GRADING for LEARNING MANUAL
Wauwatosa Public Schools

Contents

Why We Use Grading for Learning (G4L)

The main purpose of grading is to communicate clearly and consistently to a student (and their parents) how well they're mastering benchmarks. The model we use to grade and to communicate with our students is called ***Grading for Learning (G4L).*** We use this model because it's aligned with best practices for 21st Century teaching and learning. Some hallmarks of 21st Century teaching and learning are:

- A focus on critical thinking skills and big ideas
- Mastery learning that is project-based, performance based, or problem-based
- Authentic student engagement--meaning students:
    - are genuinely motivated to succeed
    - take responsibility for their learning
    - monitor/track their learning over time
- Collaborative teaching and cooperative learning
- Authentic audiences through community relationships and the digital world
- Technology and 21st Century Skills

When we communicate with students, we give them information to help them master benchmarks, which are thinking skills and big ideas. A student's benchmarks scores should give that student an idea of how they're doing in your class. Additionally, parents should be able to look at their child's benchmark scores and their assessment grades, and have a sense of how their child is doing in a course.

It's important to note that the term ***Grading for Learning*** comes from the work of Ken O'Connor, who is one of education's foremost authorities on grading and assessment. An example of his work can be found here: GRADING PPT

A Word about Equity

The Wauwatosa School District is deeply committed to ensuring equitable learning for all students. The Grading for Learning model supports equity in our schools and classrooms. In shifting the primary focus from having students memorize and reproduce recalled information to having them think critically and independently, we create more entry points to the learning process. By offering practices like multiple opportunities and recency, we honor more diverse learning styles. By separating out academic performance and behavior, we give more students access to success. Our district has relied on the work of Zaretta Hammond and Sharroky Hollie, among other educators, in examining the relationship between assessment and equity.

G4L Tenets

In the summer of 2016, a committee of teachers worked on Grading for Learning, and they identified these tenets as a foundation for grading to which we aspire.

1. We grade so we can ***communicate*** students' mastery of benchmarks to the students and their parents. In our grading, we give timely and effective feedback that promotes further learning.

2. We use ***benchmarks*** to direct our teaching and identify what is important for students to learn. We use benchmarks to organize how record and store students' grades.

3. We focus on teaching ***thinking skills*** and ***big ideas*** over the teaching of recalled information. When we do focus on recalled information, we are selective, and we teach it in a meaningful context for students.

4. We use grading practices that promote ***equitable access*** for all students to be successful in mastering the benchmarks.

5. We ***collaborate*** closely with our colleagues in creating and using ***common summative assessments.*** We also create ***rubrics*** both to help students understand what is expected of them and to assess their work. As teachers, we work on ***calibrating*** our grading practices so that they're valid and reliable from one teacher to another in common courses.

6. We base students' academic grades on ***academic performance***, not on behaviors like timeliness, neatness, respect, etc.

7. We acknowledge that learning is ***complex*** and that students learn in different ways and at different rates. Because student learning is complex:
   - We offer ample, ungraded ***formative assessment*** for students as they practice learning;
   - We offer ***multiple pathways*** to student learning;
   - We focus on ***performance-based***, ***project-based*** and ***problem-based*** learning to simulate real-world learning.
   - We give more weight to the ***most recent*** assessments;
   - We use our ***professional discretion*** in determining student grades.

Paradigm Shifts

Certain paradigm shifts--changing the way we think--support moving from traditional grading to benchmark-based [standards-based] grading.

TRADITIONAL PRACTICE	STANDARDS-BASED GRADING
Tracking grades by assessments [tests, quizzes, etc.]	Tracking grades by benchmarks
Grading academic behavior *and* academic performance	Grading only academic performance
Including all benchmark scores earned over a grading period (like a semester) in determining a grade	Using the most recent benchmark scores when determining grades in a grading period (like a semester)
Relying solely on math to determine a term grade	Using standardized grading protocols to determine a term grade, including rubrics, spiraling, multiple opportunities and **a teacher's professional discretion***
Basing grades on percent of material mastered	Basing grades on quality of student performance relative to benchmarks
Limiting grades to what happens in one specific unit	Giving students multiple opportunities to demonstrate mastery of benchmarks, including spiraled grading practices over a number of units
Focusing primarily on the recall of information--memorization of factual information	Focusing primarily on the mastery of thinking skills and big ideas. [When we do teach factual information that we want students to recall, we're selective, and we teach the information in the context of critical thinking skills and big idea s.]
Using paper-and-pen written tests or quizzes as the primary way of assessing students	Using assessments that are performance-based where students show their mastery of learning in real-life situations or simulations

**A Word about Professional Discretion*

Grant Wiggins said, "All grading by human judges is subjective. The question is not whether grading is subjective, but whether it is defensible and credible. I wish we could stop using 'subjective' as a pejorative word! So-called objective scoring [multiple choice] is still subjective test writing." January 19, 2000

We acknowledge that as teachers we have to employ our professional discretion at times--sometimes even when it conflicts with the mathematical outcome of a student's grade. We rely on our expertise, our experience, and our G4L protocols, such as recency and retakes to assist us in determining students' grades.

Summative Assessments and Formative Assessments

Summative assessments are end-of-unit assessments that capture a student's mastery at the completion of a learning period. Formative assessments are usually more frequent, "smaller" assessments that can take place every day and include classroom checks for understanding, quizzes, and homework assignments. Teachers collect and record both formative and summative assessments in the Infinite Campus gradebook for students and parents to see. These are some typical differences between formative and summative assessments.

Formative Assessments	Summative Assessments
are considered practice	are final products or performances and are thus eligible for retakes
can occur frequently	occur less frequently than formative assessments
are "smaller" in scope, covering fewer benchmarks	are "bigger" in scope, covering a number of benchmarks
can be very informal, on-the-spot, individualized for one or several students	are typically for formal, for all students in a course
can count for up to 10% of a student's grade	can determine 90% of a student's grade

The term "common summative assessment" (CSA) refers to crosstown, end-of-unit assessments written by teachers representing both sides of town (East/West, Longfellow/Whitman). While students often work collaboratively, their grades should reflect only their individual achievement; summative "group grades" are not an acceptable grading practice.

Typically, by the end of the semester, a course should have a minimum of three scores per benchmark. However, content areas can have anywhere from 3 to 7 benchmarks, so individual courses may vary in the number of benchmarks they have, and when a course has more benchmarks, it's appropriate that there may be only two summative scores for some of its benchmarks. Teachers should communicate to students and parents the number of benchmarks and summative scores in their courses.

NOTE about GRADING FORMATIVE ASSESSMENTS:

When G4L was initially implemented, formative assessments could not count in a student's grade. In the summer of 2017 a representative committee of teachers and administrators made

a number of changes to G4L, one of which relates to formative assessment. The change was that formative assessment could count for up to 10% of a grade. The decision of how much formative assessment would count in a course grade was to be made by all teachers of a given course.

Recording Summative Assessment Scores

Here's an example for one student's benchmark scores in a course. Please note this is not a literal "shot" of an Infinite Campus page. It's simply a depiction of a student's grading history in a course. In this course, the teachers have decided that the most recent three scores for each benchmark will count in the students' grades.

Assessments	**Benchmark 1 Read**	**Benchmark 2 Write**	**Benchmark 3 Research**	**Benchmark 4 Oral**
Short Story Project [9/12]	2	2	--	3
Interview Essay [10/4]	**3**	--	**3**	**2**
Poetry Presentation [10/17]	**3**	**2**	**2**	**3**
Poetry Essay [11/2]	--	**4**	--	--
Short Story Project [11/20]	**3**	**3**	**3**	**3**
Benchmark Average Scores	**3**	**3**	**2.67**	**2.67**

Aligning Summative Assessments and Benchmarks
COURSE PLANNING TEMPLATE

Here is a blank temple that can be used to align your assessments with your benchmarks. When you do this alignment, you are, in effect, planning your course. You can list your summative assessments in the column on the left and then indicate which benchmarks align with each assessment. Remember that over the course of a semester, to be valid, each benchmark in a course should have a minimum of three scores.

Also note, that sometimes teachers will not use all of their content area benchmarks for a given course. In the case below there are 5 benchmarks. It's possible that the teachers of this course might use only 4 of the benchmarks. A decision like this should be made by all the teachers of the course, on both sides of town.

Summative Assessments	**Benchmark 1**	**Benchmark 2**	**Benchmark 3**	**Benchmark 4**	**Benchmark 5**

Teaching with the End in Mind

A phrase that's heard a lot in teaching these days is "teaching with the end in mind." This means that a teacher's starting point in planning a course or a unit should be: "What thinking skills and big ideas do I want my students to master?" Answering this question is "teaching with the end in mind." After you identify what you want your students to learn, then you identify ways they can show you what they've mastered. These are your assessments. After creating your assessments, you should then think about the instruction necessary for the students to master the goals. In other words, you start with curriculum [what students should learn], you move to assessment [how they demonstrate their learning], and then end by creating engaging, equitable instruction to pull students in [how they will learn].

Grant Wiggins and Jay McTighe coined the term "teaching with the end in mind." They have created a model called UNDERSTANDING by DESIGN, which many teachers have found helpful in planning curriculum and assessments.

Benchmarks

Benchmarks (often called standards in other districts) represent broad learning goals, like problem solving, reading, applying scientific models, evaluating sources, etc. Each of our content areas (academic departments) has three to six benchmarks. Often, a teacher uses all the benchmarks in a content area to track students' progress throughout a semester. However, sometimes the team of teachers teaching a course decide to use only some of their content area benchmarks. A course should use a minimum of three benchmarks. A student's grade reflects their mastery of the benchmarks in a given course.

Here is an example. In Science, there are three benchmarks.

- Interpret a variety of ***data and evidence***.
- Analyze, evaluate, and construct, ***models, inferences and experimental results***.
- Develop ***deep understandings of the core concepts*** of science—and communicate and apply these understandings.

In Biology, a teacher tracks students' learning using these three benchmarks. Over the course of a unit, students are developing their mastery related to these three benchmarks. A unit summative assessment in Biology would typically allow students to demonstrate their learning related to the three benchmarks. Sometimes teachers "break up" a unit summative assessment into parts, and they administer these parts at different times throughout a unit. This could be the case in virtually any course in any content area. By the end of a grading period (semester), each benchmark in a course should have a *minimum* of three summative scores, because three points of data establish a valid trend or process. Benchmark scores should indicate how well a student is doing in mastering the given benchmarks. Here is a link to a list all SECONDARY BENCHMARKS.

Usually benchmarks represent important thinking skills or abilities. Here are two benchmarks from English.

- Read: Comprehend, analyze and interpret a variety of texts
- Research: Gather, evaluate and synthesize information

Here are two examples from Computer Science:

- Design and implement creative solutions and artifacts
- Analyze problems and artifacts using strategies

Sometimes benchmarks reflect big ideas or concepts. In Social Studies, one benchmark states, "Develop an understanding of core social studies concepts." Other content areas embed big ideas or concepts into the thinking skills. For instance, in English, when students work on research skills, they're also working with big ideas and concepts.

Grades

Teachers use numeric scores when assessing student work: 4, 3, 2, and 1. Additionally, when necessary teachers can give half-garde, for example 3½. When a teacher considers student work over time, over the course of several summative assessments, they can use this rubric, which is based loosely on Norman Webb's DEPTH of KNOWLEDGE.

4 Students consistently demonstrate ***thorough mastery*** of benchmarks in **new or extended contexts.**

3 Students consistently demonstrate ***adequate mastery*** of benchmarks in **familiar contexts.**

2 Students demonstrate ***partial mastery*** of benchmarks in **familiar contexts.**

1 Students inconsistently demonstrate a ***partial mastery*** of benchmarks **in familiar contexts.**

INC [Incomplete] There is insufficient evidence of student work at this time to provide a grade. An unresolved incomplete eventually defaults to azero.

Please note that the rubric above is a GLOBAL rubric: It is broad, and it is not based on a specific set of skills from a content area. If you were to use the rubric above it would be important for your colleagues and your students to have a shared understanding of critical terms like, "thorough," "adequate," or "familiar contexts." It should be clear to everyone using the rubric what these terms mean.

Here is a second GLOBAL rubric you could use that would be helpful when focusing on only one assessment at one point in time.

4 Students demonstrate a ***thorough*** understanding of ideas and knowledge and a ***thorough*** level of benchmark mastery**.**

3 Students demonstrate a ***satisfactory*** understanding of ideas and knowledge and a ***satisfactory*** level of benchmark mastery.

2 Students demonstrate a ***partial*** of understanding of ideas and benchmarks and a ***partial*** level of benchmark mastery.

1 Students demonstrate an ***emerging*** understanding of ideas and benchmarks and an ***emerging*** level of benchmark mastery.

It's important to note that the two rubrics above are GLOBAL rubrics. LOCAL rubrics, one developed for a specific content area or course, can sometimes be more effective to use with your students on tasks and assessments. Here are two examples of LOCAL rubrics, very different from each other but both very effective.

World Languages: Interpersonal Communication

2	4
Uses single words, memorized phrases and short sentences to express information. Pauses frequently in the conversation.	Combines phrases into sentences to express both information and insights. Speech flows naturally.
Responds to questions.	Responds to questions and follows up by asking logical questions.
Uses English and target words. Uses gestures/facial expressions to show confusion.	Remains in target language. Asks for rewording or slowing of speech when needed.
Converses by completing basic goals of the conversation. Speech sounds like a scripted interview or like classroom practice. Conveys information.	Completes the extended goals of the conversation. Speech sounds like an authentic conversation. Each speaker sounds unique. Conveys humor and emotion.

Science Cross-Cutting Standards: Patterns

4	3	2	1
Analyze and connect multiple patterns using mathematical methods and/or other methods.	Explain patterns in terms of scientific concepts.	Organize and interpret patterns to predict outcomes.	Identify patterns observed in physical phenomena.

Context: *Students have been assigned to do a series of physical and chemical changes in the lab. Students arrange their data and observations so as to accomplish the above tasks/skills. i*

Level One: *The student correctly identifies physical and chemical changes.*

Level Two: *The student organizes the data in the patterns that have been identified. The student interprets the results to predict what would happen if similar tests were done.*

Level Three: *The student explains the patterns in terms of the chemical and physical changes and the criteria used to identify these changes.*

Level Four: *The student analyzes other changes to determine if the changes are physical or chemical using the patterns that were identified in the lab.*

Teachers develop rubrics to assess levels of benchmark mastery on both formative and summative assessments. Rubrics are used for assessing student performance and for giving students meaningful, timely feedback to assist them in their mastery of the benchmarks. This feedback should help students move forward in their learning process.

Another very valuable tool to use when assessing students in the G4L model is exemplars, which are student-created examples of work. You can give students exemplars of masterful work, as well as exemplars of developing work. The exemplars can give students an idea of what types of performance you expect of them.

Teachers use whole points (1 through 4) when assessing students. They may also use half-point scores (½, 1½, 2½, 3½) when they feel a student's performance falls squarely between two whole-point levels. A zero can be given as a benchmark score once a grading period is completed and a teacher has no evidence whatsoever of a student's level of mastery related to a benchmark. The assumption is that the teacher has given the student multiple, accessible opportunities to demonstrate learning, but the student has done nothing to allow the teacher to evaluate their learning.

A Word about Percentages

In the Grading for Learning model, teachers use the scores of 1 through 4 on assessments. Over time, a student gets a number of scores for each benchmark, and Infinite Campus averages those scores. In order to average benchmark scores, Infinite Campus must convert scores of 1 through 4 to percentages (numbers between 1 and 100). It is simply a technical, mechanical necessity. So, Infinite Campus assigns 100% to a 4, 75% to a 3, 50% to a 2, and 25% to a 1. It then averages scores in a percent-based system and ultimately reports the scores back in the 1 through 4 scale.

It's important not to project the traditional scoring system (93 to 100 is an A, 85 to (2 is a B, etc.) onto Grading for Learning. If a student is earning a 2 in a class, that means they are doing roughly "C work." The 50% Infinite Campus assigns to the 2, in order to perform calculations, should not be construed to mean they're mastering 50% of the material, which would have been an "F" in the traditional system. Students and parents do not see percentages in the gradebooks, and teachers would do well to ignore them.

Calculating a Grade

All content areas have agreed upon benchmarks that represent their important thinking skills and big ideas. A content area or course may have from three to seven benchmarks depending on how they have organized their thinking skills and big ideas. Throughout the semester teachers assess and record a student's developing mastery of these benchmarks in Infinite Campus.

Step One

The teacher scores an assessment often using a rubric that describes the four levels of mastery related to benchmark mastery, 4.0, 3.0, 2.0, 1.0. Additionally, a teacher may assign a half number for added precision, 3.5, 2.5, 1.5, 0.5. All scores are recorded in Infinite Campus according to the benchmark assessed.

Step Two

An individual benchmark grade is calculated in Infinite Campus by averaging the summative scores for any grading period.

Step Three

The course grade is calculated by averaging the individual benchmark scores. A running snapshot of this grade is available in Infinite Campus throughout the semester.

Grading Scale

Grade	A	A-	B+	B	B-	C+	C	C-	D+	D	D-	F
Upper	4.0	3.46	3.19	2.92	2.65	2.39	2.12	1.85	1.59	1.32	1.05	0.79
Lower	3.47	3.2	2.93	2.66	2.4	2.13	1.86	1.6	1.33	1.06	0.8	0
Span*	0.53	0.26	0.26	0.26	0.25	0.26	0.26	0.25	0.26	0.26	0.25	0.79
Span**	0.8		0.79			0.79			0.79			0.79

*Span from lower point to upper point **Span for the letter grade range

GPA

A student's grade point average is calculated at the end of each semester using a 4-point scale. In high school a student's cumulative GPA is reported on the student's transcript.

Grade	A	A-	B+	B	B-	C+	C	C-	D+	D	D-	F

Points	4.0	3.67	3.33	3.0	2.67	2.33	2.0	1.67	1.33	1.0	0.67	0

Multiple Opportunities

In the Grading for Learning model, we acknowledge that students learn in different ways and at different rates. Not every student masters a given skill by the day of the assessment—sometimes it takes them more time, or they need another attempt at demonstrating their mastery. That's why we offer multiple opportunities to demonstrate benchmark mastery. Multiple opportunities can come about in a variety of ways. The most obvious way would be to let a student retake an assessment, or a variation of it. Another way would be to "spiral" in the learning outcome(s) to the following unit, giving the student another opportunity to master the outcome(s) and show their mastery. An example of this would be assessing an ongoing skill like making a claim. In some classes, a student will have many opportunities to show their ability to make a claim. So if they weren't satisfied with their work in one unit, they could work on perfecting that skill and giving evidence of it in the following unit, because that skill is being spiraled through. When a teacher offers multiple opportunities, including spiraling, they need to discount previous, lower scores if a student demonstrates mastery of a benchmark after a number of unsuccessful attempts.

In August 2017, the Grading for Learning Retake Policy was revised. Here is the revised version of the policy.

Grading for Learning Retake Policy

Teachers have the discretion to decide whether to grant a student an isolated one-event retake of a summative assessment. Isolated one-event retakes are not a routine expectation for summative assessments. Retakes are offered in the spirit of promoting a culture of teaching and learning, not as a replacement for a lack of preparation for summative assessments. Also, teachers have the discretion to use various types of multiple opportunities rather than the isolated one-event type. Some examples of these other types of retakes are spiraling, formative assessments, grading for recency, portfolios, etc. Retakes at teacher discretion require an active partnership between teacher and student.

In this partnership, the student will:

- Complete homework in a timely manner
- Participate in class
- Communicate with the teacher regarding their progress
- Reflect thoughtfully on their learning
- Fulfill teacher requirements for a retake

In this partnership, the teacher will:

- Ensure there are multiple opportunities in a semester for students to show their learning
- Use grades solely to communicate how students are mastering benchmarks
- Provide timely feedback on student work
- Provide re-teaching opportunities to support student learning
- Provide clear expectations about the learning process and how students are being assessed
- Provide clear communication about retakes by conferencing with the student

Updates to the Original Model Regarding: Formative Assessment, Responsibility for Learning, and Recency

The original Grading for Learning Model was fully implemented in the 2014-15 school year. In summer 2017, the following changes were made.

Formative Assessment

Past Practice: Formative assessments, including homework, aren't calculated into a student's course grade.

Change: Formative assessment, including both homework and classwork, can count up to 10%. The decision on how much to count formative assessment must be made by all the teachers of a course, including those on both sides of town. A percentage between 0 and 10 has the weight to affect a student's grade somewhat, but not so much weight so as to severely impact a student's grade if they don't complete assignments. Including formative assessment in the calculation of a grade underlines the importance of engaging in classwork, quizzes and homework, but still allows for true assessment of a student's mastery of benchmarks.

Conditions: If formative work is included in a grade, there must be teacher feedback on the work that informs student learning.

Responsibility for Learning

Past Practice: Teachers use the R4L score as a communication tool regarding student behavior. It is used in some buildings to determine student access to co-curricular activities and student events.

Change: Teachers can continue to use the R4L score as a communication tool, but the scores will not be used to determine student access to co-curricular activities and student events.

Conditions: Even though the new PowerSchool [now Infinite Campus] gradebook would allow us to change the way we communicate the R4L grade, we should maintain the use of "substituting" scores rather than averaging them over time, and maintain one score rather than delineating 2 or 3 scores.

Recency

Past Practice: Only the 3 or 4 most recent summative benchmark scores are counted directly into a student's course grade. Counting only the 3 or 4 most recent benchmark scores was meant to give students an opportunity to practice a skill before being graded on it. We found teachers have come up with many ways to give students practice without being graded,

and we found that, in some cases, work from the entire semester can be used to give an accurate grade.

Change: Teachers will no longer interpret *recency* to mean the institutionalized use of counting only the 3 or 4 most recent benchmark scores. Teachers will continue to honor the complexity and developmental nature of the learning process through the use of teacher discretion and the use of such practices such as spiraling, portfolios, multiple opportunities, retakes, etc. Teachers may continue to count the most recent three summative scores, or they may count more scores in determining a grade. The decision of how many summative scores to count in determining a grade must be made by all the teachers of the course. including both sides of town.

Conditions: In any course, all decisions regarding which summative assessments will be calculated into students' grades will be made in crosstown PLC teams.

RESPONSIBILITY FOR LEARNING (R4L)

When we give students a course grade, we take into account how well they are mastering course benchmarks. Course benchmarks are really important skills. In an English class, reading and writing are examples of benchmarks. In Science, interpreting data and communicating scientific information are examples of benchmarks. A course grade reflects how well a student is mastering course benchmarks. A course grade does *not* include student behaviors.

Here are some examples of student behavior:

- Shows respect for one's classmates, one's environment and one's self
- Engages in class by participating in activities and cooperating with other students
- Demonstrates responsibility by preparing for class, doing homework, and managing time

While these behaviors are important in a child's overall education, they are not directly related to a student's academic grade, because that grade is reserved for their mastery of the benchmarks. We keep academic behaviors separate from academic grades. However, because these behaviors are so important, we do want to give students and their parents information on the quality of these behaviors. That's where the Responsibility for Learning Score comes in.

The Responsibility for Learning Score reflects the quality of a student's behavior in a class. We use the following rubric to provide feedback on student behaviors such as those listed above:

Consistently (4)
The student consistently demonstrates the above behaviors.

Usually (3)
The student usually demonstrates the above behaviors.

Occasionally (2)
The student occasionally demonstrates the above behaviors.

Rarely (1)
The student rarely demonstrates the above behaviors.

It should be noted that attendance is very important and can obviously have an effect on grades. However, attendance should not directly be factored into a student's academic grade.

QUOTATIONS on GRADING

GRADING

How confident are you that the grades students get in your school are:

- accurate
- consistent
- meaningful, and
- meaningful, and
- supportive of learning?

If grades do not meet these four conditions of quality they are "broken," i.e., ineffective.
Ken O'Connor, 2012

GRADING AND TRADITION

"Grading practices are not the result of careful thought or sound evidence, rather, they are used because teachers experienced these practices as students and, having little training or experience with other options, continue their use."
Guskey, Thomas R. (Editor), Communicating Student Learning: The 1996 ASCD Yearbook, ASCD, Alexandria, VA, 1996

GRADES AS PUNISHMENT

"Don't use grades punitively… Without exception, experts in the area of student grading recommend that grades not be used in a punitive sense. When a teacher uses grades as punishment for student behaviors, the teacher establishes an adversarial relationship in which grades are no longer meaningful to students as indicators of their accomplishments. The punitive use of grades only increases the likelihood that students will lose respect for the evaluation system; consequently the appeal to students of subverting such a system will be heightened." Cizek, G. J. 2003, Detecting and Preventing Cheating; Promoting Integrity in Assessment, Corwin, Thousand Oaks, CA, 2003, 100 in O'Connor, K., A Repair Kit for Grading, Pearson, Boston, MA

GRADING IS SUBJECTIVE

"All grading by human judges is subjective. The question is not whether grading is subjective, but whether it is defensible and credible. The AP and IB programs are credible and defensible, and yet they are subjective. I wish we could stop using "subjective" as a pejorative word! So-called objective scoring [multiple choice] is still subjective test writing."
Grant Wiggins, January 19, 2000

PRIMARY PURPOSE OF GRADING

"The primary purpose for grading is to communicate with students [and parents] about their achievement of learning goals. . . .
Brookhart, S., Grading, Pearson Merrill Prentice Hall, Columbus, OH, 2004

PRIMARY PURPOSE OF GRADING

"The primary purpose of grades is to communicate student achievement to students and parents."
Bailey, J. and McTighe, J., "Reporting Achievement at the Secondary School Level: What and How?" 1996

WE DON'T NEED GRADES TO LEARN

"Teachers don't need grades or reporting forms to teach well. Further, students don't need them to learn." Thomas R. Guskey, (Ed.) Communicating Student Learning, 1996

THE SYMBOL IN GRADING

"What critics of grading must understand is that the symbol is not the problem; the lack of stable and clear points of reference in using symbols is the problem." Wiggins, G., "Honesty and Fairness: Toward Better Grading and Reporting" 1996

THE SYMBOL IN GRADING

"Grades or numbers, like all symbols, offer efficient ways of summarizing." Wiggins, G., "Honesty and Fairness: Toward Better Grading and Reporting" 1996

THE SYMBOL IN GRADING

"Trying to get rid of familiar letter grades gets the matter backwards while leading to needless political battles. Parents have reason to be suspicious of educators who want to tinker with a 120 year old system that they think they understand - even if we know that traditional grades are often of questionable worth." Wiggins, G., "Honesty and Fairness: Toward Better Grading and Reporting" 1996

WARM DEMANDERS

"Warm demanders first establish a caring relationship that convinces students that the teacher believes in them and has their best interests at heart. . . . On the basis of this relationship, warm demanders relentlessly insist that all students perform required academic work and treat the teacher and their peers with respect." Abstract of Bondy, E, and D. D. Ross. "The Teacher as Warm Demander," Educational Leadership, September 2008.

EFFECT OF LOW GRADES

"Instead of prompting greater effort, low grades more often cause students to withdraw from learning." Guskey and Bailey, Developing Grading and Reporting Systems for Student Learning, Corwin Press, 2001

SHIFTING TO GRADE BY BENCHMARKS

"The use of columns in a grade book to represent standards, instead of assignments, tests, and activities, is a major shift in thinking. Under this system, when an assessment is designed, the teacher must think in terms of the standards it is intended to address. If a (test) is given that covers three standards, then the teacher makes three entries in the grade book for each student - one entry for each standard - as opposed to one overall entry for the entire (test)." Marzano, R., and J. Kendall, A Comprehensive Guide to Developing Standards-Based Districts, Schools, and Classrooms, McREL, Aurora, CO, 1996

ALIGNMENT HELPS

"Systems that are aligned – curriculum, teaching, and assessment - have a greater chance of success for students." Glenda Lappan, NCTM News Bulletin, October, 1998

THE AVERAGE

"Educators must abandon the average, or arithmetic mean, as the predominant measurement of student achievement." Reeves, D., "Standards are Not Enough: Essential Transformations for School Success," NASSP Bulletin, Dec. 2000

ZEROES

"A zero has an undeserved and devastating influence, so much so that no matter what the student does, the grade distorts the final grade as a true indicator of mastery. Mathematically and ethically this is unacceptable. " Rick Wormeli, 2007

ASSESSMENT

"The ongoing interplay between assessment and instruction, so common in the arts and athletics, is also evident in classrooms using practices such as nongraded quizzes and practice tests, the writing process, formative performance tasks, review of drafts and peer response groups. The teachers in such classrooms recognize that ongoing assessments provide feedback that enhances instruction and guides student revision." McTighe, J., "What Happens Between Assessments," Educational Leadership, 1997

FORMATIVE ASSESSMENT

"The thrust of formative assessment is toward improving learning and instruction. Therefore, the information should not be used for assigning "marks" as the assessment often occurs before students have had full opportunities to learn content or develop skills." Manitoba Education and Training, Reporting on Student Progress and Achievement: A Policy Handbook for Teachers, Administrators and Parents. Winnipeg, 1997

FEEDBACK AND LEARNING

Feedback that Supports Learning

- Focuses on attributes of the work rather than on attributes of the student
- Is descriptive of the work; how to do better
- Clearly understood by the user
- Clearly understood by the user
- Is sufficiently detailed to be helpful, but does not overwhelm
- Arrives in time to inform the learning Chappuis, 2009

TIME, PRACTICE, FORMATIVE ASSESSMENT

"We know that students will rarely perform at high levels on challenging learning tasks at their first attempt. Deep understanding or high levels of proficiency are achieved a result of trial, practice, adjustments based on feedback and more practice." McTighe, J., "What Happens Between Assessments", Educational Leadership, 1997

G4L RESOURCES: LINKS

These are the resources/links presented in this manual and some additional ones. None of these links is necessarily 100% aligned with our Wauwatosa G4L model (except the SECONDARY BENCHMARKS), but they all present and support new ways of thinking about grading and assessment.

Ken O'Connor GRADING PPT

Zaretta Hammond

Sharroky Hollie

UNDERSTANDING by DESIGN: Grant Wiggins,

Wauwatosa's SECONDARY BENCHMARKS.

Norman Webb's DEPTH of KNOWLEDGE.

Rick WORMELLI

THOMAS GUSKEY WEBSITE

ARTICLES

- Grades that Show What Students Know
- Grading What Matters
- Making the Grades
- Seven Reasons for Standards-Based Grading

Standards-based grading article Ed Leadership

John Hattie: Assessment Article