Prologue

Mallory's Dilemma

The data couldn't be possible. Actually, it *shouldn't* be possible.

Mallory had just completed her first year as principal of Centennial College Prep Middle School, a new public charter school in Huntington Park, California. As a young, white woman leading a school that served nearly all Latino students, many living below the poverty line, Mallory had approached her job humbly, not immediately pushing initiatives and changing policies to align to her own personal vision (what she called the "new sheriff in town approach"). Instead, her priority was to first understand her school community: its context, history, strengths, and needs. She had watched, listened, and built relationships with her faculty, students, and their families. She had visited classrooms, reviewed teachers' lesson plans, and studied the school's statistics: attendance percentages, disciplinary referrals, and test scores.

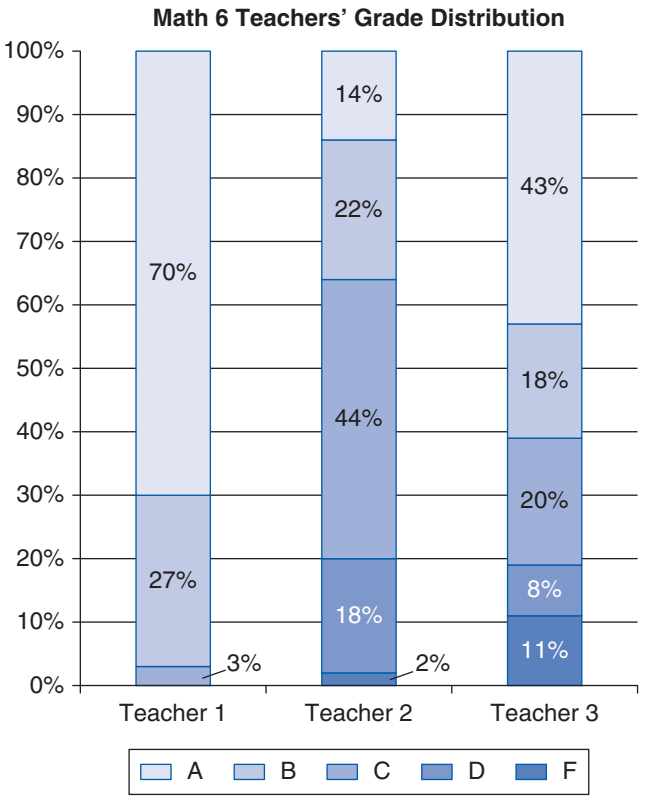
Whether the data she reviewed was "hard" data like test scores or "soft" data like her observations of teacher–student dynamics in classrooms, Mallory kept a sharp lookout for how the school could be made more equitable. Mallory's vision was that students should have equal opportunities for success regardless of their ethnicity, first language, gender, income, or special needs. She paid attention to patterns of unequal achievement or opportunity in her school. For example, were boys being referred more frequently to the office? Were poorer students showing a common weakness on a strand of skills on the writing assessment? Did students who received special education services have a higher rate of absenteeism?

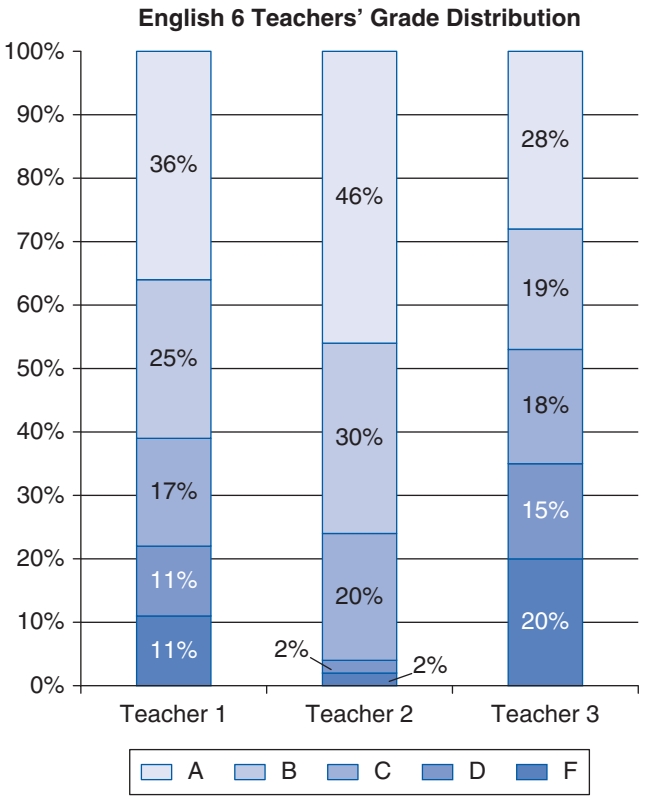
But that wasn't all. To Mallory, one of the most important indications of a high-quality, equitable school is that students are successful regardless of their teacher.

One teacher's students shouldn't learn different material or be less prepared for the next grade than another teacher's students. Fortunately, based on her classroom visits and other data, Mallory found that although teachers approached their work in ways that reflected their individual backgrounds and personalities, students' learning experiences were generally consistent across classrooms. Students in the same course taught by two different teachers—such as Ms. Thompson's and Ms. Richardson's sixth-grade English classes—were learning the same skills, reading the same books and essays, getting the same homework, receiving similar support, and taking the same tests. Mallory was confident that regardless of their sixth-grade teacher, students would be similarly prepared for seventh-grade English.

Since teachers were aligned with what and how they were teaching, and because the school didn't track students or create unbalanced classes where one sixth-grade English class would be stronger than the others, Mallory reasoned that by all accounts the performance of students should be comparable across teachers of the same course. In other words, the rate of As, Bs, Cs, Ds, and Fs in any course should be relatively similar for each teacher of that course. But that wasn't happening. Strange things were showing up in the data.

Take, for example, her school's sixth-grade math and English classes, each taught by three different teachers:





If you were a student in two of the three teachers' math classes you had about a 20 percent chance of getting a D or F, but if you were in the third teacher's math class, you had 0 percent chance of getting a D or F. In the English classes, taught by three different teachers including Ms. Richardson and Ms. Thompson, the range of D and F rates—4 percent, 22 percent, and 35 percent—was even more dramatic. Mallory double-checked the grade data, then double-checked that students in the classes weren't significantly different—in other words, one teacher's students as a group didn't have lower standardized test scores or higher rates of absences. No, the groups of students were similar; the only difference among the classes seemed to be the chances of receiving a particular grade.

Mallory put on her detective hat and considered, investigated, and then rejected several explanations: No substantive differences in instruction. Teachers were using the same curriculum with the same tests and even scored those tests as a team to ensure fairness and uniform evaluation. Mallory scoured students' previous test scores and grades, with no indication of drastically different profiles of the classes as a whole. No substantive difference in the classroom physically—it wasn't as if one classroom had a broken thermostat or was closer to a noisy playground. What was even odder was that students with identical standardized test

scores received different grades depending on their teacher. The teachers were teaching similarly, the students were demonstrating similar achievement, but the grades showed inconsistency. This data seemed unexplainable, impossible, and grossly inequitable.

On a lark, Mallory looked at the syllabus for each class—each teacher of a course had created her own personalized version—and it shocked her. Each teacher's syllabus began with a similar introduction to the course content and description of important materials for the class, but then it was as if each teacher was in an entirely different school:

- One teacher accepted no homework after the attendance bell rang, some deducted points if homework was late (although the amount deducted ranged from a few points to two letter grades' worth), and another accepted work beyond the due date up until the end of the quarter, with no penalty.
- One teacher gave each daily homework assignment a grade of 10 percent or 100 percent based on how much of the homework was completed and correct, and allowed students who had received 10 percent up to one week to correct mistakes. Another gave full credit for an assignment if the student showed effort to complete it, regardless of whether answers were correct.
- One teacher reduced points on an assignment if the student didn't completely and correctly write her or his first and last name, along with the title of the assignment. Another subtracted points if an assignment was submitted on notebook paper that had ripped holes or ripped edges.
- Most teachers organized their gradebook by grouping types of assignments into categories (Homework, Classwork, Tests, etc.), and weighted each category to denote its importance (Homework = 30% of the grade; Tests = 70%). However, no teacher had the same weightings for any categories. For example, the weight of tests ranged from 40 percent to 70 percent of a student's grade.
- Some teachers had only three categories of assignments (Tests, Classwork, and Homework), while others included categories that seemed more subjective, such as Citizenship, Participation, and Effort. There was no explanation in the syllabus of how these subjective categories were calculated or on what they were based.
- Other teachers didn't use percentage weights at all, but assigned different point values to different assignments. For example, Homework assignments might be 5 to 10 points each, with tests worth 100 points.

Teachers' different grading policies made it possible for two students with the same academic performance to receive different grades. What particularly confused and concerned Mallory was that some teachers were grading students on criteria that seemed to have nothing to do with their academic achievement—such as whether their paper had intact holes or had the proper heading—and others were basing

parts of students' grades entirely on subjective criteria, such as effort, that were susceptible to teachers' implicit biases. This grade data that couldn't be possible suddenly was.

A few days later, something happened that changed Mallory's confusion to concern. Maria, a shy but earnest eighth grader, came to her office nearly in tears. Last year as a seventh grader, she had received a B in math, her most challenging subject, but this year was barely passing with a D. What was really frustrating Maria was that even though she often handed in homework assignments late or incomplete—she had after-school responsibilities at home in addition to dance class three times a week—she consistently performed well on every exam. She obviously had learned the math and had shown it when it mattered most, and though last year this type of performance had earned her a B, her teacher this year gave zeros for late or incomplete homework, resulting in her D. Maria was feeling a crisis of confidence: Other students copied to get their homework in on time for the homework points, which Maria had resisted, but would she have no other choice? Had last year's teacher lied to her about her math skills? Was she not as good at math as she thought? Or was this year's teacher out to get her?

To Mallory, no longer were her teachers' inconsistent policies a theoretical dilemma. The school had spent months of planning and coordination to make sure teachers in the math department were using sequenced curriculum and that each teacher was preparing students to be ready for the next year—called "vertical alignment." Yet teachers' different approaches to grading was undermining all of it, sending confusing messages about learning and impacting students' grades and promotion rates, their beliefs about school, and even their self-image.

Mallory had to talk to her teachers about what was happening. The prior year, she had broached many conversations—some quite difficult and uncomfortable—with her teachers about curriculum, teaching strategies, job responsibilities, even evaluation. Surely, she assumed, they would be as astonished as she was when they saw the data and would reconsider how they graded.

But now came her second shock: When she began a discussion of grades with her teachers, it was like poking a hornet's nest. Nothing prepared her for the volatility of conversations about teachers' grading practices. Many of her teachers, previously open to exploring new ideas about nearly every aspect of their work, reacted with defensiveness and adamant justification. Teachers with higher failure rates argued proudly that their grading reflected higher standards, that they were the "real teachers." A teacher with low failure rates explained that he was the only teacher who cared enough to give students retakes and second chances. One teacher simply refused to discuss the topic, citing her state's Education Code that protected teachers from administrators' pressure to change or overwrite grades. One teacher began to cry, confessing that she had never received any training or support on how to grade and feared that she was grading students unfairly. Conversations about grading

weren't like conversations about classroom management or assessment design, which teachers approached with openness and in deference to research. Instead, teachers talked about grading in a language of morals about the "real world" beliefs about students; grading seemed to tap directly into the deepest sense of who teachers were in their classroom.

When she talked about these grading problems with principals of other schools, Mallory was surprised and dismayed to learn that grading varied by teacher in *every* school. This phenomenon was widespread, even the norm. Teachers thoughtfully and intentionally were creating policies that they believed, in their most thoughtful professional judgment, would promote learning. Yet they were doing so independently and often contradicting each other, yielding in each school a patchwork of well-intentioned but ultimately idiosyncratic approaches to evaluating and reporting student performance. Even when a department or a group of teachers made agreements—for example, to have homework count for no more than 40 percent of a grade—teachers' other unique policies and practices, such as whether homework would be accepted after the due date, made their attempts at consistency seem half-hearted and ineffectual.

What's more, even though every principal had the same problems and frustrations with inconsistent grading, no one had any success in addressing it. Other principals had tried to raise the topic of grading and had met the same kind of resistance Mallory had experienced, sometimes even with vitriol and formal allegations of attempted infringement upon teachers' academic freedom.

Mallory wondered: Was inconsistent grading an unavoidable part of schools, like the annoying bells between classes, the complaints about cafeteria food, the awkward physical education outfits, and weak turnout at Open House? Was it an inevitable side effect of teacher creativity, ownership, and initiative? Were teachers' different ways of evaluating and reporting student performance a hallmark of teachers' professionalism or an undermining of that professionalism? And did principals' avoidance of addressing the variance and inconsistency of grading represent support of their teachers, a détente between teachers and administrators, or an unspoken compromise that ignored the damaging impact on children, particularly those who are most vulnerable?

My Own Journey: Frustrations and Hope

In over twenty years of working in schools as a teacher, principal, and district administrator, I've known lots of "Mallory"s. In fact, as a principal I was a "Mallory." Grading among my teachers—my professional, awesome, hardworking, ethical, deeply committed and emotionally invested teachers—was inconsistent. Though as a professional learning community of educators we tackled the challenging topics of relevant curriculum design, high-quality instructional practices, writing across the

curriculum, our racial disparities in achievement and discipline, and, occasionally, our obligation to stand against the historically and culturally hegemonic function of American schools, we couldn't mention grading. Years later, as a district administrator responsible for supporting and coaching principals, I could never convince my principals, much less equip them, to find the language, strategies, or courage to address teachers' grading practices.

I could not agree more with Jeffrey Erickson (2010) who calls grading the "third rail" of schools. On one hand, like a train's third rail, grades provide power and legitimacy to teaching and learning. Grades are the main criteria in nearly every decision that schools make about students. Here are some examples:

- course assignment (eligibility for advanced, honors, or AP classes)
- graduation (completion of course requirements)
- academic awards (valedictorian, summa cum laude)
- extracurricular activities (athletics, clubs)
- promotion (able to progress to next grade level or sequenced course)
- retention (repeating a course or grade level)
- additional supports (mandatory tutoring or remediation)
- additional opportunities (special field trips)
- scholarships
- college admission

Grades inform decisions outside the educational world as well. Potential employers consider grades when hiring, and GPAs are often required for youth work permits and reductions in car insurance, which means students' grades can affect family income and expenses. And those are just the decisions made by institutions. Caregivers and families often provide rewards and privileges (including praise) or enforce punishments and restrictions (including shame) based on grades.

But like a train's third rail, grades are so powerful and important to classrooms and schools that no one dares touch them. As Mallory experienced, the questioning of grading practices by administrators, caregivers, students, and even teachers can invoke anxiety, insecurity, pride, obstinacy, and conflict. And so most of us avoid the topic altogether.

It wasn't until I read a few articles—including "The Case Against the Zero" by Doug Reeves (2004), "The Case Against Percentage Grades" by Thomas Guskey (2013), and *A Repair Kit for Grading* by Ken O'Connor (2010)—that I began to see that teachers use grading for many different, and contradictory, purposes:

1. To communicate the achievement status of students to parents or guardians and others
2. To provide information that students can use for self-evaluation
3. To select, identify, or group students for certain educational paths or programs
4. To provide incentives for students to learn
5. To inform instructional decisions
6. To provide evidence of students' lack of effort or inappropriate responsibility

No wonder that grading practices vary so widely. The teacher who grades to sort students into programs will use grading practices incompatible with the teacher who grades to incentivize students to learn.

And beyond the variation in grading among teachers, I found that many grading practices themselves had deep flaws. For example, I learned that the calculations that we commonly use to derive grades—and often embedded in our grading software—are mathematically unsound.

Secondly, I learned that many of us evaluate students on criteria that are nonacademic and highly susceptible to bias. For example, a teacher who evaluates a student's effort as part of a grade likely applies a culturally narrow definition of what effort looks like.

Thirdly, teachers often use grades for behavior modification, offering the reward or punishment of points and use (or threaten to use) the zero or F to motivate students even though the "motivational F" is largely a myth; research is clear that low grades, or the threat of low grades, do nothing for the student who has low confidence in their academic abilities or limited experience with academic success—the majority of students who receive Fs.

I also learned that our grading often creates "collateral consequences" that contradict our intentions. For example, we lament our students' rampant cheating and copying of homework. Yet when we take a no-excuses approach to late work in the name of preparing students for real-world skills and subtract points or even refuse to accept the work, we incentivize students to complete work on time by hook or by crook and disincentivize real learning. Some common grading practices encourage the very behaviors we want to stop.

As I continued to research and learn more, I realized that the inaccuracy of grades seemed to be only a symptom of a deeper problem. Although I had previously attributed schools' achievement and opportunity gaps of race and income entirely to unaddressed needs in our instruction and curriculum, limited cultural understanding, or a weakness in resolve, I came to realize that our common grading practices make us active accomplices in perpetuating these gaps. The ways we grade

disproportionately favor students with privilege and harm students with less privilege: students of color, from low-income families, who receive special education services, and English learners. For example, we teachers often assign students a zero in the gradebook if homework isn't handed in by the deadline. However, we don't account for all the reasons that a student wouldn't turn something in on time. One reason, of course, might be laziness or disinterest—certainly not legitimate reasons. Perhaps a student has after-school classes or sports, which could make it harder to turn in work on time, but arguably this is a self-inflicted wound. But what if a student's circumstances are beyond her control? What if there isn't a space at home that is quiet enough, or well-lit enough, or not distraction free enough for a child to complete homework? What if a student's caregiver is away at a job (or second job, or third job), so that she isn't around to provide support? What if the parent or caregiver isn't formally educated enough or doesn't speak enough English to help the child complete the homework? What if the child has home responsibilities (caring for an older relative or younger siblings) or has her own job in order to contribute to the family income? What if the student who has few supports simply doesn't know the answers to the homework? What option is there but to submit the work incomplete or late? Clearly, we don't want to grade students based on their environment or situations beyond their control, but unfortunately, when we use grading practices such as penalizing students for late work, that is often what we do.

It was a very depressing and discouraging awakening.

To my relief, I also learned that grading, if done differently, can be accurate, not infected with bias, and can intrinsically motivate students to learn. Grades can clearly and more objectively describe what students know and can do. Grading practices can encourage students not to cheat but to learn, to persevere when they fail and not lose hope, and to take more ownership and agency for their achievement. And the power of these approaches can be especially transformative for struggling students—the students who have been beaten down year after year by a punishing grading system of negative feedback and unredeemable failure.

Yet despite my own research and revelations, knowing how to make grading more accurate and equitable was only the very first step. The real challenge was to understand how teachers could learn, understand, and then implement improved grading. I had to not just touch but embrace the third rail of grading; I had to get others to embrace it with me.

It didn't work out so well at first. When I discussed these practices with teachers, I was constantly met with the same arguments: Our current grading system prepares students for the real world and if we alter it we're doing our students a disservice; "smart kids" can handle changes to grading and can be internally motivated but "remedial" or "regular" students need external motivation; these changes just inflate grades; students will just game the system. Conversations were intellectual jousts that didn't really change what teachers believed or did. Grading was so deeply

intertwined with teachers' belief systems and their daily practices that it wasn't as simple as just explaining and justifying the practices. I realized that for teachers to become convinced of the effectiveness and the equitable impact of different grading practices, they had to try them out. Through a combination of persuasion, promises, and appeals, I found some teachers willing to test out these new grading practices.

Amazingly, it worked.

Teachers who tried these grading practices were surprised and sometimes shocked by the results. The practices seemed to do the impossible: decrease student failures, reduce grade inflation, and reduce achievement gaps—all at the *same time*. Here were the results in one high school:

High School Teacher Cohort: Percentage of D or F Grades Awarded 2015–2016 (before grading initiative) vs. 2016–2017 (1st year of grading initiative)[1]

	2015–2016 SEM. 2	2016–2017 SEM. 2	PERCENTAGE POINT CHANGE	PERCENTAGE DIFFERENCE
Percentage of D or F Grades Awarded	23%	17%	−6	**26% decrease**

In the 2015–2016 school year, 23 percent of the grades that the teachers assigned were Ds or Fs, and fell by over one-quarter, to 17 percent of the grades in 2016–2017. Although this decrease may seem small (and is still too high), because these high school teachers had student loads of 125 to 150 each and assigned thousands of grades every semester, this decrease in D and F grades represents *hundreds* of fewer failed grades, meaning fewer remedial "seats" and therefore less money needed for remedial classes, to say nothing of the long-term impact on graduation rates. What was even more energizing was that the grading practices had a greater (and statistically significant) impact on groups who had been historically underserved in schools. From the same high school:

High School Teacher Cohort: Percentage of D or F Grades Awarded 2015–2016 (before grading initiative) vs. 2016–2017 (1st year of grading initiative)

	2015–2016 SEM. 2	ACHIEVEMENT GAP 2015–2016	2016–2017 SEM. 2	ACHIEVEMENT GAP 2016–2017
FRPL* Students	27%	8% points	19%	3% points
Non-FRPL Students	19%		16%	

* Free and Reduced Price Lunch

[1]The results in these figures on pages xxvi–xxviii were generated by Leading Edge Advisors, an independent evaluation firm.

With these more equitable practices, the rate of Ds and Fs the teachers assigned to students who qualified for free or reduced-price lunch, a proxy for low-income, decreased from 27 percent to 19 percent, while the percentage of Ds and Fs assigned to students who came from higher income families (who therefore did not qualify for free or reduced-price lunch) decreased much less, from 19 percent to 16 percent. The rate of Ds and Fs decreased more sharply for low-income students, meaning that the school decreased their D and F achievement gap between these groups of students from 8 percent to 3 percent.

Here are results at a middle school, where teachers' changes reduced grade inflation *and* failing grades, and narrowed the achievement gaps of income and race:

Middle School Teacher Cohort 2015–2016 (before grading initiative) vs. 2016–2017 (1st year of grading initiative)				
Percentage of A Grades Awarded				
	2015–2016	**ACHIEVEMENT GAP 2015–2016**	**2016–2017 SEM. 2**	**ACHIEVEMENT**
FRPL Students	36%	14% points	31%	9% points
Non-FRPL Students	50%		40%	

Percentage of D or F Grades Awarded				
	2015–2016 SEM. 2	**ACHIEVEMENT GAP 2015–2016**	**2016–17 SEM. 2**	**ACHIEVEMENT GAP 2016–2017**
African American Students	25%	8% points	14%	1% point
White Students	17%		13%	

When teachers used these more equitable grading practices, the disparity in the percent of As assigned to students who qualified for free or reduced price lunch compared to the percent of As assigned to students who did not qualify for free or reduced price lunch decreased by over one-third, and the disparity in the percent of Ds and Fs assigned to African American students compared to white students, which had been eight percentage points, was virtually eliminated.

Of course, it is notoriously difficult to tie changes in student achievement to a specific change in a teacher's practice; student performance and teacher effectiveness are influenced by so many variables inside and outside the school. When teachers at this middle school confidently explained that a primary cause of these changes in student achievement was their improvements to grading and assessment, I wasn't satisfied. I first asked what might be *incorrect* explanations others might give if they saw this data. They quickly responded: "That we lowered our standards; that we

were too soft; that we were pressured to give passing grades." One teacher added, almost adamantly, "Actually, we *raised* our standards. Students no longer can get good grades with fluff assignments."

I believed the teachers, but wasn't yet convinced. I was worried that the practices might yield grades that were improved, but weren't more valid. To determine whether the grades were more valid—that they more accurately and consistently described student achievement—we compared teachers' classroom grades to students' standardized test scores. We found that teachers' grades had an increased correlation to standardized test scores. Not only were grades less inflated or deflated, they were also more *accurate*:

State Test Score Results vs. Sem. 2 Grades Assigned Spring 2016 (before grading initiative) vs. Spring 2017 (1st year of grading initiative)				
	SPRING 2016 SEM. 2	SPRING 2017 SEM. 2	PERCENTAGE POINT CHANGE	PERCENTAGE DIFFERENCE
Percentage of Students for Whom State Exam ELA Score **MATCHES** Teacher-Assigned English Sem. 2 Grade (ex.: 3 = B, 2 = C, etc.)	34%	48%	+14	**41% Increase**
Percentage of Students for Whom State Exam Math Score **MATCHES** Teacher-Assigned Math Sem. 2 Grade (ex.: 3 = B, 2 = C, etc.)	21%	38%	+17	**80% Increase**

Although in 2016, before teachers used more equitable grading practices, only about one-third of semester 2 English grades matched standardized test scores in English, after teachers used the practices in 2017 nearly half of teachers' English grades matched the test scores, and the percent of semester 2 math grades that matched standardized test scores in math nearly doubled. And even though there are plenty of reasons to be skeptical of standardized tests, we'd prefer teachers' grades to be more correlated with external test results than less correlated.

Beyond the quantitative data, the impact of these more equitable grading practices on the day-to-day work of teachers and students was even more transformational. Students were relieved and grateful to not have everything "count" in their grade, to have flexibility to turn in assignments after a deadline, and to be allowed to retake exams. Teachers felt the emphasis in their classrooms had shifted from meeting due dates and earning points to *learning*. Students completed assignments because they found that doing so improved their performance on assessments, not because of the homework or classwork points they could earn or lose.

What's more, teachers felt empowered by this work. Prior to this work, the ways their students behaved—what motivated them, whether they cheated or not, how much they understood or cared about their grade—had seemed to the teachers to be fixed and often chalked up to "that's how kids are these days." But the teachers who tried these practices found that they could actually change students' attitudes and behaviors. Students who had seemed unmotivated and even resistant to learning became more engaged. Relationships between students and teachers—which had been based on compliance and a system of extrinsic threats and rewards—were now partnerships based on trust, transparency, and, perhaps most importantly, hope. Students persevered when they struggled, took initiative, stopped cheating, and wanted to learn even after the test—all because of changes to how teachers graded. After using these new more equitable practices, these once skeptical teachers had the passion of religious converts. Cathy, a middle school history and English teacher, was typical in her reaction:

"I have a different outlook now on how I want to grade and how I want to use it. Last year it was almost punishment: 'Oh, you didn't do the work, now you have a bad grade.' Doing this work really changed my perspective. This helped me realize that the main purpose of grading is to see how much the students know, to assess their learning instead of assessing their efforts; do they really understand the work, as opposed to did they do all of the assignments."

Plus, this work to improve grading didn't just change how teachers graded. It changed their beliefs about themselves, about teaching and learning, and about their students. They discovered that they didn't need to give points for assignments to make students value and complete the work. They found that they were just as respected, and more trusted and appreciated, by their students when they changed their grading. Most powerfully, they learned that by changing how they graded, their students—whether elementary children, middle school tweens, or high school teenagers, and whether overachieving or struggling and resistant—would take ownership and responsibility over their learning, would be intrinsically motivated to succeed, and would be excited about learning and their own progress.

Over the past several years, we have seen these benefits of equitable grading in many school types and environments: at large comprehensive district-run schools, charter schools, and independent schools; at schools with only white students and those with only students of color; at schools nestled in urban centers and located in suburbs; and at schools with students who enter with skills far below grade level and at Phillips Andover Academy, one of the most elite boarding schools in the country. But regardless of the school's context or its student population, this work was hard. Examining our grading practices can challenge our deepest beliefs about what we

know (or think we know) about our teaching, our students, and ourselves. Lucy, an eighteen-year veteran high school English teacher, best expressed the difficulty of considering changes to longstanding grading practices, and why the experience can be so transformational:

"This challenges what I've learned to do as a teacher in terms of what I think students need to know, what they need to show back to me, and how to grade them. This feels really important, messy, and really uncomfortable. It is 'Oh my gosh, look what I've been doing!' I don't blame myself because I didn't know any better. I did what was done to me. But now I'm in a place that I feel really strongly that I can't do that anymore. I can't use grading as a way to discipline kids any more. I look at what I have been doing and I have to do things differently."

Lucy's description captures it all: Examining grading is "important, messy, and uncomfortable." It can be difficult to amass the energy and resolve, particularly with all the mandates and sky-high expectations placed on teachers, to make grading more accurate and equitable. But it is some of the most important and rewarding work we can do. We know that students' family income, whether they have a stable, safe home (or even a home at all), their caregivers' education background, their race, and other elements outside teachers' control all have a huge influence on achievement, but at the end of the day, it's their grades—our description of students' academic performance—that opens doors or closes them. And though we can learn a new curriculum or a new instructional strategy, but if our grading doesn't change, nothing for our students, particularly those most vulnerable, will really change, and the achievement and opportunity gaps will remain.

It's time to embrace the third rail.