

END THE “EASY A”

TOUGHER GRADING STANDARDS SET MORE STUDENTS UP FOR SUCCESS

GRADE INFLATION IS PERVASIVE in American high schools. Over the past 20 years, grade point averages have soared while SAT scores and other measures of academic performance have held stable or fallen. As a result, supposedly “good” grades have become unreliable markers of knowledge and skills.

Is rampant grade inflation cause for concern? On the one hand, students who receive favorable marks despite struggling to master academic content may be encouraged in their studies; surely, many teachers who are generous in assigning grades have this logic in mind. On the other hand, grade inflation may be yet another manifestation of the “soft bigotry of low expectations” that President George W. Bush famously warned about. When students who have not mastered the material receive passing marks, they may become complacent and fail to reach their full potential.

How teachers’ grading standards affect student success is an empirical question—one that I address here in a new study of roughly 350,000 North Carolina students taking Algebra I between 2006 and 2016. I first measure the grading standards of individual Algebra I teachers in the state by

comparing the course grades they assigned their students to those students’ scores on a standardized end-of-course exam. I then ask whether students did better or worse than expected when they were assigned to a more-demanding teacher.

My results confirm that “everyone gets a gold star” is not a victimless mentality. Not only do students learn more from tougher teachers, but they also do better in math classes up to two years later. The size of these effects is on the order of replacing an average teacher with one near the top of her game.

Exactly how higher grading standards lead to greater student success is not clear, and there may be multiple, overlapping factors at play. What is clear, however, is that inflated grades lead to a host of unintended consequences, and that teachers’ grading standards are malleable. This presents a clear opportunity to act, to guard against the “easy A” and work together to enforce higher grading standards.

Parents faced with stressed-out children and an increasingly competitive college-admissions process may resist calls for more-rigorous grading. Educators and school leaders may be tempted to satisfy them, which is part of

by SETH GERSHENSON



ILLUSTRATION / THEO RUDNAK

how the grade-inflation problem was created to begin with. But policymakers and other decisionmakers would deserve a genuine A if they reminded parents, principals, and teachers that they aren't doing students any favors by depriving them of appropriate academic challenges or an accurate picture of their knowledge, skills, and abilities.

A question of standards

There is surprisingly little empirical evidence to back up the intuitive idea that high grading standards boost student learning. The best evidence to date comes from a study of elementary-school students in Florida (see "The Gentleman's A," *research*, spring 2004), in which Maurice Lucas and David Figlio found that students whose classroom teachers had high grading standards did better in math and reading, and that those effects were largest for high-achieving students. They also found that parents spent significantly more time at home helping children with a tougher-grading teacher, suggesting that the effect of high grading standards operates partly through increased parental involvement.

This research dovetails with prior evidence from Julian Betts and Jeff Grogger, who used data from a nationally representative sample of 10th graders to show that higher school-level grading standards, defined as schools' average gap between GPAs and standardized test scores, boost student achievement. Both studies found that the effects of grading standards on achievement are positive for all students and largest for high achievers. However, Betts and Grogger also found that black and Hispanic students attending a high school with higher grading standards were less likely to graduate, suggesting that higher standards could have adverse consequences for traditionally disadvantaged subgroups.

Indeed, the same high grading standards might improve some students' outcomes while harming those of others. For example, consider two classmates whose teacher has high grading standards and who both receive a C on their mid-semester report cards. If the students have different temperaments or innate ability levels, one student might be invigorated to improve her study habits while the other takes this same information as a signal that the subject is too difficult for her and further disengages from school.

The analysis I describe below builds on this prior research first by investigating how the grading standards of a high-school math teacher affect content mastery, as measured by performance on the end-of-course Algebra I exam. I also examine whether a teacher's grading standards affect students' performance in subsequent math courses and the students' likelihood of graduating from high school. I explore whether the effects of grading standards vary for students from different demographic groups and with the type of school students

Not only do students learn more from tougher teachers, but they also do better in math classes up to two years later.

attend. Finally, having shown that teachers' grading standards matter, I look to see what school and teacher characteristics are associated with having higher standards.

Data and method

I focus on Algebra I teachers and students for both practical and theoretical reasons. From a practical standpoint, Algebra I was continuously required for high-school graduation in North Carolina throughout the study period of 2006-16, subject to an end-of-course standardized test, and uniformly identified in students' transcript data. From a theoretical standpoint, math is the subject most affected by teachers and other schooling inputs, perhaps because parents and other household members are less likely to help students with their math work. If teachers' grading practices matter for student learning, math is where we'd expect to see it.

The first order of business is to define and measure teachers' grading standards. I focus on Algebra I classrooms that had a single teacher for the entire academic year, resulting in a group of about 8,000 Algebra I teachers who taught about 350,000 8th- and 9th-grade students.

Having both course grades and end-of-course exam scores allows me to define teachers' grading standards in an intuitive way. For each teacher in the state, I compute the average exam score of every one of her students who received a grade of B in the course. For example, suppose that the average test score of the students who received a B from Ms. Apple was 80 points, while the average test score of students who received a B from Ms. Banana was 90 points. This implies that Ms. Banana has higher grading standards than Ms. Apple, because Ms. Banana's students learned more to earn their Bs. We can then sort teachers by this measure and designate the bottom 25 percent as the easiest graders, the top 25 percent as the toughest graders, and so on.

The next challenge is to isolate the causal effect of teachers' grading standards on student outcomes. Because students are not randomly assigned to teachers, we might worry that concerned parents or principals ensure that certain children are assigned to teachers with high grading standards. If so, we'd be unable to distinguish the effect of having a teacher with high grading standards from the

effect of those involved parents and principals.

I control for such confounding factors in my analysis first by adjusting for the students' demographic characteristics and their performance on the previous year's end-of-grade standardized test. I also adjust for the demographics and past performance of all of the teachers' current students, as a student's classmates might influence both the teachers' behavior and the student's outcomes. And finally, to guard against concerns that school culture, district policies, or principal effects drive both teacher grading standards and student outcomes, I limit my comparisons to students of teachers with higher and lower standards who are taking Algebra I in the same school, in the same grade, in the same year.

A related concern is that teachers with strict grading standards may differ from teachers with lax grading standards in other ways, too. If so, we'd again be unable to differentiate the effect of grading standards from the effects of these other differences. I attempt to address this concern by adjusting for other observed teacher characteristics that are known to influence student test scores, such as teaching experience and the selectivity of their undergraduate institution.

Even so, it remains possible that teachers who are tough graders share other attributes or classroom practices in common that influence their students' success. Strictly speaking, my analysis isolates the effects of having a teacher with high grading standards rather than the effects of high grading standards per se. This distinction is important to keep in mind when interpreting the results.

Effects on student achievement

To simplify the analysis, I sort teachers into four evenly sized groups based on their grading standards, where group 1 has the lowest standards and group 4 has the highest standards. Teachers with the highest standards increase student test scores by a whopping 17 percent of a standard deviation compared to their counterparts in the bottom quartile (see Figure 1). To put this difference in perspective, consider that it amounts to a little more than six months of learning. It is also larger than the impact of a dozen student absences or replacing an average teacher with a teacher whose students consistently outperform expectations. Teachers whose grading standards are in the middle are not as successful in raising student achievement as their tougher peers, but they are significantly more effective than teachers with the lowest grading standards.

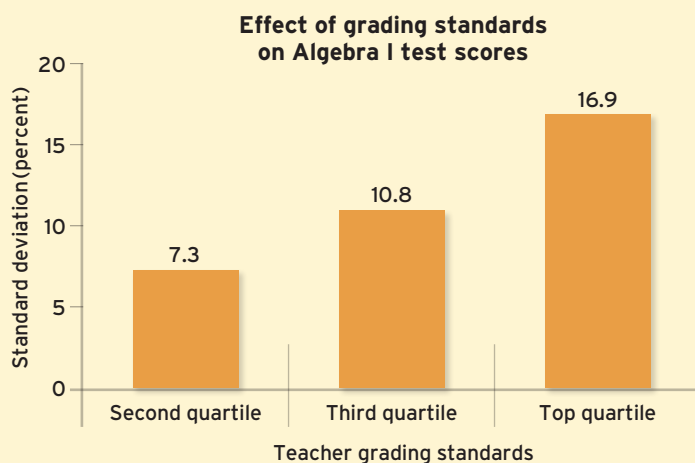
I also find that teachers with high grading standards improve their students' subsequent

performance in other math classes up to two years later. In looking at students' performance on end-of-course exams in geometry and Algebra II, students whose Algebra I teachers had the highest grading standards consistently experience higher achievement in subsequent math exams: 7 percent of a standard deviation in geometry and 9 percent of a standard deviation in Algebra II (see Figure 2). Again, these effects translate into meaningful differences of about 2.5 and 3.2 months of learning, respectively. Since these tests are in somewhat different subjects and are taken one and two years later, it is not surprising that the effects on these longer-range outcomes are smaller than the same-year Algebra I effects.

I then explore whether grading standards affect longer-run measures of educational attainment—specifically, high school completion and college intentions. These outcomes are measured three to four years after taking Algebra I in the 8th or 9th grade. I find no effect on high school completion, perhaps because students who take Algebra I early or on time are already unlikely to drop out.

Students Learn More From Teachers With Higher Grading Standards (Figure 1)

Teachers with the most rigorous grading standards increase student test scores by 17 percent of a standard deviation compared with their peers with easiest grading standards. Teachers in the middle of the grading standards distribution also improve student test scores more than teachers with less strict grading standards.



NOTE: Bars represent the effect of having a teacher in each quartile compared to having a teacher in the bottom quartile. All estimates are statistically significant at the 99-percent confidence level.

SOURCE: Author's calculations.

However, I do find some suggestive evidence that stricter grading standards increase students' intent to attend a four-year college or university after high school. Specifically, teachers with above-median grading standards appear to increase students' stated college intent by about 1 percentage point, or 2.4 percent. This result falls short of conventional levels of statistical significance, but it suggests that exposure to higher grading standards may change students' attitudes toward school outside of mathematics classrooms and performance. At a minimum, it casts doubt on concerns that higher grading standards could discourage students from pursuing higher education.

Finally, I look at how teachers' grading standards affect the performance of students from various demographic groups and at different types of schools. In both cases, higher grading standards appear to be universally beneficial.

I find that the 75 percent of teachers with the strictest grading standards significantly improve the learning outcomes of all subgroups of students defined in terms of race, gender, economic disadvantage, and prior math achievement (see Figure 3). On the whole, having a teacher with higher grading standards improves achievement by about 10 percent of a test-score standard deviation, which amounts to about 3.6 months of learning. These effects

There is surprisingly little empirical evidence to back up the intuitive idea that high grading standards boost student learning.

are similar in size for each subgroup: the effect ranges from about 8 percent to 10 percent of a test-score standard deviation. The fact that *all* student subgroups benefit from exposure to higher grading standards should alleviate any concern that some students, especially low performers, may be harmed by strict standards.

Similarly, teachers with higher grading standards benefit students in all types of schools: middle and high schools; suburban, urban, and rural schools; and schools that predominantly enroll economically advantaged and economically disadvantaged students. Students attending suburban schools benefit somewhat more from higher grading standards, with an effect of 12 percent compared to 7 percent for urban schools. This is the only difference that is statistically significant. Otherwise, I find that high standards are equally beneficial in all school types.

Which teachers and schools have higher standards?

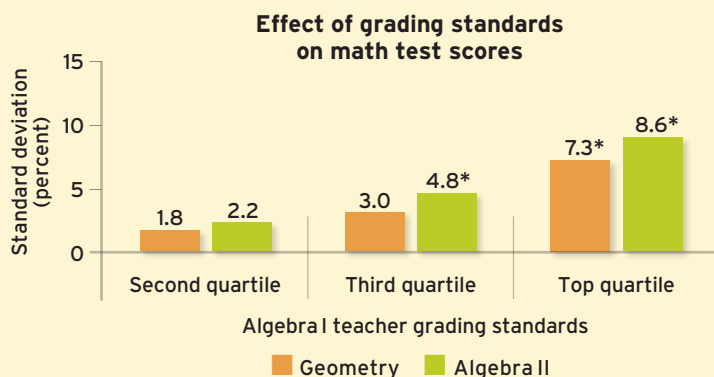
My findings so far provide compelling evidence that having high grading standards is an important attribute of effective teachers, but what leads some teachers to have higher standards than others? Comparing the characteristics of teachers with higher and lower standards reveals a range of factors both before and during teachers' careers that may influence their approach to grading.

First, teachers' own educations appear to influence how rigorously they grade their students' work. Teachers who attended selective undergraduate institutions and teachers who have completed advanced degrees both tend to have higher grading standards.

I compare the average grading standards of teachers who did and did not earn their undergraduate degrees from institutions rated as "most" or "highly" selective by *Barron's Profiles of American Colleges*. The difference is sizable:

Benefits of High Grading Standards Persist in Later Math Courses (Figure 2)

Students whose Algebra I teachers had the highest grading standards continued to earn higher test scores in Geometry and Algebra II, courses taken one and two years later.



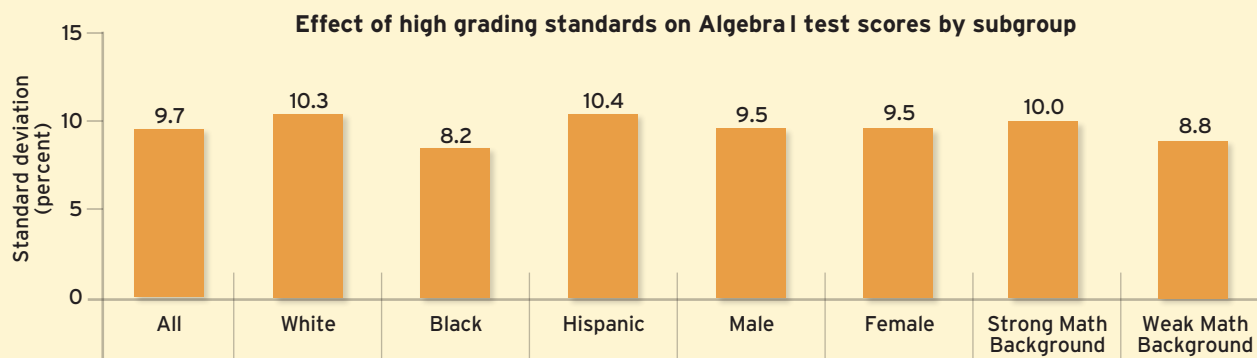
NOTE: Bars represent the effect of having an Algebra I teacher in each quartile compared to having a teacher in the bottom quartile. *indicates statistical significance at the 95-percent confidence level.

SOURCE: Author's calculations.

All Students Learn More From Teachers With Higher Grading Standards

(Figure 3)

All student subgroups benefit from exposure to higher grading standards. On average, having an Algebra I teacher in the top 75 percent of the grading standards distribution improves achievement by about 10 percent of a test-score standard deviation compared to having a teacher with the easiest grading standards.



Notes: Bars represent the effect of having a teacher in the top three quartiles of the grading standards distribution compared to having a teacher in the bottom quartile. All estimates are statistically significant at the 99-percent confidence level.

SOURCE: Author's calculations.

the average grading standards of teachers who attended selective colleges are 50 percent higher than those of teachers who earned their undergraduate degrees from less-selective schools. I find an even larger difference when comparing teachers with a graduate degree to those without: the average grading standards of teachers who have earned a graduate degree are more than twice as high as the grading standards of teachers without a graduate degree. Together, these results suggest that experiences in more-challenging academic environments may promote higher standards.

Second, my analysis shows that grading standards adjust based on teacher experience and school settings—findings relevant to policymakers and school leaders considering grading-standard interventions or policy changes.

In looking at teacher experience, I see that as years on the job increase, grading standards increase as well. On average, teachers' grading standards grow more rigorous the longer they remain in the profession, particularly during their first 15 years. Grading standards tend to be higher in middle schools, suburban schools, and schools serving more advantaged students.

I then look at teachers' grading standards by school type. Unlike my analysis of student performance, this is necessarily a descriptive exercise; teachers are not randomly assigned to schools and school-level factors might influence student test scores. First, I compare middle schools to high schools,

since students in the sample took Algebra I in either the 8th or 9th grade, and school cultures likely vary considerably between middle and high schools. Grading standards are markedly higher in middle schools, on average.

A second analysis compares schools with higher rates of student poverty to schools with more affluent students. Once again, there is a dramatic difference, with significantly higher standards in the more advantaged schools. Finally, in looking at school types by location, I find grading standards are highest in suburban schools and lowest in rural schools.

Implications

These findings are a call to action. Students assigned to teachers with the lowest standards do far worse on an end-of-course exam than their peers with tougher teachers, and continue to underperform those students one and two years later. We know that teachers' grading standards are an important component to their students' success, and we have started to identify the characteristics of teachers associated with higher standards—including those that can be influenced through training or experience. Three main lessons stand out.

First, education leaders at all levels can acknowledge that grade inflation is the path of least resistance and that it

takes active measures to uphold high standards. As Success Academy Charter Schools founder Eva Moskowitz has put it,

When teachers give high grades for mediocre work, no one asks any questions and they can carry on as before. When they give more realistic grades, they have an obligation to follow up with detailed feedback, more support, and better instruction. It's not surprising then that most—often unconsciously—opt for the first course of action.

By monitoring grading practices and ensuring that teachers are not pervasively awarding “easy As,” leaders can promote higher standards and the positive effects they have on student learning. This effort can include leaders of schools, districts, states, and schools of education, especially since

Teachers with the highest standards increase student test scores by a 17 percent of a standard deviation compared to teachers with the lowest standards—a difference that amounts to more than six months of learning.



PHOTOGRAPH / JAMES ESTRIN, THE NEW YORK TIMES, REDUX

Success Academy Founder and CEO Eva Moskowitz: “When teachers give high grades for mediocre work, no one asks any questions and they can carry on as before.”

newer teachers tend to have the lowest standards.

Some teacher-training programs are already onboard. For example, Teach for America’s summer institute includes a module explicitly dedicated to the power and importance of holding high expectations for all students, in which fellows discuss the importance of viewing students as individuals and not as members of a particular demographic group. Similarly, the teacher professional-development program Great Expectations emphasizes the importance of having high expectations for all students.

Second, teachers’ grading standards can serve as a useful measure of effectiveness to schools and districts when offering professional-development opportunities and deciding which teachers to retain and promote. Observable markers of effective teaching are in short supply. Grading-standard measures of the sort I use in this analysis can allow schools and districts to identify and retain teachers who implement high standards. These same measures also may help schools and districts to provide opportunities and resources for improvement to teachers with low standards.

Finally, it is incumbent on policymakers, researchers, and education leaders to make clear the damaging consequences of both low grading standards and grade inflation. Inflated grades can lead to a sense of complacency that prevents students from reaching their full potential and prevent parents from understanding what challenges their children face and holding them accountable for their performance. Moreover, socioeconomic gaps

in this type of grade inflation can contribute to analogous gaps in students’ educational outcomes.

Of course, changing both policy and practice is easier said than done. As researchers continue to enhance our understanding of why and how grading standards matter, practitioners can ensure “high standards” are a more common part of teaching culture through improved training and professional development. There is much work to be done.

Seth Gershenson is an associate professor at the School of Public Affairs at American University. This essay is adapted from the report “Great Expectations: The Impact of Rigorous Grading Practices on Student Achievement,” published by the Thomas B. Fordham Institute.