405841.2022.2107338

Marzano, R. J. (2006). *Classroom assessment and grading that work.* ASCD.

Marzano, R. J. (2010). *Formative assessment & standards-based grading*. Marzano Research.

O'Connor, K. (2018). *How to grade for learning: Linking grades to standards* (4th ed.). Corwin.

Peters, R., & Buckmiller, T. (2014). Our grades were broken: Overcoming barriers and challenges to implementing standards-based grading. *Journal of Educational Leadership in Action, 2*(2). https://doi.org/10.62608/2164-1102.1070

Reibel, A. R., Gobble, T., Onuscheck, M., & Twadell, E. (2025). *Pathways to proficiency: Implementing evidence-based grading* (2nd ed.). Solution Tree.

Schimmer, T. (2016). *Grading from the inside out: Bringing accuracy to student assessment through a standards-based mindset*. Solution Tree.

Townsley, M. (2019). Considering standards-based grading: Challenges for secondary school leaders. *Journal of School Administration Research and Development, 4(*1), 35–38. https://doi.org/10.32674/jsard.v4il.1941

Townsley, M. & Wear, N. (2020). *Making grades matter: Standards-based grading in a secondary PLC*. Solution Tree.

Townsley, M. & Wilcox, J. (2024). Standards-based grading: Core principles and current implementation debates. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 97*(3), 76–86, https://doi.org/10.1080/00098655.2024.2357615

SHAWN CARNEY, MEd, is a High School Science Teacher and serves as the AP Physics Course Captain in Fargo Public Schools. His major professional interests lie in the areas of standards-based grading, inquiry-based learning, and effective PLCs. Email: carneys@fargo.k12.nd.us

TERESA SHUME, PhD, is an Associate Professor and serves as the Director of Teacher Education in the School of Education at North Dakota State University. Her major research interests lie in the area of inclusive science teacher education. Email: teresa.shume@ndsu.edu

Manuscript submitted: ***September 7, 2025***
Manuscript revised: ***December 21, 2025***
Manuscript revised: ***July 2, 2026***
Accepted for publication: ***July 6, 2026***