Grading in the Differentiated Classroom

Announcing a New and FREE Website for Perspective and Practicality on Assessment and Grading Issues!

www.stenhouse.com/fiae

1. Two new, substantial study guides for *Fair Isn’t Always Equal*
2. Q&A’s - abbreviated versions of correspondence with teachers and administrators
3. Video and audio podcasts on assessment and grading issues
4. Testimonials from educators
5. Articles that support the book’s main themes

I could assess and grade students in my differentiated classroom if…

The thing about grading in a differentiated class that I don’t understand yet, is…

How is it fair when….??

What do I say to colleagues, parents, or students who ask about….??

What do I do when….??

Where do you stand on these?

- Differentiated and fair grading
- Rubrics
- Modified or adjusted curriculum
- Student self-assessment
- Extra credit
- What grades mean
- Definitions of individual grades
- Grading scales (100 vs 4.0)
- Formative vs summative assessments
- Averaging grades vs using median/mode
- Grading classwork
- Grading homework
- The purpose of homework
- How much curriculum should be on one test and tiering tests

The role of alternative assessments
- Weighting grades
- The percent influence of varied assessments
- Dealing with late work
- Setting up the gradebook according to categories, assessment formats or standards
- Re-doing work or tests for full credit
- The purpose of grades and grading

What is fair…

…isn’t always equal.
Define Each Grade

A:
B:
C:
D:
E or F:

Until Report Card Formats catch up to pedagogy, we may have to translate into three languages:

<table>
<thead>
<tr>
<th>Rubric Symbol</th>
<th>English</th>
<th>Report Card Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Mastery</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>Just below mastery</td>
<td>90</td>
</tr>
</tbody>
</table>

Three Reasons to Not Refer to Average, Above Average, Below Average

• Society changes its perception of what is average.
• “Criterion-reference” is standards-based and more helpful to everyone involved, not “norm-reference.”
• Averaging was invented in statistics to get rid of sample error, but in order to apply it, the experimental (assessment) design must be constant. Classroom assessments are not constant, and error is inherent.

A Perspective that Changes our Thinking:

“A ‘D’ is a coward’s ‘F.’ The student failed, but you didn’t have enough guts to tell him.”

-- Doug Reeves

• A
• B
• C
• I, IP, NE, or NTY

If we do not allow students to re-do work, we deny the growth mindset so vital to student maturation, and we are declaring to the student:

• This assignment had no legitimate educational value.
• It’s okay if you don’t do this work.
• It’s okay if you don’t learn this content or skill.

None of these is acceptable to the highly accomplished, professional educator.

Once we cross over into D and F(E) zones, does it really matter? We’ll do the same two things: Personally investigate and take corrective action.
Conclusions from Sample DNA Essay Grading

The fact that a range of grades occurs among teachers who grade the same product suggests that:
• Assessment can only be done against commonly accepted and clearly understood criteria.
• Grades are relative.
• Teachers have to be knowledgeable in their subject area in order to assess students properly.
• Grades are subjective and can vary from teacher to teacher.
• Grades are not always accurate indicators of mastery.

What is Mastery?

“Tim was so learned, that he could name a horse in nine languages; so ignorant, that he bought a cow to ride on.”

Ben Franklin, 1750, Poor Richard’s Almanac

Working Definition of Mastery
(Wormell)

Students have mastered content when they demonstrate a thorough understanding as evidenced by doing something substantive with the content beyond merely echoing it. Anyone can repeat information; it’s the masterful student who can break content into its component pieces, explain it and alternative perspectives regarding it cogently to others, and use it purposefully in new situations.

• Determine the surface area of a cube.
• Determine the surface area of a rectangular prism (a rectangular box)
• Determine the amount of wrapping paper needed for another rectangular box, keeping in mind the need to have regular places of overlapping paper so you can tape down the corners neatly.
• Determine the amount of paint needed to paint an entire Chicago skyscraper, if one can of paint covers 46 square feet, and without painting the windows, doorways, or external air vents.

There’s a big difference: What are we really trying to assess?

• “Explain the second law of thermodynamics” vs. “Which of the following situations shows the second law of thermodynamics in action?”
• “What is the function of a kidney?” vs. “Suppose we gave a frog a diet that no impurities—fresh organic flies, no pesticides, nothing impure. Would the frog still need a kidney?”
• “Explain Keynes’s economic theory” vs. “Explain today’s downturn in the stock market in light of Keynes’s economic theory.”

From, Teaching the Large College Class, Frank Heppner, 2007, Wiley and Sons

What will you and your colleagues accept as evidence of full mastery and of almost mastery?

• Spelling test non-example
• No echoing or parroting
• Regular conversations with subject-like colleagues
• Other teachers grading your students’ work
• Pacing Guides
• Common Assessments
One of the Stronger Tips of the Day:

Know the difference between formative and summative assessments. Use them in ethical, instructionally sound ways.

We shouldn’t take the time to assess students unless we’re going to use the information to:

1) guide instructional decisions
2) provide descriptive feedback to students.

Feedback vs Assessment

Feedback: Holding up a mirror to students, showing them what they did and comparing it what they should have done – There’s no evaluative component!

Assessment: Gathering data so we can make a decision

Greatest Impact on Student Success:

Formative feedback

Two Ways to Begin Using Descriptive Feedback:

- “Point and Describe” (from Teaching with Love & Logic, Jim Fay, David Funk)
- “Goal, Status, and Plan for the Goal”

1. Identify the objective/goal/standard/outcome
2. Identify where the student is in relation to the goal (Status)
3. Identify what needs to happen in order to close the gap

What does our understanding of feedback mean for our use of homework?

Is homework more formative or summative in nature? Whichever it is, its role in determining grades will be dramatically different.

“If we don’t count homework heavily, students won’t do it.”

Do you agree with this? Does this sentiment cross a line?
Two Homework Extremes that Focus Our Thinking

- If a student does none of the homework assignments, yet earns an "A" (top grade) on every formal assessment we give, does he earn anything less than an "A" on his report card?

- If a student does all of the homework well yet bombs every formal assessment, isn’t that also a red flag that something is amiss, and we need to take corrective action?

Be clear: We mark and grade against standards/outcomes, not the routes students take or techniques teachers use to achieve those standards/outcomes.

Given this premise, marks/grades for these activities can no longer be used in the academic report of what students know and can do regarding learner standards: maintaining a neat notebook, group discussion, class participation, homework, class work, reading log minutes, band practice minutes, dressing out in p.e., showing up to perform in an evening concert, covering textbooks, service to the school, group projects, signed permission slips, canned foods for canned food drive...

---

Accuracy of the Final Report Card Grade versus the Level of Use of Formative Assessment Scores in the Final Report Grade

**High Final Grade Accuracy**

**Low Final Grade Accuracy**

Why Do We Grade?

- Provide feedback
- Document progress
- Guide instructional decisions

- Motivate
- Punish
- Sort students

What about incorporating attendance, effort, and behavior in the final grade?

Consider...

- Teaching and learning can and do occur without grades.
- We do not give students grades in order to teach them.
- Grades reference summative experiences only - cumulative tests, projects, demonstrations, NOT formative experiences.
- Students can learn without grades, but they must have feedback.
- Grades are inferences based upon a sampling of student’s work in one snapshot moment in time. As such they are highly subjective and relative.

Premise

A grade represents a valid and undiluted indicator of what a student knows and is able to do – mastery.

With grades we document progress in students and our teaching, we provide feedback to students and their parents, and we make instructional decisions.
‘Time to Change the Metaphor:

Grades are NOT compensation.
Grades are communication: They are an accurate report of what happened.

10 Practices to Avoid in a Differentiated Classroom
They Dilute a Grade’s Validity and Effectiveness

- Penalizing students’ multiple attempts at mastery
- Grading practice (daily homework) as students come to know concepts [Feedback, not grading, is needed]
- Withholding assistance (not scaffolding or differentiating) in the learning when it’s needed
- Group grades
- Incorporating non-academic factors (behavior, attendance, and effort)

• Assessing students in ways that do not accurately indicate students’ mastery (student responses are hindered by the assessment format)
• Grading on a curve
• Allowing Extra Credit
• Defining supposedly criterion-based grades in terms of norm-referenced descriptions (“above average,” “average”, etc.)
• Recording zeroes on the 100.0 scale for work not done

0 or 50 (or 60)?

100-pt. Scale:
0, 100, 100, 100, 100, 100 -- 83% (C+)
60, 100, 100, 100, 100, 100 -- 93% (B+)

Imagine the Reverse...

A = 100 – 40
B = 39 – 30
C = 29 – 20
D = 19 – 10
F = 9 – 0

What if we reversed the proportional influences of the grades? That “A” would have a huge, yet undue, inflationary effect on the overall grade. Just as we wouldn’t want an “A” to have an inaccurate effect, we don’t want an “F” grade to have such an undue, deflationary, and inaccurate effect. Keeping zeroes on a 100-pt. scale is just as absurd as the scale seen here.

Be clear: Students are not getting points for having done nothing. The student still gets an F. We’re simply equalizing the influence of the each grade in the overall grade and responding in a way that leads to learning.
A (0) on a 100-pt. scale is a (-6) on a 4-pt. scale. If a student does no work, he should get nothing, not something worse than nothing. How instructive is it to tell a student that he earned six times less than absolute failure? Choose to be instructive, not punitive.

Consider the Correlation

| 100 | 4 |
| 90  | 3 |
| 80  | 2 |
| 70  | 1 |
| 60  | 0 |
| 50  | -1|
| 40  | -2|
| 30  | -3|
| 20  | -4|
| 10  | -5|
| 0   | -6|

Temperature Readings for Norfolk, VA:
85, 87, 88, 84, 0 (Forgot to take the reading)
Average: 68.8 degrees

This is inaccurate for what really happened, and therefore, unusable.

Clarification:

When we're talking about converting zeroes to 50's or higher, we're referring to zeroes earned on major projects and assessments, not homework, as well as anything graded on a 100-point scale. It's okay to give zeroes on homework or on small scales, such as a 4.0 scale. Zeroes recorded for homework assignments do not refer to final, accurate declarations of mastery, and those zeroes don't have the undue influence on small grading scales.

Grading Late Work

• One whole letter grade down for each day late is punitive. It does not teach students, and it removes hope.
• A few points off for each day late is instructive; there’s hope.
• Yes, the world beyond school is like this.

Helpful Consideration for Dealing with Student’s Late Work:

Is it chronic…
...or is it occasional?

We respond differently, depending on which one it is.

This quarter, you’ve taught:
• 4-quadrant graphing
• Slope and Y-intercept
• Multiplying binomials
• Ratios/Proportions
• 3-dimensional solids
• Area and Circumference of a circle.

The student’s grade: B

What does this mark tell us about the student’s proficiency with each of the topics you’ve taught?
Unidimensionality – A single score on a test represents a single dimension or trait that has been assessed

<table>
<thead>
<tr>
<th>Student</th>
<th>Dimension A</th>
<th>Dimension B</th>
<th>Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>6</td>
<td>12</td>
</tr>
</tbody>
</table>

Problem: Most tests use a single score to assess multiple dimensions and traits. The resulting score is often invalid and useless. -- Marzano, CAGTW, page 13

Setting Up Gradebooks in a Differentiated Classroom

- Avoid setting up gradebooks according to formats or media used to demonstrate mastery: tests, quizzes, homework, projects, writings, performances
- Instead, set up gradebooks according to mastery: objectives, benchmarks, standards, learner outcomes

Set up your gradebook into two sections:

**Formative**
Assignments and assessments completed on the way to mastery or proficiency

**Summative**
Final declaration of mastery or proficiency

Summative Assessments

<table>
<thead>
<tr>
<th>Standards/Outcomes</th>
<th>XYZ Test, part 1</th>
<th>PQR Project</th>
<th>EFGB Observ.</th>
<th>XYZ Test, part 2</th>
<th>GHI Perf. Task</th>
<th>Most Consistent Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 [Descriptor]</td>
<td></td>
<td>3.5</td>
<td></td>
<td>3.5</td>
<td></td>
<td>3.5</td>
</tr>
<tr>
<td>1.2 [Descriptor]</td>
<td>2.5</td>
<td>5.0</td>
<td>4.5</td>
<td>4.5</td>
<td></td>
<td>4.5</td>
</tr>
<tr>
<td>1.3 [Descriptor]</td>
<td>4.5</td>
<td>3.5</td>
<td>3.0</td>
<td>3.5</td>
<td></td>
<td>3.5</td>
</tr>
<tr>
<td>1.4 [Descriptor]</td>
<td>3.5</td>
<td>3.5</td>
<td></td>
<td></td>
<td></td>
<td>3.5</td>
</tr>
<tr>
<td>1.5 [Descriptor]</td>
<td></td>
<td>2.0</td>
<td>1.5</td>
<td></td>
<td></td>
<td>1.75</td>
</tr>
</tbody>
</table>

Gradebooks and Report Cards in the Differentiated Classroom: Ten Important Attributes

1. Everything is clearly communicated, easily understood
2. Use an entire page per student
3. Set up according to Standards/Outcomes
4. Disaggregate!
5. No averaging – Determine grades based on central tendency, trend, mode

Gradebooks and Report Cards in the Differentiated Classroom: Ten Important Attributes

6. Behavior/Effort/Attendance separated from Academic Performance
7. Grades/Marks are as accurate as possible
8. Some students may have more marks/grades than others
9. Scales/Rubric Descriptors readily available, even summarized as possible
10. Grades/marks revisable
Responsive Report Formats

Adjusted Curriculum Approach:

Grade the student against his own progression, but indicate that the grade reflects an adjusted curriculum. Place an asterisk next to the grade or check a box on the report card indicating such, and include a narrative comment in the cumulative folder that explains the adjustments.

Responsive Report Formats

Progression and Standards Approach:

Grade the student with two grades, one indicating his performance with the standards and another indicating his own progression. A, B, C, D, or F indicates the student’s progress against state standards, while 3, 2, or 1 indicates his personal progression.

Responsive Report Formats

Multiple Categories Within Subjects Approach:

Divide the grade into its component pieces. For example, a "B" in Science class can be subdivided into specific standards or benchmarks such as, "Demonstrates proper lab procedure," “Successfully employs the scientific method," or "Uses proper nomenclature and/or taxonomic references."

The more we try to aggregate into a single symbol, the less reliable that symbol is as a true expression of what a student knows and is able to do.

Report Cards without Grades

Course: Standard

---

Standard 1 Usage/Punct/Spelling
Standard 2 Analysis of Literature
Standard 3 Six + 1 Traits of Writing
Standard 4 Reading Comprehension
Standard 5 Listening/Speaking
Standard 6 Research Skills

Standards Rating

(1) (2) (3) (4)

Standard 1 2.5
Standard 2 1.75
Standard 3 3.25
Standard 4 3.25
Standard 5
Standard 6 4.0

Additional Comments from Teachers:

Health and Maturity Records for the Grading Period:

Consider:

• A 4.0 scale has a high inter-rater reliability. Students’ work is connected to a detailed descriptor and growth and achievement rally around listed benchmarks.

• In 100-point or larger scales, the grades are more subjective. In classes in which teachers use percentages or points, students, teachers, and parents more often rally around grade point averages, not learning.

100 point scale or 4.0 Scale?

• Pure mathematical averages of grades for a grading period are inaccurate indicators of students’ true mastery.

• A teacher’s professional judgment via clear descriptors on a rubric actually increases the accuracy of a student’s final grade as an indicator of what he learned.

• A teacher’s judgment via rubrics has a stronger correlation with outside standardized tests than point or average calculations do.
Accurate grades are based on the most consistent evidence. We look at the pattern of achievement, including trends, not the average of the data. This means we focus on the median and mode, not mean, and the most recent scores are weighed heavier than earlier scores.

**Median:** The middle test score of a distribution, above and below which lie an equal number of test scores

**Mode:** The score occurring most frequently in a series of observations or test data

**Suggested Language to Use in Parents' Handbook:**

Parents, as we are basing students' grades on standards for each discipline, final grades are first and foremost determined by our teachers' professional opinion of your child's work against those standards, not by mathematical calculations. Teachers have been trained in analyzing student products against standards and in finding evidence of that learning using a variety of methods. Please don't hesitate to inquire how grades for your child were determined if you are unsure.

**Allowing Students to Re-do Assignments and Tests for Full Credit:**

- Always, "...at teacher discretion."
- It must be within reason.
- Students must have been giving a sincere effort.
- Require parents to sign the original assignment or test, requesting the re-do.
- Require students to submit a plan of study that will enable them to improve their performance the second time around.

**Grading Inclusion Students**

**Question #1:**

"Are the standards set for the whole class also developmentally appropriate for this student?"

- If they *are* appropriate, proceed to Question #2.
- If they *are not* appropriate, identify which standards are appropriate, making sure they are as close as possible to the original standards. Then go to question #2.

**Question #2:**

"Will these learning experiences (processes) we’re using with the general class work with the inclusion student as well?"

- If they *will* work, then proceed to Question #3.
- If they *will not* work, identify alternative pathways to learning that will work. Then go to Question #3.

**Allow Students to Re-do Assignments and Tests for Full Credit:**

- Identify a day by which time this will be accomplished or the grade is permanent.
- With the student, create a calendar of completion that will help them achieve it.
- Require students to submit original with the re-done version so you can keep track of their development.
- Reserve the right to give alternative versions.
- No-re-do’s the last week of the grading period.
- Sometimes the greater gift is to deny the option.
Grading Inclusion Students

Question #3: “Will this assessment instrument we’re using to get an accurate rendering of what general education students know and are able to do regarding the standard also provide an accurate rendering of what this inclusion student knows and is able to do regarding the same standard?”

- If the instrument will provide an accurate rendering of the inclusion student’s mastery, then use it just as you do with the rest of the class.
- If it will not provide an accurate rendering of the inclusion student’s mastery, then identify a product that will provide that accuracy, and make sure it holds the student accountable for the same universal factors as your are asking of the other students.

Grading Gifted Students

- Insure grade-level material is learned.
- If it’s enrichment material only, the grade still represents mastery of on-grade-level material. An addendum report card or the comment section provides feedback on advanced material.
- If the course name indicates advanced material (Algebra I Honors, Biology II), then we grade against those advanced standards.
- If the student has accelerated a grade level or more, he is graded against the same standards as his older classmates.

Great New Books on Feedback, Assessment, and Grading:

- Elements of Grading, Doug Reeves, Solution Tree Press, 2010
- How to Give Feedback to Your Students, Susan M. Brookhart, ASCD, 2008
- Developing Performance-Based Assessments, Grades 6-12, Nancy P. Gallavan, Corwin Press, 2009
- Balanced Assessment, From Formative to Summative, Kay Burke, Solution Tree, 2010

Recommended Reading on Assessment and Grading

- Borich, Gary D.; Tombari, Martin L. Educational Assessment for the Elementary and Middle School Classroom (2nd Edition), Prentice Hall, 2003

Recommended Reading on Assessment and Grading

- www.exemplars.com
- Lewis, Larry; Shoemaker, Betty Jean. Great Performances: Creating Classroom-Based Assessment Tasks, John Wiley & Sons, 1998
- Marzano, Robert. Transforming Classroom Grading, ASCD 2001
- Marzano, Robert. Classroom Assessment and Grading that Work, ASCD 2006
- Marzano, Robert; McTighe, Jay; and Pickering, Debra. Assessing Student Outcomes: Performance Assessment Using the Dimensions of Learning Model, Association for Supervision and Curriculum Development, 1993

Recommended Reading

- Stiggins, Richard J. Student-Involved Classroom Assessment (3rd Edition), Prentice Hall, 2000

Grant Wiggins Web site and organization:
Center on Learning, Assessment, and School Structure (CLASS)
info@classnj.org   www.classnj.org
gpw@classnj.org