

**Elements of Grading:
A Guide to Effective Practice**
By Douglas Reeves (Solution Tree Press, 2011)

S.O.S. (A Summary of the Summary)

The main ideas of the book are:

- ~ This book shows that by improving your grading system, you can actually increase student performance.
- ~ To make your grading system more effective, it needs to be accurate, fair, specific, and timely.

Why I chose this book:

Grading is one of those very emotional subjects in education that has prevented traditional methods from changing. It's time that educators become aware of the distortions in our current system and learn alternative ways of approaching grading.

The book helps school leaders understand the reasons why current grading practices are harmful and inaccurate and it also introduces some simple strategies for improving grading methods (such as getting rid of zeroes for missing work).

It's hard to argue with grading suggestions that encourage students to do more and better work while saving the teacher time and headaches at the same time!

The Scoop (In this summary you will learn...)

√ *The importance of making grading systems accurate*

Why grade averages might be mathematically accurate but not reflect a student's current ability at all.

√ *The importance of making grading systems fair*

Why a single letter grade often means that students with vastly different performance levels all achieve the same grade!

√ *The importance of making grading systems specific*

How a standards-based report card gives students much more specific feedback that they can use.

√ *The importance of making grading systems timely*

Why the difference between two pieces of writing – one labeled “poor” and the other “gifted” – from the same student may be due to timely feedback.

√ *How school leaders can conduct productive conversations about grading with teachers*

Introduction, Chapter 1, and Chapter 2

Grading is a crucial issue in schools today. Research has shown that *feedback* is one of the most influential factors affecting student learning. As John Hattie, Robert Marzano, and other researchers have shown, feedback is one of the most powerful instructional techniques in influencing student achievement. In the same way that athletic coaches and music teachers influence the performance of their students through feedback, teachers have the potential to impact their students' performance tremendously when they provide appropriate feedback. However, not just any type of feedback is effective. In order to be effective, feedback needs to be accurate, fair, specific, and timely.

Grading is only *one type* of feedback, but it is the one that gets a great deal of attention. If a school has an excellent system of feedback but has *ineffective* grading practices, this can undermine efforts to improve student achievement. Reeves argues that we need to do a better job of basing grades more on student *performance* than on other more subjective factors less related to achievement. In order to do this, we need a grading system that is:

- (1) Accurate – The grade must reflect the academic performance of the student.
- (2) Fair – The grade must not be influenced by factors such as gender, ethnicity, socioeconomic status, etc.
- (3) Specific – The grade must be specific enough to provide information to help improve the student's performance.
- (4) Timely – Students must receive a steady stream of feedback to help them improve their performance.

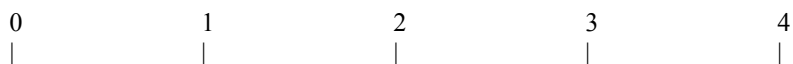
It is important to note that feedback is useless if it does not meet *all four* of these characteristics. For example, if feedback is accurate but not timely, this will not improve performance. An autopsy can be accurate, but it clearly comes too late to revive the deceased. Feedback can be timely – like in a video game – but if it provides no specific guidance as to how to improve performance, then it also may be futile.

So, how do distortions in accuracy, specificity, or fairness occur? People typically assume that grades faithfully represent student performance. In fact, some feel absolutely certain about this. This is why discussions about grading can become heated. In fact, grades can be subject to significant distortions that have nothing to do with performance. Below are two examples that illustrate this point. For example, you would find a wide range of resulting final grades if you asked your colleagues to determine the final grade for a student who has received the following grades: C, C, MA (missing assignment), D, C, B, MA, MA, B, and A. When Reeves gave this task to more than ten thousand teachers and administrators, he got the following:

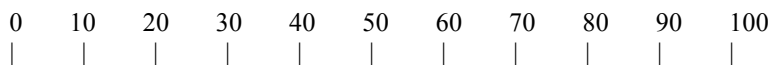
- 7 percent assigned an A
- 13 percent assigned a B
- 39 percent assigned a C
- 21 percent assigned a D
- 20 percent assigned an F

This shows the completely idiosyncratic judgment used by teachers and administrators in assigning this grade. This one student above would be given an entirely different grade depending on the teacher he or she had. Rather than work ethic or intelligence being the factors that lead to an A or an F, the grade is the result of teacher or administrator judgment.

Another way grades are distorted comes from the common practice of giving students a zero when they do not turn in their work. Look at the contrast between the two different weights given to a grade of zero. If teachers use a 4-point scale (4 = A, 3 = B, 2 = C, D = 1), then the distance between a zero and getting a D is the same distance as the difference between a D and a C and all the other intervals. You can see these two segments are the same length below:



But if teachers using a 100-point scale give a student a zero for not turning in work, the distance between the zero and the D is 60, or six times as much as the distance between a C and a D (see the scale below). So not turning in the work is six times worse than doing an awful job on the work!



Even the most well-intentioned grading system has distortions. No one can create a system with perfect accuracy and complete fairness. The goal is to begin to have constructive discussions with school stakeholders in order to move grading policies more in this direction. This is not easy to do given that historically grading has been an emotionally, politically, and financially charged issue. The rest of the book will explain where grading distortions come from and how to begin to make improvements in grading practices.

Chapter 3: How to Improve ACCURACY

Grading Inaccuracy

Some of the ways grades can be inaccurate were described above – the assignment of a zero in a 100-point system for missed work and also the greatly varying methods teachers use to come up with final grades. In addition, although *averaging* grades may make perfect mathematical sense, that is, averages can be calculated accurately, the use of an average actually *undermines accuracy in grading*. For example, if you wanted to know whether to wear a coat on January 31 you would not determine the average for the month, you would check the temperature on that day. In the same way a thermometer gives us information about the current temperature, a grade should tell us what a student knows and can do at the *current* time, not what he or she knew a month ago. A grade is only accurate if it reflects what the student knows and can do when the grade is assigned. If a student has steadily increased her performance throughout the quarter – starting with poorer grades and working her way up to higher grades – her final grade should reflect the level she has achieved at the end of the quarter, not an average of all of her work along the way. In fact, there is more of an incentive for students to improve if, as the result of receiving feedback and improving performance, they will attain a higher grade at the end of the marking period.

Another reason that averages are inaccurate is that the same grade can be assigned to students with widely varying levels of performance. Imagine two students who receive a B – one as the result of slow and steady effort up from a D and another whose levels of performance are in constant flux. Should students with such different levels of performance receive the same grade? If so, the grading system must be inaccurate. While an average may be deceptively precise, it is not an accurate reflection of student performance.

Practical Strategies to Improve Accuracy

While it may be impossible to have a completely accurate system of grading, below are some techniques to improve grading accuracy.

Reality Checks

There are ways to do a reality check to see if the grades given to students are accurate. The first way is to compare the grades to an external measure. While there may be no external exam for most classes, for those who teach International Baccalaureate (IB), Advanced Placement (AP), Cambridge International Examination (CIE) or other similar courses can compare the results of class grades with the results from these exams. There are a number of reasons why discrepancies might occur, but one possibility is the inaccuracy of the grading system.

The most common cause for the discrepancy between grades and actual student performance is the result of including behavior as a factor in grading. For example, girls are often rewarded for more quiet and compliant behavior. As a result, girls are overrepresented in honor rolls and their matriculation to college. Conversely, poor behavior often results in lower grades as well. The problem is that teachers mistakenly label behavior as “algebra” or “chemistry.” To ensure that grades reflect student performance, teachers can use academic performance scoring rubrics that focus exclusively on the *knowledge* and *skills* of students.

Collaborative Scoring

One of the best and most practical ways to improve grading accuracy is through collaborative scoring of student work. This is best done with the use of a protocol, actual student work with the names removed, and a clear scoring guide to grade the work. Teachers individually grade the student work and then have a discussion about their disagreements. There are a number of benefits to collaborative scoring. First, by spending time practicing grading with colleagues, teachers become more accurate and more efficient at grading. Second, it is a fantastic professional learning opportunity. Finally, by doing this as a team, the teachers become more consistent so that students are graded more accurately across the school as well as more accurately within their class.

Avoidance of Unintentional Mathematical Distortions

One way to avoid mathematical distortions is to make sure teachers are accurate when assigning zero for missed student work. Earlier the book showed how assigning a zero is inappropriate in a 100-point grading system. Since there are strong reactions to giving students 50 points for missed work (that would be the same interval from a “zero” to a D as a D to C), to avoid this distortion, schools should consider switching to a 4-point scale.

Furthermore, to avoid the distortions that come from assigning an average for a grade (discussed earlier), schools should find a way to weigh end-of-year assignments more heavily than earlier ones. This will more accurately reflect how students are *currently* performing when they receive their grade.

Chapter 4: How to Improve *FAIRNESS*

What is Fairness?

No matter how *accurate* grades are, this will not suffice if the grades are not *fair*. In order for grades to be fair, they need to reflect student performance and not be inappropriately influenced by other nonacademic factors such as gender, ethnicity, and socioeconomic status. Being fair is about more than having rules and enforcing them equally. Most of us think we are being fair, but it does not always turn out that this is the case. For example, there are vastly more girls than boys on the honor rolls in middle and high schools. Unless girls are significantly smarter than boys, it is likely that these scores have been inflated based on gender and do not merely reflect performance.

In fact, it is hard for a lone grade to completely reflect one student's performance. Grades often end up reflecting other nonacademic factors. For example, the same grade of B- might be given to a wide variety of students for different reasons such as:

- Outstanding effort and perseverance, but the student has not yet met grade-level standards
- Outstanding performance above grade level, but the student's attitude, work ethic, and class participation are inadequate
- Superior performance, except for one incident of cheating that resulted in a score of zero on a major exam
- Failure to meet all academic standards, but the student earned several extra credit points to merit the final grade

If a college admissions officer or a parent were looking at the report card, how would they know what the B- really means? How can a B- fairly represent all of these different situations without any of the explanations or context provided?

Why Does Fairness Matter?

Fairness matters for several reasons. On a basic level, when educators give grades that fairly represent performance, then students, parents, and other stakeholders trust those grades more. If our grading systems are inaccurate, then those stakeholders feel cynical and distrustful. Furthermore, if the grading system is perceived to be flawed, then people begin to rely more on standardized tests as measures of accurate student performance assessment. Finally, when grades are fair, students feel more engaged and motivated to succeed. The opposite is true as well – unfair grades lead students to withdraw and lose motivation to improve.

So, how can you tell if the grading system in your school or classroom is fair? Reeves provides a handout, the "Equity in Grading Self-Assessment" on pages 127-130 in the book or available for free at <http://go.solution-tree.com/assessment> for teachers and administrators to examine their own grading systems in a nonjudgmental way. Below is an excerpt:

Equity in Grading Self-Assessment				
The purpose of this form is to help teachers and administrators explore the relationship between student grades, academic performance, and nonacademic factors.				
1. Treasure Hunt: Find the A's				
Look at the most recent report cards and enter the names of the students who earned A's, or, if you have a different system, enter those students' names with the highest available mark.				
Name	Gender	Low Income?	Ethnic Minority?	Language Minority?
1.				
2.				
Etc.				
What do you notice about the profile of your A students? What demographic characteristics do they have in common?				

A school might go through this exercise and find nothing, but educators often find the following results:

- Students with higher honor-roll grades but who also have low scores on external tests are more often female. In urban systems these are most often minority female students.
- Students with lower grades but higher scores on external tests tend to be male. In urban settings, these students tend to be minority male students.

While some of these might strike you as offensive overgeneralizations, it is important to take an honest look at the patterns. Are there more girls than boys on the honor roll? What is the ratio of girls to boys who failed classes? What is the gender ratio in remedial and advanced classes? Based on the patterns that emerge, are students at your school being graded on compliant behavior rather than pure academic performance?

Chapter 5: How to Improve *SPECIFICITY*

Even if you manage to implement a grading system that is both accurate and fair, to be truly effective, it must also provide specific and timely feedback. This chapter focuses on specificity. Although specificity is a key ingredient in an effective grading system, most grading practices involve taking a complex set of factors into account and reducing them to a single number or variable. How can a single letter or number communicate clearly about a student’s performance? Although teachers often spend a great deal of time on their grades, ultimately these grades mean a lot more to the teacher than to the audience for whom they are intended. This is not unlike the idea of a thoughtful and well-intentioned teacher providing feedback written in ancient Egyptian hieroglyphics – it is unlikely the student would learn much from this feedback. We need a grading system with a higher communicative value for students, parents, fellow teachers, and other school systems.

The Standards Achievement Report

One of the biggest issues in grading effectively is the degree to which grades should be limited to academic ability and how much they should include other characteristics such as work habits and personal character. Should a grade for “math” include effort, participation, attitude, and behavior? One way to address this concern is to grade students on *both* academic performance and the other work habit traits, but keep these grades separate. To do this, schools can use a version of a standards achievement report. In this sample report below, rather than a student receiving a “B” for a lab report, the student gets a grade for each academic standard the lab report covered. Then, to separate out behavior, there are separate columns to provide grades for timeliness and teamwork. This type of report makes it clear that this student is either Proficient or Exemplary in most of the standards, but needs a lot of assistance in learning to turn in work on time. See the excerpt below (for the full report, see p. 65 in the book):

Standards Achievement Report							
Student Name:	Class:				Teacher:		
E = Exemplary; P = Proficient; IP = In Progress; N = Not Meeting Standard							
Assignment	***** Standard *****						
	1. Biodiversity	2. Genetics	3. Cell Structure	4. Environmental Interrelationships	5. Explanation of Scientific Conclusions	6. Teamwork	7. Submission of Work on Time
Lab 1	P			P	P	E	N
Analytical Paper 1	P				P		IP
Lab 2	P				P	E	IP
Analytical Paper 2		P			P		IP
Etc.							

Note that this report need not *replace* the traditional letter grade. Many schools feel compelled to keep that system. In that case, the standards achievement report can be given *in addition* to the single letter or number grade. If you use a standards achievement report to capture student performance more specifically, how do these grades translate into a letter grade? Below is a suggestion (note that **E** = 4, **P** = 3, **IP** = 2, and **N** = 1). The report above is an excerpt, but in the original, there were 14 assessments (papers and labs) in total:

- A = At least four assessments with a final score of 4 and two assessments with a final score of at least 3
- B = At least four assessments with a final score of at least 3 and two assessments with a final score of at least 2
- C = At least three assessments with a final score of at least 3

A grading system structured like this carries with it some implications. First, the use of the term *final score* suggests that what is valued is that students submit their work, respect the feedback from the teacher, and then make improvements. What is important is not receiving a 4 on a first draft, but *ultimately* receiving that 4 as a result of putting in the hard work and improving bit by bit. The emphasis is on earning enough Proficiency and Exemplary grades to show that a student has mastered a standard. The lower grades are not averaged in to the final grade because they are not being used to assign a final grade. Instead, they are used to provide feedback on student learning along the way.

Another implication of this type of report is that there is no grade of D, or as Reeves calls it, “the coward’s F.” Previously, a D has encouraged the least motivated students to do the least amount of work. A grade of D has been used to send the message, “You really don’t have to listen to the teacher, resubmit the assignment, or redo the work. Just do enough papers, quickly and badly, and we’ll let you skate through to the next grade.” Some defenders of the traditional grading system would contend that in “real life” people are not given multiple opportunities to succeed, “You know it or you don’t.” However, if you speak to most people in most professions (from physicians to pilots to plumbers), most will tell you that they get more than a one-shot attempt at success; instead, they try something, get feedback, and then work on improvement. Teachers should require that students revise work rather than simply accept “D” work.

A Note on Grading Behavior

Traditionally, when teachers have wanted their students to improve their behavior, they included behavior as part of their grades as a form of reward or punishment. For inattentiveness, tardiness, incomplete work, or disrespect the response was to dole out a “C Minus!” However, this has rarely worked to improve student behavior. If teachers want to improve their students’ behavior, first they

need to start with specificity – outlining precisely what behavior needs to change. One way to provide the type of clarity and specificity when giving students feedback on their behavior is to use some type of rubric. Below is an example of a 4-point spectrum of performance for one aspect of behavior, *attentiveness*:

Level 1. You are asleep, distracted, conversing with others, playing with electronic devices, or otherwise disengaged from the class. You are showing me that you do not even care about the class, your fellow students, or me as a professional educator.
Level 2. You are pretending to pay attention, but we both know it is only superficial engagement. You look forward and have your book open, but you are not participating actively in individual and group work.
Level 3. You are seated when the bell rings and have your book and papers ready. You volunteer to participate in class and group activities. You ask questions and contribute actively to class discussions.
Level 4. You take an active leadership role in the class, noticing when other students need help and encouragement. You regard it as a personal mission to help other students move from “Level 1” to higher levels of engagement that you know will lead to better success for the entire class.

In addition to being specific about the desired behavior, teachers need to avoid the same trap of averaging behavior grades and including zeroes for tardiness and incomplete work done several months ago. If a person who has been on a strict diet but binges on dessert for New Year’s Eve, and then goes back to that diet from February through May, would a physician say, “I’d give you a clean bill of health and encourage your progress, but I just can’t get over the weekend you ate that ice cream.”? If students know they are going to be punished in their final grades in June for behavior committed earlier in the year, they lose motivation, “What does it matter? I can’t win. There’s nothing I can do. I might as well give up.” Instead, teachers need to provide clarity and specificity around behavior expectations and a grading system that does not unfairly punish students and erode their motivation.

Chapter 6: How to Improve *TIMELINESS*

The Importance of Timeliness

In addition to providing feedback that is accurate, fair, and specific, schools need to work on the final element of effective grading: improving *timeliness*. However, in most school systems, the feedback that is most important – the data that leads to rewards and sanctions – often arrives months after assessments are given. This makes it impossible for this feedback to affect student performance. However, there are examples when feedback is extremely timely and has a profound impact on performance. For example, watch any great athletic coach or music teacher in action and you will see students receiving feedback immediately, and in fact improving their performance on the spot, without even having to wait for instruction to end. In another example, Professor Lucy Calkins distributes two pieces of writing to her audience of educators – one piece is considered “poor” and the other is deemed a “gifted” writing sample. However, these were *not* written by two students with greatly varying writing abilities. Instead, both were written by the same student. However, one was written *without* timely feedback and the other was the result of feedback that was given with enough time for the student to improve performance.

When Reeves taught a graduate level assessment and research class he would take his students to observe the school’s basketball team coached by Craig Ross. He would tell his students to observe for the following:

- The percentage of players who received feedback
- The frequency of feedback for each player
- The nature of the feedback – positive, negative, challenging, prescriptive
- The impact of the feedback – did the feedback lead to improved performance?

This is what the students observed:

- 100% of the players received feedback.
- During each 15-minute quarter, each player received feedback between five and fifteen times, depending on the player.
- The nature of the feedback varied depending on the player – some needed to be challenged while others needed encouragement.
- The impact was clear – despite being a small school with no basketball scholarships, due to the coach’s feedback, they beat teams that were bigger, stronger, and better.

When the students were asked to observe for feedback at a class in *school*, they found quite different results:

- Often the feedback was positive and only provided for the eager few who raised their hands.
- Feedback was frequently given to only a few students.
- Feedback was more binary – great/awful, right/wrong – rather than specific or tailored to the student.
- The impact was difficult to tell because there was no evident score keeping and the report card was nine weeks away.

It is important to note that there was nothing outstanding about Coach Craig Ross. In fact providing feedback in every class that is specific, differentiated, and impacts student performance, is not impossible.

Ways to Improve Timeliness

Below are three suggestions for ways teachers can improve the timeliness of their feedback to students.

(1) Involve Students in Establishing Academic Criteria

The fastest way to give feedback to students would be *before* the teacher even sees the work product. How is this possible? By going beyond clearly outlining what students need to know and be able to do in a scoring rubric, and allowing students to help create that rubric. While this may seem threatening at first, if students are involved in creating these scoring guides they are more likely to have a clear idea of the different levels of performance. You can observe students internalizing complex rules and directions when playing games, “You can go here, but you can’t go there. You can do this, but you can’t do that, but if you see someone else do this, then you get to do that.” The same is true for academics. If they help create the rules of the game, then students can better evaluate their own performance and the performance of peers.

(2) Use the Three-Column Rubric

Another way to improve timeliness is for teachers to use a three-column rubric. The first column contains the performance criteria, the second is for student self-assessment, and the third is for teacher assessment. When students submit their work, they turn in this rubric with the second column already filled in. This saves the teacher time because the teacher only focuses on the performance criteria in which the teacher’s assessment differs from that of the student. The student’s self-assessment provides a shortcut for the teacher and prevents the teacher from having to spend too much time on the mechanical and low-level feedback concerning issues or problems that the student already noticed. By saving hours of grading, the teacher can turn around the assignment and get feedback to the student much faster. A sample of this rubric is below (see p.88 for the full rubric):

Scoring Guide	Student Assessment	Teacher Assessment
1. Organization	1.1 <i>My outline included two levels of structure, including Roman numerals and letters.</i> 1.2 <i>My outline matches the content of my written and oral presentation.</i> 1.3 <i>My outline follows a logical progression, with facts to support my arguments.</i>	1.3 <i>I don't understand why global warming leads to famine in Africa – could you please elaborate on this?</i>
2. Research	Etc.	
3. Written Arguments	3.1 <i>My written presentation follows the outline structure.</i> 3.2 <i>Each sentence is complete, grammatically correct, and every word is spelled correctly.</i> 3.3 <i>Each section of my paper flows logically to the next, and every paragraph has a logical transition to the next paragraph.</i>	3.2 <i>Your writing is excellent! Please see the places where I circled incomplete sentences and rewrite them for the final draft.</i>
4. Graphics	Etc.	
5. Oral Presentation	Etc.	

(3) Offer Mid-Course Corrections

A third way for teachers to provide timely feedback is to do so *before* a project is completed. This is particularly important with larger projects or assignments. It is *least efficient* for teachers to collect all large projects on one day, spend countless hours grading them, and then, because this takes so much time, return them to the students weeks later with a grade. Instead, it is a more efficient and effective way to give feedback if teachers identify and highlight problems along the way. Rather than the well-intentioned offer to “come see me if you need help” teachers need to structure the time in class for regular feedback. This involves creating a calendar with time allocated to give two to four students a day feedback. Because some students will struggle more with the assignment, they may receive more frequent feedback than other students.

How School Administrators Can Help Timeliness

Although a number of schools are engaging in what they call “formative,” “interim,” or “benchmark” assessments, this does not mean that teachers are providing timely feedback to students. There is a lot that the school leader can do to profoundly enhance the timely use of feedback to students.

First, school leaders can commit to keeping assessments brief. Assessment experts will claim that a certain minimum number of assessment questions are needed to adequately assess a topic. Although that may be statistically true, principals need to keep in mind the impracticality of administering two-hour tests and then scoring them in a timely fashion. Ideally, by focusing on the most important standards, assessments can be short. This will allow teachers to score and analyze them on the same day so they can apply feedback the very next day. It is the job of the principal to make wise choices in order to focus assessments on what is essential so teachers can provide timely feedback.

The second way school leaders can support timely feedback is to provide teachers with the necessary *time* to analyze and use the feedback. For example, they can plan an early dismissal on days when formative assessments are given so teachers have time to analyze the results and plan for instructional improvements. Furthermore, rather than using professional development time for announcements and administrative items, educational leaders can reserve that time for focusing on student learning.

The third way educational leaders can encourage timely feedback is by monitoring teachers' responses to student assessments as part of their evaluations. This is not to say that school leaders need to evaluate teachers based on the test results themselves, but rather on how teachers respond to those results. Do teachers give students more time on literacy or other appropriate interventions in response to assessments?

Chapter 7: Time-Saving Strategies for Busy Teachers

To support teachers with the goal of providing timely feedback, this chapter presents several strategies for helping teachers save time when grading. In fact, the most *effective* grading systems actually save time for teachers thus serving both the interests of students and teachers at the same time. Below are some suggestions for more effective grading practices that also end up saving the teacher time.

An Alternative to the "Quick" Zero

Introducing a new approach to grading will fail if teachers believe it will be more time consuming. Of course teachers assume that a new grading system that gets students to do more and better work will result in a greater time commitment for them. Consider the example of giving a zero for missed work. What could be quicker than that? With the single entry of that zero in the gradebook, the teacher is done. However, there are several time-zapping consequences that teachers may not anticipate:

- A student with numerous zeroes likely needs the teacher's help and may end up using a lot of the teacher's free time during breaks, before, and after school for assistance.
- A student receiving zeroes is likely to approach the rest of the year in an angry, unmotivated, and disruptive way.
- If repeated zeroes lead to failure, the teacher may spend a great deal of emotional energy on that student if he or she repeats the class the next year.
- Zeroes and failures can lead to lengthy discussions with parents and administrators.

One alternative to the quick zero is to change the consequence so that students actually have to complete the work. At the elementary level this may mean the student misses fifteen to twenty minutes of choice time to complete the work. Students who choose not to complete the assignment when it is due are choosing to lose their freedom during a free-activity choice time. For secondary students this time can be taken from advisories, study halls, lunch, before and after school. "Randall, it appears that you've made a choice today to sit at the quiet table at lunch and complete your homework." This works for a surprisingly large number of students. In one comprehensive high school, as part of a larger reform for failing students, this reform resulted in *more than one thousand fewer course failures* and a significant reduction in discipline and behavioral issues.

The "Menu System"

The "menu system" is another alternative to the quick zero. It is an approach to giving assignments in which students take more responsibility for their work because they set their own goals and track their own progress. When they fail to complete an assignment, rather than receive a zero or offer excuses, students choose another assignment from the menu. Below is an excerpt from a sample menu. The entire menu is on p.98.

Student Name:
To earn an A you need 900 points and to earn a B you need 800 points. You must work sufficiently hard to earn one of these grades. For each assignment you will have a rubric with ratings of exemplary, proficient, progressing, and not meeting standards. You will only receive credit for assignments done at the exemplary or proficient levels.
In our colonial America unit this quarter you must choose at least one assignment from each category: government, geography, economics, and culture.
1. Government
1.1 Research project (300 points for exemplary, 240 points for progressing)
1.2 Book review (100 points for exemplary, 80 points for progressing)
1.3 Article review (50 points for exemplary, 40 points for progressing)
1.4 Electronic game creation (200 points for exemplary, 160 points for progressing)
1.5 Unit tests (100 points for exemplary, 80 points for progressing)
1.6 Concept map of unit test content (50 points for exemplary, 40 points for progressing)
2. Geography
2.1 Etc.
3. Economics Etc.
4. Culture Etc.

Chapter 8: Leading Change for Effective Grading Practices

Regardless of how rational the arguments are for creating grading practices that are accurate, fair, specific, and timely, being right is not enough. It is difficult to implement changes to a grading system when the emotion, indignation, and indifference of colleagues gets in the way. This chapter presents a thoughtful three-pronged approach for implementing changes to grading practices.

1. Revisiting the Purpose of Grading

There is a wide variation of what teachers believe to be the purposes of grading. Educators' beliefs about the goals of grading usually fall into the following categories:

- Giving feedback to improve student performance
- Reporting to parents on student progress
- Communicating with teachers at the next level of instruction
- Giving rewards to students for good behavior and attitudes
- Assigning punishments to students for poor behavior and attitudes
- Making public distinctions between good and bad students

While most agree on the top three purposes for grading, the bottom three categories find their way into many discussions about grading. Consider the following examples from teachers and school administrators:

- "I know that she didn't meet the academic standard, but how else I am supposed to recognize that she tried hard, came to class, and finished her homework? Honor roll grades are the only recognition that poor kid will ever get!"
- "He's not even sorry for missing his assignments! His attitude is casual and contemptuous. I don't care if he aced the end-of-course exam – he's failing my class until he shows some contrition."

However, the research shows that using grades as rewards and punishments does not work. In fact, when students are rewarded *just* with comments and narrative feedback, their performance improves more than when they are graded. Furthermore, when students receive zeroes, have lower grades averaged into their final grades, and are not allowed to resubmit work, rather than motivating them to do better, these policies send the message that no matter what you do you cannot improve your final grade, so you might as well give up! School leaders need to emphasize that this is not the message we want grades to send to students and that everyone should agree with the following two ideas:

*** *The primary purpose of grading is feedback to students to improve performance.*

*** *Rewards and punishments for attitudes and behavior are not acceptable purposes of grading.*

2. Implementing Unpopular Changes

Educational leaders often need to make policy changes that are unpopular. Rather than waiting for the majority to buy in to the new policy, progress in education is often the result of visionary leadership. Otherwise progress -- such as ending segregation or the use of evidence-based practices – might never occur. To modify a grading policy – an often unpopular policy change – leaders can follow the four-level action-oriented change model presented below.

A. Explicit Vision

The leader needs to create a very clear and explicit vision of what the school would look like with the new grading policy in place. *Do not* begin with a discussion of grading policies. This will just lead to hardened positions. Instead discuss two futures – one in which the current grading policy is in place and a second in which there are *fewer* student failures and *more* successes in the school. Have colleagues discuss the following in order to paint a compelling vision of the future with the grading changes in place:

- How would our school be different if we had fewer student failures?
- What would it mean to have fewer students in each class *repeating* that class?
- If we did not have to devote so many resources to students failing classes, how else could we use those resources?
- If we had fewer failures, suspensions, expulsions, and fewer low-level discipline problems, how would our professional and personal lives be better?

B. Specification of Behavior

After painting a specific and compelling vision, the leader does *not* say, "I want you to *believe* this way." Instead, the leader lays out specific behavioral expectations for implementing the changes like the following:

"We agreed that it is essential to reduce course failures to achieve a vision that includes better student success, improved discipline, and a significantly better professional environment for the faculty. In order to achieve that vision, we are making two clear changes. First, the consequences for missing or low-quality work will no longer be grades of zero or F, but rather a requirement that students complete the work. We will collaborate to create time, support, and appropriate consequences, including time before, during, and after school, for students to finish their work. To be clear, we will no longer use the zero on a 100-point scale. We will no longer use the average to calculate the final grade..."

C. Assessment of the Degree of Implementation

The next step is for the leader to assess the degree to which the grading policy changes have been implemented. It is usually not the case that people either implement a new change or they don't. Rather, there is often a range of how successful the implementation has been. Below is an example of how the school leader might assess the success of the grading policy implementation:

Level 1 – The teacher received the professional development on grading policy but there is no evidence of implementation – use of zeroes, averages, and grades as punishments and rewards still exist. There is also no evidence of students using feedback to improve performance.

Level 2 – Classroom grading practices have been somewhat revised. Late work and revisions to low-quality work are accepted, but all work is done at home and there is little evidence of students using feedback to improve performance.

Level 3 – Class grading policies meet school standards and there is clear evidence of in-school time allocated for students to improve and complete work. There is evidence of an improvement in the quantity and quality of student work along with a decline in student failure. Improved academic performance is leading to an improved classroom environment.

Level 4 – All grading criteria has been met, there are innovative structures to encourage higher levels of student performance, and the teacher is experimenting with assignments with no grades at all.

D. Continuous Refinement

Different staff members will have the same grading policy, the same professional development, and the same tools, and yet implement the changes differently in their classes. The primary reason for this is a failure on the part of the leader to specify, with enough detail, what changes are to be made. At this point, the leader needs to have follow-up discussions with individuals failing to fully implement the grading changes such as the one below:

“Mr. Walters, in my last visit to the class, I noticed that four students had failed to complete the assignment for the day, but it was not clear to me how you would handle that situation. Could you please explain how the work will be completed and what support from me, if any, you may need to ensure that these students get back on track?”

While this may sound overly specific to some people, leaders will undermine their change efforts if they are not clear in what they expect. Teachers have many areas over which they have discretion, but grading policies should not be one of them.

3. Aligning Systemic Support

Because schools and school systems are complex, changing one aspect inevitably affects other areas. Making a change in grading policy without considering the impact on other aspects of school life may undermine the implementation. For example, changing grading policies may impact failure rates, discipline, morale, equity, college opportunities, and many other areas. While the consequences may include improved opportunities for more students, at first this may mean that advanced classes are larger or that the composition of those classes may become more diverse. It is important for school leaders to first identify all of the possible implications of the new grading policy and then anticipate (rather than react to) potential complaints, “Thanks for your concern—we have actually anticipated that, and here is why we think it’s a good idea...”

Overall, the test of any new grading system is whether it is working. By “working” this means that students, parents, and teachers can use the feedback from the grading system to improve performance. Even in the early elementary grades, students should be able to respond to the question, “What do you think you need to do to get better?” If teachers have used a standards-based approach and language students can understand, then the response should be something like, “I only got a 2 because I forgot to... but I can get a 3 because I will...” This shows that the student understands his or her performance and knows what to do to improve it.

An effective grading system will only be put into place if the leader does the following:

- Rather than announce the changes, engages the community in extensive dialogue about those changes
- Has a clear idea that the purpose of grading is to improve student learning, not as a tool to punish and reward
- Tolerates dissent but is not intimidated by it
- Actually impacts student performance through the grading changes by improving student success

PROFESSIONAL DEVELOPMENT SUGGESTIONS

NOTE: You can conduct the following discussions with just school leaders or with your entire faculty depending on how much you want to open up the topic of grading for discussion.

I. Lay the Groundwork for Changes in Grading Practices: Discussion of a Better Future

Reeves clearly warns school leaders *not* to introduce new ideas about grading with a discussion of *grading policies* in Chapter 8. This will just lead to a heated discussion and a hardening of views. Instead, lead staff through a discussion of a different and better future using questions like the ones below:

- How would our school be different if we had fewer student failures?
- What would it mean to have fewer students in each class *repeating* that class?
- If we did not have to devote so many resources to students failing classes, how else could we use those resources?
- If we had fewer failures, suspensions, expulsions, and fewer low-level discipline problems, how would our professional and personal lives be better?

II. Discussion of the purposes of Grading

1. Have staff rank what they believe to be the top three purposes of grading from the list below from 1 – 3. Then have them discuss their rankings with a partner.

- A. Giving feedback to improve student performance
- B. Reporting to parents on student progress
- C. Communicating with teachers at the next level of instruction
- D. Giving rewards to students for good behavior and attitudes
- E. Assigning punishments to students for poor behavior and attitudes
- F. Making public distinctions between good and bad students

2. Share with staff that research shows that using grades to reward and punish does *not* improve student achievement. Only the top three purposes (A, B, and C) are valid reasons for giving grades. We should NOT be giving grades as rewards. And yet, a number of our actions and grading practices reflect other reasons for giving grades. Share these quotations with staff and have them discuss whether these are valid reasons for giving grades:

- “I know that she didn’t meet the academic standard, but how else I am supposed to recognize that she tried hard, came to class, and finished her homework? Honor roll grades are the only recognition that poor kid will ever get!”
- “He’s not even sorry for missing his assignments! His attitude is casual and contemptuous. I don’t care if he aced the end-of-course exam – he’s failing my class until he shows some contrition.”

III. Some Problems with Our Current Grading System

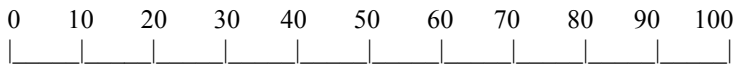
1. *Varying Methods for Calculating Grades:* Have your staff (each person independently) determine the final grade for a student who has received the following grades throughout one marking period: C, C, MA (missing assignment), D, C, B, MA, MA, B, and A. Then have teachers share the final grades they calculated. There will be a wide variety of results. Discuss: How can our grading system be accurate when we have such widely differing final grades for this student?

2. *Including Behavior in Grades:* To share another example of a problem with our current approach to grading, show teachers how including behavior in grades means that the same grade can be given to a wide range of students. In the example below, a grade of B- has been given to four different students for four different reasons:

- Outstanding effort and perseverance, but the student has not yet met grade-level standards
- Outstanding performance well above grade level, but the student’s attitude, work ethic, and class participation are inadequate
- Superior performance, except for one incident of cheating that resulted in a score of zero on a major exam
- Failure to meet all academic standards, but the student earned several extra credit points to merit the final grade

Have teachers discuss this. Should a B- be allowed to represent all of these different situations? If not, what’s the alternative? If this is too abstract, have teachers think of two students who received the same grade last marking period and determine how well the grade represented the differences in the two students. If your teachers include behavior when they calculate grades have them discuss what type of behavior (in-class disruptions, turning in homework late, etc.) they include in their grades and how this affects the legitimacy of those grades.

3. *The Distortion of Zeroes*: If teachers at your school use a 100-point scale, ask them to mark on the number line below where an A, B, C, and D fall. Then ask them how many points they think a student should get for not turning in an assignment and have them mark this on the number line as well.



Discuss that giving a student a zero for not turning in work would mean that the distance between the zero and the D (if a D is 60) is 60, or six times as much as the distance between a C and a D (which is 10). So not turning in the work is six times worse than doing an awful job on the work!

IV. Alternatives: Changes We Can Make in Our Grading System

1. *The Four Criteria of Effective Feedback*: To make changes in the grading system, teachers need to understand the four criteria necessary for effective feedback. Distribute the chart below and have teachers assess themselves. Have teachers think back to the last assessment they gave to students and rate themselves on these four criteria – how accurate, fair, specific, and timely was their feedback to their students? Have them think and then discuss this in pairs.

The Four Criteria of Effective Feedback	
Accurate	Feedback must accurately reflect what a student has learned rather than how well the student behaved. Note that a grade may be mathematically accurate (in the case of a “correct” average) yet not reflect what the student has learned.
Fair	Feedback must not be influenced by the gender, ethnicity, socioeconomic status, or other characteristics of the students.
Specific	Feedback must be specific enough to help students improve their performance. Furthermore, the more specific the feedback guidelines (in a rubric, for example), the more reliable and consistent the feedback will be.
Timely	Feedback is timely when it is delivered to students with sufficient promptness to influence their performance.

2. *Grading Changes We Can Make at Our School*: Have small groups brainstorm together to come up with possible suggestions for ways to revise your current approach to grading in order to make it more accurate, fair, specific, and timely. Before sending off the groups to brainstorm, share some of the suggestions from the book to give them ideas:

- Eliminate the use of zero and instead require students make up missed work during the school day
- Give separate grades for behavior
- Have teachers collaboratively score student work to improve accuracy
- Use a Standards Achievement Report (p. 4 of the summary) to clarify how students are performing on each individual standard and to separate out grades for behavior
- Eliminate the practice of averaging to find final grades and instead weigh end-of-year assignments more heavily
- Eliminate the grade of D (“the coward’s F”) and require that students rework assignments instead
- Use the “Menu System” (p. 7 of the summary) as an alternative to the quick zero

Have the small groups share their brainstormed ideas for reforming the grading system, discuss the merits of each, and then charge a committee to move forward and work out the details of making some school-wide changes to improve grading.

An Idea for Conducting Classroom Observations: Observe for FEEDBACK

Assessing teacher feedback when conducting classroom observations

Use the questions Reeves gave to his students when observing a coach at a basketball game to observe how effective your teachers are at giving feedback. Go into several classrooms and *only* observe for feedback. Look for:

- The percentage of students who received feedback – Did everyone get some type of feedback?
- The frequency of feedback for each student – Did those who need more help get it?
- The nature of the feedback – Was the feedback positive, negative, challenging, or prescriptive? Was the feedback differentiated based on student need?
- The impact of the feedback – Did the feedback lead to improved performance?

Collaborative Scoring of Student Work

Email Jenn if you want three suggestions for ways your teachers can collaboratively score student work.