How Traditional Grading Contributes to Student Inequalities and How to Fix It

By Laura J. Link and Thomas R. Guskey

Grades have long been identified by those in the measurement community as prime examples of unreliable measurement (Brookhart, 1994; Stiggins, Frisbie, & Griswold, 1989). What one teacher considers in calculating students’ grades may differ greatly from another teacher (Guskey & Link, 2019; McMillan, 2001; McMillan, Myran, & Workman, 2002). A major factor contributing to the unreliability of grades is teachers’ inclusion of aspects of students’ behavior in the grades they assign. Despite the recommendation of experts to separate behavior from academic achievement in formulating students’ grades, teachers at all grade levels typically include student behavior as a contributing factor in determining grades (Brookhart, Guskey, Bowers, McMillian, Smith, J., Smith, L., & Welsh, 2016; Frary, Cross, & Weber, 1993; Gullickson, 1985; Link, 2018; McMillian & Nash, 2000; Randall & Engelhard, 2010).

In assigning grades, teachers typically divide the evidence they gather from students into different categories such as tests, quizzes, homework, labs, participation, effort, attendance, etc. Using a computerized grading program, they then assign a percentage weight to each category specifying its contribution to each student’s subject area or course grade. This combination of evidence yields an amalgamated “hodgepodge” grade (Brookhart, 1991, p. 36) that mixes achievement and other non-academic factors related to various aspects of students’ behavior. Including indicators of students’ behavior distorts their meaning of grades, however, and drastically diminishes their communicative value. In addition, because teachers vary in the weight they attach to these factors in determining students’ grades, it also makes grades less reliable indicators of students’ performance. Grades that include factors such as effort and participation become tools for managing students’ behavior as much as they are indicators of students’ learning (Olsen & Buchanan, 2019).

Despite their noted unreliability, grades remain the basis for making many important decisions about students (Brookhart et al., 2016; Guskey, 2015). Report card grades determine whether or not students are promoted from one grade level to the next. They also determine honor roll status, enrollment in advanced or remedial classes, special education services, and college or university admissions (Brookhart & Nitko, 2008). Because grades typically include a mix of academic and behavioral factors, however, students’ academic opportunities may be unevenly affected when implicit racial and gender biases influence how teachers consider behavioral factors when assigning grades.

Race and Behavioral Grades

Research shows that teachers treat students differently depending on students’ race, and these differences contribute to racial inequalities in grading, especially when behavioral factors are considered (McKown & Weinstein, 2008; Okonofua, Walton, & Eberhart, 2016; Rubie-Davies, Hattie & Hamilton, 2006). Studies indicate, for example, that white teachers tend to perceive black students as more disruptive than white students (Downey & Pribesh, 2004; Ferguson, 2000), and as less mature (Alexander, Entwisle, & Thompson, 1987). These...
differences in teachers’ perceptions shape their treatment of students in the classroom and their use of disciplinary actions. Other studies reveal that black students are more likely than their white peers to be reprimanded for behavioral offenses such as insubordination, disrespect, and excessive noise (Diamond & Lewis, 2015; Ford, 2016). Black students are also more likely than white students to be referred to the office or suspended, even when the misbehaviors are similar (Lleras, 2008). Results of suspensions often translate into reduced teaching and learning access, which can negatively impact students’ success in the classroom. Subsequently, when teachers include indicators of student behavior in determining students’ grades, black students are more likely to be negatively affected than their white peers.

When teachers interpret student behaviors through the lens of race, credit for behaviors such as being seated when the bell rings, following directions, cooperation, and dressing appropriately may be inequitably assigned. In an early study, for example, Brophy and Good (1974) found that some teachers develop simplistic and rigid stereotypes, and they react more to the stereotypes than to the students themselves. A more recent meta-analysis examining teachers’ expectations of students based on race and other teacher stereotyping studies support similar findings (Tenenbaum & Ruck, 2007; Willard, Isaac, & Carney, 2015). As a result, racial stereotypes may lead teachers to award more behavioral credit to white students and less to black students for their perceived classroom conduct.

Such differences can have profound influence on students’ grades. If, for instance, a combination of behavioral factors (e.g., effort, participation, class conduct, homework completion, etc.) counts 20 percent of the final grade, awarding maximum points for behavior could move a student from a C to an A in the typical percentage grading system. Conversely, students who are perceived as not meeting behavioral expectations could drop from a grade of C to a D or F.

In addition, teachers work under conditions that tend to heighten the negative impact of racial stereotypes. Throughout the school day, teachers make numerous micro-decisions about students’ behavior amid working conditions that are highly stressful and cognitively demanding. This is particularly true in low-resourced schools that serve disproportionately large numbers of minority students of color (Warikoo, Sinclair, Fei, & Jacoby-Senghor, 2016). These are precisely the kinds of situations in which implicit biases and stereotypes have their greatest effect. Implicit associations have an even stronger impact when teachers are unable to devote cognitive resources to their own behaviors and decisions, instead relying on spontaneous, gut reactions (Cameron, Brown-Iannuzzi, & Payne, 2012; Olson & Fazio, 2009). These reactions play out in teachers’ grading decisions. In moments of cognitive overload, teachers are more likely to impose grade reductions on students who aren’t following established classroom procedures or who display disruptive behavior. Evidence indicates that when teachers are trying to balance multiple demands, they are more susceptible to the influence of implicit racial biases and to use grades as a means of control (Warikoo, Sinclair, Fei, & Jacoby-Senghor, 2016).

Gender and Behavioral Grades

Teachers’ grading practices are also influenced by students’ gender. Girls have long received higher grades in school than boys. Even in the 1950s and 1960s, girls earned better grades and had higher class standing in high school (Alexander & Eckland, 1974; Mickelson, 1989). Today, from kindergarten through high school and even in college, girls get better grades in all major subjects, including math and science – subjects traditionally viewed more suitable for boys (Perkins, Kleiner, Roey, & Brown, 2004; Terrier, 2016). This may be explained in part because girls typically display better social skills and classroom behavior.

As early as kindergarten, boys exhibit more disruptive conduct in class and less positive orientations to learning activities (Zill & West, 2001). According to elementary school teacher reports, twice as many boys as girls have difficulty paying attention (Buchman & DiPrete, 2006). Girls also demonstrate greater persistence in completing tasks and greater eagerness to learn (Buchman & DiPrete, 2006; McDaniel, 2007). During adolescence, high school teachers consistently rate girls as putting forth more effort, being more attentive, more organized, and less
disruptive than boys (Downey & Vogt Yuan, 2005). Furthermore, girls are generally more adept at reading test instructions before proceeding to the questions, paying attention to the teacher rather than daydreaming, choosing homework over TV, and persisting in long-term assignments despite boredom and frustration than are boys. These differences in non-cognitive skills may be central in explaining why girls generally get higher grades. Boys’ less developed self-discipline skills leave them at a disadvantage in school settings where grades weigh self-regulation and organizational skills alongside demonstrations of acquired knowledge.

Including behavior in grades plays right into most girls’ strengths—and most boys’ weaknesses. A host of cross-cultural studies show that females tend to be more conscientious than males (Hogan, 1981; King & Hill, 1993; Kobrin, Sathy, & Shaw, 2007). In school, girls are more apt to take more detailed notes in class, transcribe more accurately what teachers say, complete homework on time, and invest in impressing their teachers with their efforts (Buchmann & DiPrete, 2006; McDaniel, 2007).

On a whole, boys approach schoolwork differently. They are less satisfied with the whole enterprise of organizing their work and tending to details. As a result, they are more apt to be inattentive, leave completed assignments at home, and fail to turn the page and complete the questions on the back (Gnaulati, 2014). Boys are also more likely to blurt out answers, doodle instead of taking notes, have messy backpacks, and even poke students who sit in front of them (McLeod & Kaiser, 2004). When such transgressions are considered in determining students’ grades, fairness issues come into play, especially if teachers assign zeroes for work that is missing, turned in late, or incomplete. A single zero can doom a student to failure, regardless of what dedicated effort or level of performance might follow (Guskey, 2015). When combined with the common practice of averaging scores from different sources of evidence, a single zero can have a devastating effect on a student’s percentage grade. The overall grade is unfairly skewed by that one score, leaving boys’ achievement underestimated and feeling alienated in an environment where self-regulation and conscientiousness account for a good portion of their grades.
Grades versus Other Measures of Achievement

Even though minority students and boys are more susceptible to lower course grades due to perceptions of classroom behavior, they are paradoxically experiencing increasing levels of success on external assessments of their achievement. Although still not outscoring their white peers, black and Hispanic students, in particular, are earning higher scores than ever in math and reading on the National Assessment of Educational Progress, or NAEP. According to the U.S. Department of Education (2015), while the overall math averages for 9-year-olds grew by 25 points between 1978 and 2012, average NAEP scores among black and Hispanic students increased by 34 and 31 points, respectively. Among 13-year-olds, math scores for white students increased by 21 points, while results for blacks and Hispanics increased by 34 points and 33 points, respectively. White 17-year olds, many of whom are one year away from enrolling in college, nudged upward by six points overall between 1978 and 2012 on the math portion of NAEP, but scores for black and Hispanic students increased by 20 and 18 points, respectively. The same holds true for NAEP reading scores. Between 1975 and 2016, black and Hispanic students’ reading assessment scores grew by more than 20 points on average across all grade levels (NCES, 2017). Additionally, the number of minority students earning a passing score on at least one Advanced Placement course exam has nearly doubled from 2004 to 2018 (College Board, 2018).

A similar grade paradox holds true for boys: Girls may earn higher grades than boys throughout elementary, middle and high school, but they do not outperform boys on achievement or IQ tests. In a landmark study by Duckworth and Seligman (2006) investigating the role of gender in grades and achievement, girls earned significantly higher final grades than boys in high school Algebra II, English, and social studies. Despite these high grades, however, since 1972, boys have overshadowed girls on the SAT, registering higher overall scores every year by an average of 45 points (College Board, 2018).

How to Fix Grade Inequities

To fix these grade inequalities and limit the potential influence of bias in grading, we must do three things: (1) Determine students’ grades based on learning criteria; (2) Distinguish product, process, and progress criteria; and (3) Report each type of criteria separately.

Determine Students’ Grades Based on Learning Criteria

When asked to identify the purpose of grading, most teachers indicate that grades should describe how well students have achieved the learning goals established for a grade level or course. In other words, grades should reflect students’ performance based on specific learning criteria, not their relative standing among classmates. Teachers as well as students prefer this approach because they consider it both fair and equitable (Kovas, 1993).

Distinguish Product, Process, and Progress Criteria

As we described earlier, teachers use widely varying criteria in determining students’ grades. In most cases, these different criteria can be grouped into three broad categories: product, process, and progress criteria (Guskey, 1996).

* Product criteria reflect what students know and are able to do at a particular point in time. Teachers who use product criteria typically base students’ grades on final examination scores, final products (reports or projects), overall assessments, and other culminating demonstrations of learning.

* Process criteria emphasize behaviors that enable or facilitate learning. Teachers who consider effort or work habits when assigning grades are using process criteria. So are teachers who count formative assessments, homework, punctuality of assignments, class participation, or attendance.

* Progress criteria describe how much students gain from their learning experiences. Other names for progress criteria include “learning gain,” “improvement scoring,” “value-added learning,” and “educational growth.” Teachers who use progress criteria typically look at how much improvement students have made over a particular period of time, rather than just where they are.
Because of concerns about student motivation, self-esteem, and the social consequences of grading, most teachers routinely base their grading procedures on some combination of all three types of criteria. Many also vary their grading criteria from student to student, taking into account individual circumstances. Although teachers defend this practice on the basis of fairness, it seriously confounds the meaning of any grade. A grade of A, for example, may mean the student knew what was intended before instruction began (product), did not learn as well as expected but tried very hard (process), or simply made significant improvement (progress).

Report Each Type of Criteria Separately
After establishing explicit indicators of product, process, and progress learning criteria, teachers should assign separate grades for each. In other words, they provide a “dashboard” of information rather than a single hodgepodge grade. In this way grades for homework, effort, work habits, responsibility or learning progress are kept distinct from grades that reflect academic achievement and performance. The intent is to provide a better, more accurate, and much more comprehensive picture of what students accomplish in school.

While schools in the U.S. are just beginning to catch onto the idea of separate grades for product, process, and progress criteria, many Canadian educators have used the practice for years (Bailey & McTighe, 1996). Each marking period, for example, teachers in Ontario assign an “achievement” grade to students based on their academic performance on projects, assessments, and other demonstrations of learning. In addition, they assign separate grades or marks for behaviors related to responsibility, organization, independent work, collaboration, initiative, and self-regulation. Ontario teachers say that reporting such factors separately compels students to take these behaviors more seriously. In addition, it offers parents a more comprehensive picture of their children’s performance in school (Tierney, Simon, & Charland, 2011).

Teachers often presume that reporting multiple grades will increase their grading workload. But those who use the procedure claim that it actually makes grading easier and less work. Teachers gather the same evidence on student learning that they did before, but no longer worry about how to weight or combine that evidence in calculating an overall grade. As a result, they avoid irresolvable arguments about the appropriateness or fairness of different weighting strategies.

Perhaps most important, reporting separate grades for product, process, and progress criteria also makes grading more meaningful and less prone to the influence of bias. By pulling out non-achievement factors from an achievement grade, the grade-inflating or deflating influence of students’ behavior is eliminated. Yet by including separate grades or marks on behavioral factors in the reporting procedures, however, their importance to teachers and students is maintained. It simply makes grading a more accurate and more meaningful form of communication. In turn, report cards and transcripts become more robust documents that present a better and more discerning portrait of students’ performance in school.

Conclusion
Developing meaningful, reliable, and equitable grading policies and practices will continue to challenge educators. Distinguishing specific product criteria and reporting achievement grades based on these criteria allow teachers to offer a more precise description of students’ academic achievement and performance. Reporting on specific process criteria related to homework, class participation, attitude, effort, responsibility, behavior, and other non-academic factors ensures they remain important but distinct. Doing so will clarify the meaning of grades, enhance their communicative value, and ensure far greater equity in grading at all education levels.

References

Innovative Approaches to Access and Equity for All Learners


Innovative Approaches to Access and Equity for All Learners


Laura J. Link, EdD, is an Assistant Professor of Educational Leadership & Policy in the College of Public Service at the University of Houston Downtown. She is the co-author of *Cornerstones of Strong Schools: Practices for Purposeful Leadership* and author of several articles, book chapters, and professional papers on school leaders, grading, and assessments. Dr. Link also leads K-16 research-practice partnerships and has won multiple awards for her community engagement. Dr. Link has served in many central office and school-based leadership roles and has taught elementary, middle, high school and college students. While she was Assistant Superintendent of Teaching & Learning in Memphis, TN, she was 1 of 7 administrators charged with leading the largest school district merger in United States’ history.

To learn more about her work, visit [www.GradingRx.com](http://www.GradingRx.com) or follow her on Twitter @laurajlink

Thomas R. Guskey, PhD, is a Senior Research Scholar in the College of Education and Human Development at the University of Louisville and Professor Emeritus in the College of Education at the University of Kentucky. He began his career as a middle school teacher, served as an administrator in the Chicago Public Schools, and is the author/editor of 24 award-winning books and more than 250 book chapters, articles, and professional papers on educational measurement, evaluation, assessment, grading, and professional learning. His articles have appeared in prominent research journals as well as practitioner publications, and he has received many honors and awards in recognition of his contributions to the field. To learn more about his work, visit [http://tguskey.com](http://tguskey.com) or follow him on Twitter @tguskey.