

Rethinking “The 4” in Standards-Based Grading: Advanced Mastery, Standards Alignment, and Implications for K–12 Reporting

James Mattiace

Assessment and Competency Collaborative, Oregon, USA

ABSTRACT

Debates over the use of four-level scales in standards-based grading and reporting, particularly the meaning and necessity of the highest level, persist in K–12 education. This conceptual, practice-informed article examines whether “the 4,” commonly labeled *advanced mastery*, *exceeds*, or *extending*, is theoretically sound and instructionally useful within standards-aligned grading systems. Illustrative classroom scenarios demonstrate that students may produce exceptional work without exceeding the requirements of the standard being assessed. The article considers arguments for and against a top level that requires students to exceed the standard, and it distinguishes valuable qualities such as creativity, sophistication, and superior performance from evidence of standards attainment. Particular attention is given to the misapplication of Webb’s depth of knowledge framework as a proxy for mastery levels, clarifying the difference between cognitive complexity and quality of performance. The article recommends precise performance descriptors, moderation and calibration of student work, narrative feedback, student conferencing, and other approaches for recognizing excellence without distorting proficiency scales. It concludes that schools should initially use a three-level system focused on developing a shared understanding of proficiency. Additional levels may be considered later, following reflection, calibration, and analysis of implementation evidence.

Keywords: standards-based grading, advanced mastery, student achievement

A recurring question in standards-based grading and reporting reform is what to do with the top level, commonly represented by a 4 on a four-point progression. Should it exist at all? If it does, what exactly does it mean? When should it be conferred? And if a school still converts standards-based marks into letter or number grades, however ill-advised that may be, what should a 4 represent?

One commonly used standards-based grading structure uses labels looking something like the scale shown below:

4: ****

3: Meeting/Achieving

2: Progressing

1: Beginning

For the sake of argument, the label for the 4 is intentionally left blank. Schools label and use this level in many ways, so it is worth asking whether it is necessary before assuming what it should mean.

The central distinction this article explores is not between three-level and four-level scales. It is between two different uses of the four-level scale. In some systems, Level 4 is simply the point at which the standard has been met. In others, proficiency is represented by a 3, and the 4 is intended to mean something beyond the standard: *advanced mastery*, *exceeds*, *extends*, or some other version of *more than proficient* (Table 1).

In Table 1 below, the concern is the second scale usage: proficiency at Level 3. The problem is not the number 4 itself; it is what the grading system claims the fourth level represents. In such a scale usage, stakeholder concerns are common. Teachers worry about fairness, parents worry about recognition, and administrators worry about consistency. A grading system designed to simplify communication about learning can therefore become complicated at precisely the moment it should be clearest. This tension can show up in calibration meetings, report-card conferences, and policy revisions, where educators are asked to explain why a student who did something genuinely impressive still “only” met the standard.

Table 1***Two Common Uses of a Four-Level Scale***

Use of the Scale	Level 3 Label	Level 4 Label
Proficiency at Level 4	Approaching/Progressing	Meeting/Achieving
Proficiency at Level 3	Meeting/Achieving	Exceeds/Extending/Advanced Mastery

Scholars of grading have long argued that standards-based systems are, by design, criterion-referenced: Students are evaluated against defined learning standards rather than compared with one another (Brookhart et al., 2016). From this perspective, proficiency represents successful attainment of the standard itself. As Guskey (2014) argues, labels that imply performance beyond the standard—such as “exceeds”—introduce conceptual confusion by suggesting that the target is fixed and surpassable at the same time.

This article offers a conceptual, practice-informed analysis of that problem. It does not present original empirical findings or attempt a comprehensive review of the grading literature. Instead, it uses existing scholarship and illustrative K–12 classroom examples to examine whether a level beyond proficiency can be coherently defined within a criterion-referenced grading system and how schools might recognize exceptional work without distorting the purpose of standards-based reporting. The central argument is straightforward: Excellence is worth recognizing, but a standards-based grade may not always be the right place to encode it.

WHAT DOES IT MEAN TO EXCEED A STANDARD?

The first dilemma concerns what exactly we are measuring (Brookhart et al., 2016). Consider three real-life situations shown in Table 2, slightly truncated for clarity. In which of these situations has the student achieved *beyond* the expectation or at an *advanced mastery level*?

Table 2***Sample Student Responses to Standards***

Standard	Student Performance Description
<i>CCSS.ELA-LITERACY.W.6.1</i> Write arguments to support claims with clear reasons and relevant evidence. (National Governors Association Center for Best Practices & Council of Chief State School Officers [NGA & CCSSO], 2010)	Grade 6 student writing a persuasive essay on animal welfare used arguments dating back to Plutarch and a 1613 painting by Brueghel to make the case that this is not a recent trend or phenomenon. The student also cited multiple religious texts to supplement the scientific argument.
<i>CCSS.ELA-LITERACY.L.9-10.2</i> Demonstrate command of the conventions of standard English capitalization, punctuation, and spelling when writing. (NGA & CCSSO, 2010)	Grade 10 student writes flawlessly, and their papers never contain errors in conventions.
<i>NGSS 3-ESS2-1 Earth's Systems</i> Represent data in tables and graphical displays to describe typical weather conditions expected during a particular season. (NGA & CCSSO, 2010)	Grade 3 student figured out how to use Google Sheets to create charts and graphs in an exploration of monthly average precipitation in dry and rainy seasons.

Below, consider educator arguments for and against a top level that requires students to exceed the standard, along with a parent argument.

The argument for: In each case, the student has demonstrated a high level of skill and intelligence that appears to go beyond the normal grade level. Since the standard is written for a grade level, the student has gone beyond that grade level and should be recognized as such. A second argument is that the student did something no one expected and should be recognized for it.

The argument against: The demonstration of higher intelligence is not part of the standard (Sadler, 1989). The question is whether the student meets the expectation. If the answer is yes, the standard has been met. Creativity, fluency, or sophistication may make the work exceptional, but those qualities are not automatically part of the criterion being assessed.

The parent's argument: "My child is a genius and needs to be celebrated!" This concern should not be dismissed. Recognition matters. The question is whether a proficiency score is the best instrument for providing that recognition.

Now, here are some commonly used definitions of when a student has gone beyond the expectations:

- They can teach it to others and have it make sense to them.
- They have done something unique and innovative that was not assisted by an adult.
- They made a contribution that furthered the teacher's own evaluation of the expectation (e.g., they "set the curve").

At the core of standards-based grading is the idea that standards define the learning expectations that students are responsible for meeting. In a criterion-referenced system, the judgment is independent of how other students perform (Brookhart et al., 2016). A veteran teacher may occasionally encounter a student whose work is unusually impressive, but the evaluation still must return to the evidence and the criteria. Otherwise, *advanced* risks becoming a judgment about how remarkable the student appears compared to other students rather than how the student performed against the standard.

The second dilemma is whether a category labeled *advanced mastery* or *extending* can exist at all (Sadler, 1989). It is like the old joke that one cannot be "a little bit pregnant." In this context, you either meet the standard or you don't. However, this does not mean performance against a standard cannot vary in quality. Rubrics can and do distinguish degrees of quality within defined criteria. The problem arises when the highest category requires students to demonstrate something outside the standard itself. After all, no standard says, "Student can solve quadratic equations and teach it to others," or "Student will demonstrate correct subject-verb agreement in such a way that it makes the teacher reconsider whether they are even needed in the classroom anymore."

A familiar bicycle analogy can further illuminate this problem, as shown in Table 3. In this case, riding independently represents meeting the standard (Level 3). To reach Level 4, the rider pops a wheelie, jumps ramps, and performs stunts. It is memorable and easy to explain to families. However, it also changes the task. The student is no longer simply demonstrating the original standard at a higher quality; the student is doing something additional.

Table 3

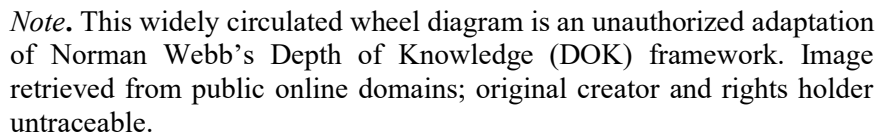
The Bicycle Analogy: Rubric Criteria and the "Extending" Problem

Level	Label	Descriptor
4	Extending	You can ride a bike by yourself and also do jumps, wheelies, and other stunts.
3	Achieving	You can ride a bike by yourself with no assistance.
2	Approaching	You can pedal and stay upright, but you need assistance.
1	Beginning	You have not mastered the basics, such as pedaling, turning, and balancing.

This analogy has been adapted to swimming, hiking, baking, and even swinging ("I can swing fast and high") to better explain the *extending* level of a four-point grading scale. These versions make the progression intuitive, but they are less useful when evaluating conventions in English or problem-solving in mathematics. An IB rubric for Spanish Language A, for example, can distinguish between ideas being "well organized" and "adequately organized" (International Baccalaureate, 2023). That is a difference in quality against defined criteria. It is not necessarily the same thing as going beyond the standard.

The confusion is reinforced by the widely circulated “DOK Wheel,” an educator-created visualization based on Webb’s work (Figure 1). Webb and colleagues have distanced the framework from this wheel because assigning verbs to fixed DOK levels can create precisely the kind of distortion at issue here (Webb et al., 2023).

The Misleading “DOK Wheel”



If the standard being assessed uses a DOK level 4 verb, and a student meets it and then demonstrates additional, deeper understanding, that could support awarding advanced mastery. Similarly, a standard requiring a DOK level 2 verb like “classify” could justify recognizing an additional advanced mastery if a student not only classifies but also proposes new or synthetic groupings based on other perceived similarities (Brookhart, 2010).

DOES “EXCEEDS” SERVE A USEFUL PURPOSE?

A classic argument for the top level of a scale allocated to achievement beyond the standard (however named) is the need to recognize excellence and maintain motivation for high-performing students. Student conferencing is a useful tool in a grading system that recognizes advanced mastery. Through direct discussions, teachers can assess whether a student's performance goes beyond the standards. This method is more effective when clear grade descriptors are established, allowing students to demonstrate how they meet specific criteria independently and consistently.

However, showcasing student excellence doesn't require adding an additional performance level above the standard. Instead, teachers can directly communicate outstanding achievements to parents through emails, notes, or postcards home. Additionally, recognizing students in award assemblies and through external competitions like math or debate contests can further highlight their talents. Teachers can also document these accomplishments in a growing portfolio that students can use in the future (Guskey & Bailey, 2010).

In high school, recognition matters for university applications. While many student applications feature high GPAs and impressive titles, schools can also provide a school profile that highlights a school's unique grading system, such as valuing *meeting the standard* as the top achievement. As many schools move away from class ranking (Strauss, 2016), this approach becomes increasingly relevant.

Finally, the importance of report card comments should be emphasized. These comments, often a challenge for teachers and parents, can play a crucial role in university applications, offering a personalized showcase of a student's excellence. When used effectively, these comments can convey achievements and abilities in a meaningful way, adding depth to the student's academic profile.

WHAT COULD BE DONE INSTEAD OF “EXCEEDS”?

Recognizing excellence does not require pretending that exceptional work never happened. It requires choosing the right tool for the message. As mentioned earlier, schools have a variety of methods to recognize high achievement, many of which are often more meaningful and personal to the student and their families than a simple designation of extending, exceeds, or advanced mastery.

But there is also a simple option that often gets lost in the debate: Retain four levels, but make Level 4 the point at which the standard is met. A scale might start at beginning and move to progressing, approaching, and meeting. This preserves four reporting categories without requiring the system to define a mysterious space beyond proficiency. When this option is chosen, the design question becomes not whether a school uses three levels or four, but whether the scale clearly communicates each level it claims to measure.

To ensure consistent expectations for each level of the scale, mentoring new teachers is crucial. All teachers, including veterans, should collaborate to review student work objectively through the process of moderation. This process, involving the entire faculty, specific departments, or peer mentorship, helps in identifying true excellence in student performance (Guskey, 2017).

COUNTERARGUMENTS AND RECOMMENDATIONS

Arguments for the *extending*, *exceeds*, or *advanced mastery* level generally emphasize recognition, alignment with GPA conversion systems, or its usefulness as a reward. These arguments are not trivial. Recognition matters, conventional GPA systems may create practical constraints, and schools reasonably want high-performing students to remain challenged. The question is whether those concerns require a performance level beyond the standard or whether they can be addressed through other parts of the reporting system.

This article strongly recommends initially implementing a standards-based learning, grading, and reporting system that includes only criteria for meeting the standard, without additional levels for exceeding it. Although it might not be immediately popular, this approach is crucial to ensuring faculty, students, and parents fully understand what meeting the standard entails. Implementing this system without additional levels at the start prevents the complication of having to defend decisions as to why a student is achieving at Level 3 rather than Level 4.

This foundational period allows everyone involved to gain a clear and solid understanding of the system's expectations. After an initial implementation period (approximately 1–2 years), once there is collective recognition of what truly constitutes performance that goes above and beyond the standard, and how frequently it occurs, the discussion can shift to whether and how to integrate a higher level into the system. This phased approach helps in gradually acclimating all stakeholders to the system and in making informed decisions regarding *what to do about that 4*.

CONCLUSION

At first glance, the debate about “the 4” can look like a technical question about scales, but it is really a question about the purpose of grading. If a standards-based grade is intended to communicate whether a student has met a defined expectation, then every level in a grading scale must have an equally clear meaning. Schools should absolutely recognize creativity, sophistication, independence, and exceptional achievement. The question is not whether those qualities matter. The question is whether assigning a higher number is the clearest and most defensible way to communicate these accomplishments. Before adding a level beyond proficiency, schools should be able to explain exactly what it measures, why it belongs in the grading scale, and how educators will identify it consistently. If they cannot, meeting the standard may be enough.

REFERENCES

- Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. ASCD.
- 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>
- Guskey, T. R. (2014, October 17). Why the label “exceeds standard” doesn’t work. *Education Week*. <https://www.edweek.org/education/opinion-why-the-label-exceeds-standard-doesnt-work/2014/10>
- Guskey, T. R. (2017, January 31). New direction in the development of rubrics. *Corwin Connect*. <https://corwin-connect.com/2017/02/new-direction-development-rubrics/>
- Guskey, T. R., & Bailey, J. M. (2010). *Developing standards-based report cards*. Corwin.
- Hess, K. (2013). *Applying Webb’s Depth of Knowledge levels in classroom instruction*. National Center for Assessment. <https://www.karin-hess.com/cognitive-rigor-and-dok>
- International Baccalaureate. (2023, December 19). *Assessment & exams*. <https://www.ibo.org/programmes/diploma-programme/assessment-and-exams/>
- Level Questions. (n.d.). *Depth of Knowledge (DOK)*. <https://levelquestions.weebly.com/depth-of-knowledge-dok.html>
- Lexington School District Two. (n.d.). *Proficiency scales*. <https://www.lex2.org/page/proficiency-scales-elementary>
- National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts & literacy in history/social studies, science, and technical subjects*. <https://thecorestandards.org/ELA-Literacy/>
- Next Generation Science Standards. (n.d.). MS-LS4-3: *Biological evolution: Unity and diversity*. <https://www.nextgenscience.org/pe/ms-ls4-3-biological-evolution-unity-and-diversity>
- Next Generation Science Standards. (n.d.). *ESS-1 Earth’s Place in the Universe*. <https://www.nextgenscience.org/overview-topics>
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119–144.
- Strauss, V. (2016, December 13). The case for abolishing class rank. *The Washington Post*. <https://www.washingtonpost.com/news/answer-sheet/wp/2016/12/13/the-case-for-abolishing-class-rank/>
- Webb, N. L. (2002). *Depth-of-knowledge levels for four content areas*. Wisconsin Center for Educational Research.
- Webb, N., Christopherson, S., & Morelan, B. (2023, March 27). *An inside look at Webb’s Depth of Knowledge*. Edutopia. <https://www.edutopia.org/article/how-use-norman-webb-depth-of-knowledge>

JAMES MATTIACE is the Executive Director of the Assessment and Competency Collaborative, which focuses on certifying and accrediting schools and divisions that are aligned with standards-based grading and reporting best practices. Email: james@assessmentcollaborative.org

Manuscript submitted: **September 6, 2025**

Manuscript revised: **January 31, 2026**

Manuscript revised: **May 28, 2026**

Manuscript revised: **August 28, 2026**

Accepted for publication: **September 3, 2026**