

Changes in K-12 Grading Research Since 2016

Susan M. Brookhart
Duquesne University, USA

ABSTRACT

This essay explores additions to the grading research literature observed in the decade since the large review “A Century of Grading Research: Meaning and Value in the Most Common Educational Measure” (Brookhart et al., 2016). Sections discuss studies of traditional grading, the effects of the pandemic on grading, studies of standards-based grading, and studies about rubrics. A final section discusses future directions for grading research.

Keywords: grading, standards-based grading, classroom assessment

In 2016, a group of colleagues and I published a review of the grading research literature (Brookhart et al., 2016), which at that point included over a century of study. For this special issue of JSARD on current and future grading practices in K-12 schools, I have been asked to comment on what I have observed in the field since that paper was published and where the field might be going in the future. I make no claim that this paper presents a systematic review of all the grading research in the last decade, which is beyond the scope of this commentary. However, I am happy to take up the two-fold challenge that was put to me and present thoughts on grading research since 2016 and possible future directions.

GRADING RESEARCH SINCE 2016

The 2016 paper included five sections: (a) early studies of the reliability of grades, (b) quantitative studies of the composition of K–12 report card grades, (c) survey and interview studies of teachers’ perceptions of grades, (d) studies of standards-based grading (SBG), and (e) grading in higher education. Leaving aside the early historical studies and the higher education studies (because of the K-12 focus of this issue), we can identify K-12 empirical studies in the other three categories published since 2016: studies of traditional grading (b and c above), the effects of the pandemic on grading (c and d), and studies of SBG (d).

Recent Research on Traditional Grading

Recent (post-2016) research on traditional grading seems to me to extend what has gone before, either through increased methodological sophistication or through increased theoretical grounding. Two examples of sophisticated statistical modeling include Doz (2023) and Klapp (2017).

Doz (2023) used a large national dataset from Italy and found that teacher-assigned mathematics grades, for students in their final year of high school, were related to students’ gender, socio-economic status (SES), and scores on a national exam and to school type and location. He documented some findings already reported in the literature, such as a gender gap and an effect of SES, as well as school-type and geographical differences that had not been studied extensively in Italy.

Klapp (2017) took advantage of a natural experiment that was possible because, for a while, Swedish municipalities could decide whether to use grades in sixth grade. Thus, she could compare the 7th-grade GPA of students who had been graded in 6th grade with those who had not (for whom 7th grade was the first year they received grades). She found a negative

effect of grading for low-ability students on 7th-grade GPA, which was fully mediated (that is, the effect disappeared) through academic self-concept in math and Swedish and motivation to improve. She interpreted this to mean that low-ability students with prior experience with grading believed they were poor at math and Swedish and needed to improve, and that they received lower subsequent grades than low-ability students without prior experience with grading.

Studies designed in terms of theory and motivation around learning and engagement, for example, self-regulation and self-confidence, strengthened the theoretical grounding of research on grading, moving from the simple atheoretical studies of yesteryear. One example is the Klapp (2017) study just reported. This trend toward firmer theoretical grounding is also evident in interview and survey studies.

An interview study was conducted in Norway and Sweden with upper-secondary students; despite some differences in grading policies and procedures, grades are high-stakes for students in both countries (Liljeröd et al., 2025) and are an important determinant of admission to higher education. All participants thought grades were important, and most said they studied more because they knew they were being graded. Students saw grades as a public indicator of their competence and, sometimes, a distraction from learning. Good grades provoked positive emotions that enabled students to study harder, and grades also prompted strategic decisions about studying, including sometimes deciding they didn't need to study more. Most students who perceived their grades as low reported negative emotions, reduced willingness to study, and a negative effect on their academic self-concept. A few students reported feeling activated by low grades, taking them as a challenge to show they could do better.

Chen and Bonner (2023) surveyed pre-service and in-service teachers about using student classroom conduct in assigning grades, referencing the learning theories teachers drew from. A number of respondents did support conduct-based grading, despite its being contrary to current recommended practices. Secondary content-area teachers were more likely than elementary teachers to endorse this view. Different teachers held different views, however, and many endorsed achievement-only grading. Teachers who endorsed mixing achievement and conduct in grades pointed out that student conduct affects the learning environment for everyone in the classroom, and they were more likely to endorse traditional or didactic teaching approaches and behaviorist approaches to classroom management.

In my view, studies that add to the theoretical foundation for understanding grading practices are expanding the field in an important way. Of course, there are still some studies of teachers' grading beliefs and practices that situate their investigations in the context of prior research instead of learning theory. Thus, they replicate or extend findings from previous research with other and more contemporary populations. These studies continue to find similar results as in the past (Brookhart et al., 2016). Guskey and Link (2019) found differences in grading practices by grade level; for example, secondary teachers put more weight on formal exams and assignments than do elementary teachers. Kunnath (2017) found that secondary teachers identified concerns about external perceptions, desire to facilitate student success, and administrator pressure for assigning low grades as drivers of their grading practices and mixed achievement and non-achievement factors, especially student ability and effort, in assigning grades. Some survey studies of grading practices replicate or extend previous findings about grading, updated with more recent samples or samples from different countries. For example, Arsyad Arrafii (2020) found that Indonesian secondary school teachers consider student achievement as well as effort and external factors in grading. Moreover, they reported feeling some pressure to inflate grades.

The Pandemic and Research on Grading

Because of the COVID-19 pandemic, most K-12 education was online, at least for a while, in 2020 and 2021. Some students, especially high school students, began to worry about their grades and their futures (Ashta et al., 2023). State departments of education published guidance for grading during emergency remote learning (Townesley & Kunnath, 2022). Troubles in traditional grading practices, many of which SBG had been designed to address, became more obvious in the online environment, like over-reliance on tests and worksheets and inflexible policies that measured work habits as much as learning. For example, some teachers experienced problems grading participation when students had trouble logging in or turned off their video cameras. Tests and quizzes were no longer guaranteed to be done by students individually, without help. Worksheets and other assignments lost context, as teachers could no longer be sure what instructional experiences had actually occurred for each student. Districts experimented with various stopgap grading policies, most of which were not much help.

The pandemic affected grading systems in countries other than the US, as well. For example, in Norway, secondary school students' grades are typically composed of 80% teacher assessments and 20% central exams, and grades determine admission to higher education. When the exams were canceled due to the pandemic, students' final grades rose, which prompted questions about fairness and justice in college admission and both positive and negative consequences for students

(Gamlem et al., 2023). I have no doubt that a search for research about grading during the pandemic in other countries would illustrate, as this does, that the effects of the pandemic interacted with the grading system in place.

At the time of this writing, the COVID-19 pandemic is over. Based on the US experience, Townsley (2020) recommended three principles for pandemic-era learning that are also core SBG principles: base academic grades on prioritized course objectives or standards, report non-cognitive and behavioral conduct separately, and design grading systems that prioritize what students have learned over when they have learned it. Research on SBG principles continues to proliferate. In the next two sections, I provide an overview of some of the advancements in research on SBG implementation and practice and on school leadership for SBG, respectively.

Implementation in SBG

What stands out as new in the K-12 grading research is the proliferation of empirical studies of SBG since 2016. In 2016, we claimed (p. 828), “Literature on SBG implementation recommendations is extensive, but empirical studies are few,” and we discussed only seven empirical studies dated from 2003 to 2014. That is certainly not the case today.

When research begins in new topic areas in education, it often begins with descriptive studies investigating practices—what is done—and perceptions—how people think and feel about what is done. In the time since the 2016 grading review article, researchers have begun to publish descriptions of SBG practice and teacher or student beliefs about SBG. Many of these studies used small samples in one or two contexts. Others used larger sample sizes but, because of that, lost the ability to describe the context. The results are mixed, and this could simply be the result of the context-bound nature of this research, with its attendant variation in school and community readiness for SBG, teacher knowledge and skill with SBG practices, and cultural or micro-cultural differences between contexts.

Descriptive Studies of SBG Implementation and Beliefs

The following studies paint a picture of SBG as it emerges into the mainstream. Unlike traditional grading practices where everyone “just knows” what an A means, teachers and students alike are learning about SBG, and halting progress—slow and sometimes hesitant—is being documented. The following five studies provide evidence of this halting progress. I include sample sizes to show that most of them are small-sample, context-bound investigations. While I think this is a good thing, because grading practices like other educational practices are situated in school and community cultures and micro-cultures, it does mean that no sweeping generalizations about SBG reform can yet be made.

Knight and Cooper (2019) studied seven teachers from five high schools with varying levels of SBG in place. They described positive effects for instruction, assessment, the classroom environment, and students’ approaches to learning. However, all the teachers needed to convert their standards-based grades to percentages and letter grades for reporting, which essentially meant that what was studied was a standards-based approach to traditional grading. Even so, Knight and Cooper found evidence for an implementation dip, where newly adopted practices get more confusing before they become clear and routine.

Olsen and Buchanan (2019) studied 15 teachers in two high schools who were undergoing a year-long professional development on grading reform. The goal of the professional development was to reform grading practices that included behavior, effort, and motivation in the same academic grade by using separate academic grades. One of the two high schools had already moved to a 1-4 scale, away from percentage grading. Teachers had little prior professional learning on grading and expressed varied and unclear purposes for grading. They also expressed differences in philosophies about the purpose of education in general and, therefore, what elements of student performance they believed were gradable. In these respects, teachers were influenced by their own backgrounds and experiences and by school characteristics. Teachers did make some changes, but they were only partial and sometimes reverted to the grading approach the program was trying to reform.

Ketsman and Reeves (2023) surveyed 136 K-12 teachers in one district participating in SBG reform. They asked questions about SBG practices, teacher perceptions of those practices, perceptions of barriers or challenges to implementation, and teachers’ prior opportunities to learn about SBG during pre- and in-service education. Overall, teachers reported mixed levels of implementation of various SBG practices and neutral perceptions of SBG. They tended to agree that SBG provided better information about students’ strengths and weaknesses. Barriers identified included time, incompatibility with current grading systems, and parent/guardian perceptions. Teacher confidence partially mediated the relationship between pre-service opportunity to learn and frequency of use of SBG practices, and teacher confidence fully mediated the relationship between in-service opportunity to learn and frequency of use of SBG practices.

Also regarding pre-service preparation, Battistone et al. (2019) investigated whether 11 new teachers were prepared to work in a school that uses SBG. They found that pre-service training programs and field experience were inconsistent in their assessment training.

In one of the larger-scale studies, Wisch et al. (2018) surveyed 429 secondary teachers from 17 schools on their beliefs about SBG practices, especially redo/retake policies. They found that whether teachers allow redos and retakes depended on school policies, content area, and teachers' personal beliefs about the importance of deadlines and student accountability. Many teachers used grading practices that fell somewhere between SBG and traditional grading, although only one of the 17 schools had an explicit SBG policy, and even that school needed to transform its grades to the 100-point scale for reporting. Forty-one percent of the teachers indicated their school had an SBG policy, and 98% said their school allowed redos and retakes. Only 8% of teachers did not allow redos or retakes of any sort; the rest had varying policies but allowed at least some revisions.

The following studies, in music education and physical education, respectively, demonstrate that research on SBG implementation is also arising within content areas. "Halting progress" also seems to describe these efforts.

St. Pierre and Wuttke (2017) conducted a survey of SBG use among music educators in one state. Among the 96 volunteer respondents, just over half reported they were not familiar with SBG. Almost half (47%) reported that they were using SBG at least partially, and they defined what they meant by SBG. After removing some participants whose definitions did not reflect understanding of SBG, the researchers determined that 40% were actually using the practice. The most common reason teachers gave for using SBG was that they were required to, and the most common reason teachers gave for not using SBG was lack of knowledge about it.

Townsley and McNamara (2021) studied implementation of SBG by four physical education teachers and two administrators at one school. They found that while teachers were somewhat confused about what exactly they were supposed to do to implement SBG, the work did result in teachers' desire to more fully understand their state's physical education standards (Townsley & McNamara, 2021). Townsley and McNamara (2024) found that preservice teachers participating in a one-hour SBG workshop helped students appreciate the importance of physical education and grading on defined standards but needed more support before they began implementation in their own classrooms.

The following two studies, one small and one larger but still within one school, show how high school students who have become used to traditional grading are concerned—perhaps with a little of the flavor of "fear of the unknown"—about how SBG will affect their immediate needs, both social and educational.

Guskey et al. (2020) studied students' beliefs about the impact of SBG in high school on their transition to college or university. They used surveys to ascertain the degree of SBG implementation in high school for 13 students, followed by interviews with 7 of those students. They found that high schools were not implementing SBG with fidelity, but to the degree they did implement SBG, students did not perceive these practices had caused problems for their college performance. Concerns about SBG and the transition to college, which some parents voice, may be misplaced; for the students in this study, the social and time management aspects of their transition to college were their biggest concerns.

Peters et al. (2017) studied students' perceptions of SBG at one high school. The school had implemented SBG policies and had worked to help students and parents understand them. However, there was still some "disconnect" (Peters et al., p. 14) for parents and students. They used fall survey data from 376 students and spring surveys from 230 students, out of 478 enrolled in the school. Students had implementation concerns, as not all teachers were consistent in their implementation of the policy. Some students were concerned about the 1-4 scale, worried their grades would be lower, and had fears about their preparation for higher education. Less common were concerns about student complaints, stress, the appearance of being "stupid" if one needed to reassess, and motivation. The authors point out that resistance to change is unsurprising and that specific student concerns can be a source of feedback for the school as it moves forward with SBG implementation.

Finally, the following study gives evidence that communication about SBG, important for any school reform, is important for grading reform, as well.

Larsen (2023) asked 10 teachers, 4 parents, and 2 high school principals to discuss their beliefs about SBG. Ten of the 12 educators believed that SBG improved student learning, even though all educators agreed that letter grading was easier to do. Some parents expressed dismay that homework and other practice work was not recorded, a finding that can be addressed with education and communication.

What are we to conclude from these studies of SBG implementation? The evidence offered in these studies suggests that learning and implementing new grading practices is gradual and influenced by the attitudes and beliefs of the educators involved. As we look at these studies of "halting progress," it is possible to be a pessimist, emphasizing the "haltingly" slow and hesitant nature of changes, mixed results, and the fact that SBG reform is not yet implemented fully and with fidelity to basic SBG principles. It is also possible to be an optimist, emphasizing the "progress" and the fact that over the last

decade these reforms have been pursued and are beginning to be studied, with the intention that future implementation can benefit from understanding the barriers and uncertain beginnings this research has brought to light. A question one might ask is whether SBG reforms are worth it: Is there evidence that SBG is effective? The next section looks at that question, using research that has been published since the 2016 review.

Investigations of the Effectiveness of SBG

For new topics in educational research, efficacy studies often follow descriptive studies when the field has developed to a point where an intervention can be defined clearly enough to evaluate. A few such studies have begun to appear for SBG, investigating the effectiveness of SBG practices on clarity of communication, math anxiety, achievement and homework completion, and achievement and motivation, respectively. Unlike the mixed results of the descriptive studies, these studies all reported positive effects of some kind.

Link and Guskey (2022) reviewed literature to answer the question “Is SBG effective?” They defined SBG as adherence to three principles: (1) report student performance on standards rather than as a single content-area grade; (2) use a limited number of reporting categories; and (3) report academic grades separately from non-cognitive and behavioral evaluations. Interestingly, they did not make multiple opportunities to show mastery, or redos and retakes, one of the defining categories of SBG. Their review supported the conclusion that SBG was an effective communication tool because applying these three principles clarifies the meaning of grades. They did not address effects on students’ learning and performance or behavior.

Fergus and Smith (2022) investigated how SBG practices affected five students in two middle-level mathematics classes. These students had self-identified as having math anxiety. Using a case study method, the researchers found that assessment was the characteristic of proficiency-based learning or SBG most closely associated with reducing students’ math anxiety. Specifically, establishing clear learning objectives and providing effective feedback were two SBG practices students reported as being helpful.

Huey et al. (2022) studied the two-year effects of removing practice work from grade computations in a geometry course on student performance and behaviors. The students all had the same teacher. The students were primarily high-achieving, and changes in performance due to reassessment were negligible. Two cohorts were compared: for Cohort 1, homework completion counted for 15% of the final grade, and the rest was based on exam scores; for Cohort 2, the final grade was based solely on exam scores. Cohort 1 had higher exam scores and greater homework completion rates. Some students reported seeing homework as optional if it didn’t count toward their grade.

Kramer et al. (2024) studied the impact of an SBG program called PARLO (Proficiency-based Assessment and Reassessment of Learning Outcomes) in 9th grade mathematics. This study was a large-scale ($n = 2,736$) randomized controlled trial, the only such design I have come across so far. Interestingly (to me, anyway), PARLO principles included basing student final grades on the number of learning outcomes for which students achieved proficiency or better, providing opportunities for reassessment for full credit after further learning, using clearly stated learning goals, and providing students with formative feedback. Because the latter two (learning goals and feedback, which are the formative aspects of SBG) were beginning to be widespread and therefore present in both control and treatment (PARLO) schools, the study really evaluated the effects of basing grades on proficiency and using reassessment. The PARLO program (that is, proficiency grades with reassessment) improved student achievement on end-of-course tests. Students with higher expectancies for and value of their mathematics learning benefited more.

One might wonder about these largely positive studies of effects of SBG reform on communication, achievement, motivation, and behavior, given the mixed results or “halting progress” reported in the descriptive studies in the previous section. Is it publication bias? Or is it, perhaps, that the clarity of learning goals and focus on proficiency—by both teachers and students—is so powerful that it’s actually helpful even when diluted by less-than-full implementation? I don’t think we have enough evidence in the studies reported here to draw a conclusion, but it’s worth thinking about.

Two additional studies are reported below, one investigating the effects of SBG on predicting ACT scores and the other on the long-term economic effects of the alternative grading practice of minimum grading. Both are distal, rather than proximal, effects of SBG grading reform, but interesting questions to ponder. The results were not quite as positive as the studies of proximal effects reported just above.

Townsley and Varga (2018) compared students’ GPAs from two demographically matched high schools that practiced traditional grading and SBG, respectively. They also compared the ACT scores of these two groups and investigated whether GPA and grading method predicted ACT scores. GPAs did not differ between schools, but students from the traditional-grading school had higher ACT scores (English, Math, and composite). Regarding prediction, English or Math GPA was

the best predictor of ACT English or Math score, respectively; males scored higher than females in each, and students from the traditional-grading school scored higher than those from the SBG school.

Morris and Johnson (2024) simulated the economic impact of implementing a statewide minimum grading policy in ninth grade in Arkansas. Minimum grading is an alternative method used with percentage-based grading that assigns a lower limit to low grades, typically 50 instead of 0, to mitigate the effects of failing grades on grades assigned on the 0-100 scale. The simulation investigated how minimum grading, a no-cost intervention, might increase school effectiveness (by reducing expenses required to teach students who don't graduate), increase the cost-effectiveness of public-school spending, and raise the return on investment in terms of students' lifetime earnings divided by the cost of educating them. They estimated that a 50% minimum grading policy would affect 1 to 3 percent of students (pass at least one additional course), increase educational cost-effectiveness by 6 to 23%, and have no impact on return on investment in terms of students' future earnings.

Leadership in SBG

There is a small group of studies of principals and other educational leaders' perceptions of SBG. These studies are descriptive, similar in tone to the descriptive studies reported above but describing the thoughts and actions of educational leaders instead of teachers. They were all conducted in the same state and by the same team of colleagues, with one exception: a study done with school principals in China (Cheng et al., 2023). These studies in one state do not generalize to the rest of the US; however, I report them here because of this journal's focus on administration and because leadership is one of the classic, well-established, and well-researched foundations that can make or break educational reform. If SBG practices are to reform grading and have the desired effects on students' beliefs, achievement, and motivation, educational leaders will need to know how to proceed.

Buckmiller et al. (2020) surveyed principals of rural high schools in one state. They overwhelmingly reported envisioning implementing SBG practices in the next 5 years, indicated a good understanding of the major ideas behind SBG grading reform, and believed they had the resources to implement. Slightly less, but still a majority, agreed that their teachers would have the disposition to engage in grading reform and that their upper-level administrators and school boards would support SBG reform. They also expected some resistance from parents and community members. Using data from the same survey, Townsley et al. (2019) predicted a second wave of SBG implementation would follow the early adopters or first wave of SBG (Peters & Buckmiller, 2014) in the state, and that principals felt teacher resistance to be the most common barrier. In addition, they believed that SBG grading reform would require time and planning, teacher professional development, and community education.

Townsley and Knight (2020) studied SBG reform in one high school. Their case study described implementation, barriers—including mixed levels of buy-in and depth of understanding of SBG—and the amount of effort required to produce proficiency scales that are well calibrated for cross-classroom use but meaningful enough to use within classrooms.

Research on Rubrics

Another area of study in which empirical literature has expanded greatly in the last decade is rubrics. Research in this topic area is past the stage of purely descriptive studies and has developed to the point that some meta-analyses are beginning to be published. These studies show that, overall, the use of rubrics has positive effects on student achievement and self-regulation of learning (Camargo Salamanca et al., 2024; Panadero et al., 2023). Research on rubrics is not usually linked with research on SBG, but it seems to me that a case can be made.

One important change in contemporary grading practices involves the use of proficiency or mastery scales instead of points and percentages. These scales typically have labels like "Advanced, Proficient, Approaching, and Beginning," or any one of several similar level names. These labels support performance level descriptions used for both formative assessment and grading; in other words, many SBG initiatives rely on rubrics.

To be clear, rubrics for assignments or sets of assignments (e.g., narrative writing, math problem solving) are the kind of rubrics that have been studied in the rubric research literature. Ketsman and Reeves (2023) point out the difference between scoring rubrics for assignments and proficiency scales for evaluating a body of evidence. While SBG does use proficiency scales designed for evaluating a body of evidence, it also quite commonly uses rubrics for assignments coordinated with those proficiency scales.

So far, I have not seen anyone try to combine the SBG and rubric literatures, but it seems obvious to me that the descriptions of proficiency on standards matter, and the research on rubrics is relevant to grading and especially to SBG, which rests on identifying and communicating clear standards. I suggest that readers interested in contemporary grading research should also be encouraged to read contemporary rubrics research.

As an example of a study that is relevant to both topic areas, rubrics and grading, I'll describe a study by Quinn (2020), who, through the use of the internet, was able to design a randomized experiment involving over 1,500 teachers. He used this study to investigate whether well-designed grading practices might help with bias in grading. Student writing samples were created that differed only by cues about the author, which were evident in the names used (distinctively Black or distinctively White). He found that when teachers evaluated a writing sample with a vague scale (e.g., "far below grade level" and so on), teachers scored the same writing sample lower for the Black author than the White author. However, when they used a descriptive rubric for personal narratives, there was no evidence of racial bias in grading.

FUTURE DIRECTIONS FOR GRADING RESEARCH

In addition to continued research about SBG, I see at least three additional areas where grading research may grow, advance, and extend in the future. These are, in no particular order of likelihood or prediction, AI-assisted grading, competency-based grading (and perhaps some other alternative grading practices that haven't been invented yet), and equitable and culturally relevant grading.

Already existing, but mostly so far studied in the context of higher education, is a literature about the possibilities and benefits of AI-assisted grading (e.g., Flodén, 2025). I expect that to continue, maybe even explode, and to extend further into the K-12 context. The interest is in AI grading as a labor-saving device for teachers, which is important enough in its own right. However, my own view is that if AI can be harnessed to assist with formative feedback, as a labor-saving device for teachers, that would be even more important. Presumably, regarding assessment for summative purposes, teachers will grade what they must anyway, and as the SBG literature shows us, more grades don't mean better grading. However, regarding assessment for formative purposes and most notably feedback, teachers might give more feedback if it were less burdensome to do so, to the benefit of students and their learning.

Competency-based instruction and assessment have developed greatly since 2016, and in my view, it finds itself where SBG was in 2016: there is a literature of recommendations for implementation, but not much in the way of empirical studies. At present, every state in the U.S. allows for competency-based learning, but it is not yet a reality in schools (Stanford, 2024). I predict that, as for the wave of SBG research before it, there will be a wave of research into competency-based grading—at least, if it is to survive as a practice. Typically, as with SBG research, that would mean a wave of descriptive studies first (to identify what practices are being used in what ways and the perceptions of educators and students about these practices) followed by efficacy studies (to investigate effects on learning and motivation) and more theoretically grounded studies (to investigate the reasons for the effects).

Finally, and importantly, in the decade since 2016, there has been a growing interest in grading practices that are equitable and culturally relevant (Brookhart & Chism, 2025). As for competency-based instruction now (and for SBG ten years ago), there is a large and growing literature on recommendations for implementation of equitable and culturally relevant (e.g., Feldman, 2024) grading practices, but not much yet in the way of empirical studies. I would expect research in this area to proceed in the same manner: descriptive studies, followed by efficacy studies and more theoretically grounded studies.

In summary, research in grading is moving along nicely. I am particularly happy to see development in several topic areas at once that are somewhat different from the traditional grading research we summarized ten years ago (Brookhart et al., 2016). The next 100 years of grading research should be very exciting to researchers and, I hope, consequential for the improvement of practice.

REFERENCES

- Arsyad Arrafii, M. (2020). Grades and grade inflation: Exploring teachers' grading practices in Indonesian EFL secondary school classrooms. *Pedagogy, Culture and Society*, 28(3), 477-499. <https://doi.org/10.1080/14681366.2019.1663246>
- Ashta, J. K., Weingart, R., & Gazmararian, J.A. (2023). The impact of COVID-19 on education experiences of high school students in semi-rural Georgia. *Journal of School Health*, 93(4), 257-265. <https://doi.org/10.1111/josh.13269>
- Battistone, W., Buckmiller, T., & Peters, R. (2019). Assessing assessment literacy: Are new teachers prepared to assume jobs in school districts engaging in grading and assessment reform efforts? *Studies in Educational Evaluation*, 62, 10-17. <https://doi.org/10.1016/j.stueduc.2019.04.009>
- Brookhart, S. M., & Chism, D. (2025). Equitable and culturally relevant grading practices. In C. Evans & C. Taylor (Eds.), *Culturally responsive assessment in classroom and large-scale contexts: Theory, research, and practice* (pp. 306-323). Routledge.

- 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>
- Buckmiller, T., Townsley, M., & Cooper, R. (2020). Rural high school principals and the challenge of standards-based grading. *Theory and Practice in Rural Education*, 10(1), 92-102. <https://doi.org/10.3776/tpre.2020.v10n1p92-102>
- Camargo Salamanca, S., Parra-Martinez, A., Chang, A., Maeda, Y., & Traynor, A. (2024). The effect of scoring rubrics use on self-efficacy and self-regulation. *Educational Psychology Review*, 36, 70. <https://doi.org/10.1007/s10648-024-09906-w>
- Chen, P., & Bonner, S. (2023). Teachers' beliefs about grading, grades, and student classroom conduct. *Educational Practice and Theory*, 45(2), 69-91. <https://doi.org/10.7459/ept/45.2.06>
- Cheng, L., Baidoo-Anu, D., & DeLuca, C. (2023). Examining school principals' conceptions of assessment and grading practices. *Alberta Journal of Educational Research*, 69(1), 41-65. <https://journalhosting.ucalgary.ca/index.php/ajer/>
- Doz, D. (2023). Factors influencing teachers' grading standards in mathematics. *Oxford Review of Education*, 49(6), 819-837. <https://doi.org/10.1080/03054985.2023.2185217>
- Feldman, J. (2024). *Grading for equity* (2nd ed.). Corwin.
- Fergus, S., & Smith, C. P. (2022). Characteristics of proficiency-based learning and their impacts on math anxiety in the middle grades. *Research in Middle Level Education Online*, 45(4), 1-19. <https://doi.org/10.1080/19404476.2022.2045810>
- Flodén, J. (2025). Grading exams using large language models: A comparison between human and AI grading of exams in higher education using ChatGPT. *British Educational Research Journal*, 51(1), 201-224. <https://doi.org/10.1002/berj.4069>
- Gamlem, S., Smith, K., & Sandvik, L. V. (2023). Stakeholder opinions about canceling exams in Norwegian upper secondary school during the pandemic, and its consequences—An illuminative study. *Assessment Matters*, 17. <https://doi.org/10.18296/am.0063>
- Guskey, T. R., & Link, L. J. (2019). Exploring the factors teachers consider in determining students' grades. *Assessment in Education: Principles, Policy & Practice*, 26(3), 303-320. <https://doi.org/10.1080/0969594X.2018.1555515>
- Guskey, T. R., Townsley, M., & Buckmiller, T. M. (2020). The impact of standards-based learning: Tracking high school students' transition to the university. *NASSP Bulletin*, 104(4), 257-269.
- Huey, M. E., Silvey, P. R., Vaughan, A. G., & Fisher, A. L. (2022). Assessing the impact of standards-based grading policy changes on student performance and practice work completion in secondary mathematics. *Studies in Educational Evaluation*, 75, 101211. <https://doi.org/10.1016/j.stueduc.2022.101211>
- Ketsman, O., & Reeves, T. (2023). Standards-based grading in a small, suburban district: Teacher education, confidence, and implementation. *Journal of School Administration Research & Development*, 8(2), 65-75. <https://doi.org/10.32674/jsard.v8i2.5565>
- Klapp, A. (2017). Does academic and social self-concept and motivation explain the effect of grading on students' achievement? *European Journal of the Psychology of Education*. <https://doi.org/10.1007/s10212-017-0331-3>
- Knight, M., & Cooper, R. (2019). Taking on a new grading system: The interconnected effects of standards-based grading on teaching, learning, assessment, and student behavior. *NASSP Bulletin*, 103(1), 65-92. <https://doi.org/10.1177/0192636519826709>
- Kramer, S. L., Posner, M. A., Browman, A. S., Lawrence, N. R., Roem, J., & Krier, K. (2024). The impacts of a standards-based grading system emphasizing formative assessment, feedback, and re-assessment: A mixed methods, cluster randomized control trial in ninth grade mathematics classrooms. *Journal of Research on Educational Effectiveness*, 18(1). <https://doi.org/10.1080/19345747.2023.2287594>
- Kunnath, J. P. (2017). Teacher grading decisions: Influences, rationale, and practices. *American Secondary Education*, 45(3), 68-88.
- Larsen, A. (2023). Investigating the potential benefits of standards-based grading practices at urban secondary schools in Southern Utah: A qualitative study. *Journal of Nonprofit Innovation*, 3. <https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=1076&context=joni>
- Liljeröd, H., Jönsson, A., Klapp, A., & Jonsson, A.-C. (2025). Students' perceptions of how grades influence their motivation: Voices of upper secondary school students in Norway and Sweden. *Educational Assessment, Evaluation and Accountability*, <https://doi.org/10.1007/s11092-025-09454-z>
- Link, L. J., & Guskey, T. R. (2022). Is standards-based grading effective? *Theory into Practice*, 61(4), 406-417. <https://doi.org/10.1080/00405841.2022.2107338>

- Morris, S. R., & Johnson, A. H. (2024). Could minimum grading enhance high school graduation rates and cost-effectiveness across Arkansas? *Education Policy Analysis Archives*, 32(40). <https://doi.org/10.14507/epaa.32.8496>
- Olsen, B., & Buchanan, R. (2019). An investigation of teachers encouraged to reform grading practices in secondary schools. *American Educational Research Journal*, 56(5), 2004-2039. <https://doi.org/10.3102/0002831219841349>
- Panadero, E., Jonsson, A., Pinedo, L., & Fernández-Castilla, B. (2023). Effects of rubrics on academic performance, self-regulated learning, and self-efficacy: A meta-analytic review. *Educational Psychology Review*, 35, 113. <https://doi.org/10.1007/s10648-023-09823-4>
- 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), Article 2. <https://digitalcommons.lindenwood.edu/ela/vol2/iss2/2>
- 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 initiative in the Midwestern United States. *American Secondary Education*, 45(3), 9-28.
- Quinn, D. M. (2020). Experimental evidence on teachers' racial bias in student evaluation: The role of grading scales. *Educational Evaluation and Policy Analysis*, 42(3), 375-392. <https://doi.org/10.3102/0162373720932188>
- Stanford, L. (2024, September 17). All states allow competency-based learning. Will it become a reality in schools? *Education Week*. [All States Allow Competency-Based Learning. Will It Become a Reality in Schools? \(edweek.org\)](https://www.edweek.org/all-states-allow-competency-based-learning-will-it-become-a-reality-in-schools/2024/09/17)
- St. Pierre, N. A., & Wuttke, B. C. (2017). Standards-based grading practices among practicing music educators: Prevalence and rationale. *Update: Applications of Research in Music Education*, 35(2), 30-37. <https://doi.org/10.1177/8755123315604468>
- Townsley, M. (2020). Grading principles in pandemic-era learning: Recommendations and implications for secondary school leaders. *Journal of School Administration Research and Development*, 5(S1), 8-14. <https://doi.org/10.32674/jsard.v5iS1.2760>
- Townsley, M., Buckmiller, T., & Cooper, R. (2019). Anticipating a second wave of standards-based grading implementation and understanding the potential barriers: Perceptions of high school principals. *NASSP Bulletin*, 103(4), 281-299. <https://doi.org/10.1177/0192636519882084>
- Townsley, M., & Knight, M. (2020). Making change stick: A case study of one high school's journey towards standards-based grading. *Journal of Educational Leadership in Action*, 6(3), Article 3. <https://doi.org/10.62608/2164-1102.1014>
- Townsley, M., & Kunnath, J. (2022). Exploring state department of education grading guidance during COVID-19: A model for future emergency remote learning. *Education Policy Analysis Archives*, 30(163). <https://doi.org/10.14507/epaa.30.7448>
- Townsley, M., & McNamara, S. (2021). "I thought I was supposed to get an A in PE!" Successes and challenges of teachers and administrators implementing standards-based grading in physical education. *Studies in Educational Evaluation*, 70. <https://doi.org/10.1016/j.stueduc.2021.101012>
- Townsley, M., & McNamara, S. (2024). "PE is more than playing kick ball the entire hour": Using standards-based grading to elevate physical education's status among preservice physical educators. *Journal of Teaching in Physical Education*, 43(4), 733-739. <https://doi.org/10.1123/jtpe.2023-0064>
- 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. <https://eric.ed.gov/?id=EJ1168171>
- Wisch, J. K., Ousterhout, B. H., Carter, V., & Orr, B. (2018). The grading gradient: Teacher motivations for varied redo and retake policies. *Studies in Educational Evaluation*, 58, 145-155. <https://doi.org/10.1016/j.stueduc.2018.06.005>

SUSAN BROOKHART, PhD, is a Professor Emerita in the Department of Educational Foundations and Leadership at Duquesne University. Her major research interests are the role of both formative and summative classroom assessment in student motivation and achievement, the connection between classroom assessment and large-scale assessment, and grading. Email: suebrookhart@gmail.com

Manuscript submitted: June 26, 2025
Accepted for publication: October 30, 2025