

Grading and Reporting Student Learning

Thomas R. Guskey

**Distinguished Service Professor
Georgetown College
400 East College Street
Georgetown, KY 40324**

Phone: 502-863-7010

E-mail: Guskey@georgetowncollege.edu

Thomas R. Guskey

Distinguished Service Professor – Educational Measurement and Evaluation
Georgetown College, 400 East College Street, Georgetown, KY 40324
(Phone: 502-863-7010 -- E-mail: Guskey@georgetowncollege.edu)

Dr. Guskey is the Distinguished Service Professor and Co-Director of the Center for the Advanced Study of Assessment at Georgetown College in Georgetown, Kentucky, and widely known for his research in education reform, assessment, grading practices, and mastery learning. A graduate of the University of Chicago, he has taught at all levels, served as an administrator in Chicago Public Schools, and was the first Director of the *Center for the Improvement of Teaching and Learning*, a national educational research center. His books have won numerous awards and his articles have appeared in prominent research journals as well as *Educational Leadership*, *Kappan*, and *School Administrator*. He served on the Policy Research Team of the *National Commission on Teaching & America's Future*, on the Task Force to develop the *National Standards for Staff Development*, and recently was honored by the American Educational Research Association for his work relating research to practice. He co-edits *the Experts in Assessment Series* for Corwin Press and was featured on the National Public Radio program, "Talk of the Nation." As a consultant to schools throughout the United States as well as Europe and Asia, he helps bring clarity and insight to some of education's most complex problems.

Publications on Grading and Reporting

Developing Standards-Based Report Cards

(with J. Bailey). Thousand Oaks, CA: Corwin, 2008.

Practical Solutions for Serious Problems in Standards-Based Grading. Thousand Oaks, CA: Corwin, 2008.

How's My Kid Doing? A Parents' Guide to Grades, Marks, and Report Cards.

San Francisco: Jossey Bass, 2002.

Developing Grading and Reporting Systems for Student Learning (with J. Bailey). Thousand Oaks, CA: Corwin, 2001.

Implementing Student-Led Conferences.

(with J. Bailey). Thousand Oaks, CA: Corwin, 2001.

Communicating Student Learning.

Alexandria, VA: Association for Supervision and Curricular Development, 1996.

Presentation Digest XI: Developing Grading and Reporting Systems for Student Learning.

(with Bailey). Alexandria, VA: Association for Supervision and Curriculum Development, 2003.

Grading and Reporting Student Learning – Professional Development Inquiry Kit. (with Marzano). Alexandria, VA: Association for Supervision and Curriculum Development, 2002.

"Grading Policies and Standards-Based Reforms: Conflicts and Solutions." In S. Mathison & E. W. Ross (Eds.), *Battleground Schools: An Encyclopedia of Conflict and Controversy*. Westport, CT: Greenwood, 2007.

"Standards-Based Grading and Reporting: A Model for Special Education (with L. Jung). *Teaching Exceptional Children*, 2007, 40(2), 48-53.

"Efficacy of Online Reporting". *Alberta Assessment Consortium Communiqué*, p. 2. Also available at: <http://www.aac.ab.ca/freenews.html>

"'It Wasn't Fair!' Educators' Recollections of Their Experiences as Students with Grading. *Journal of Educational Research and Policy Studies*, 2006, 6(2), 111-124.

"Making High School Grades Meaningful."

Phi Delta Kappan, 2006, 87(9), 670-675.

"The Problem of Grade Inflation."

Principal Matters, 2006, 66, 38-40.

"The Challenges of Standards-Based Grading."

(with Jung). *Leadership Compass*, 2006, 4(2), 1-5.

"Mapping the Road to Proficiency."

Educational Leadership, 2005, 63(3), 32-38.

"A Historical Perspective on Closing Achievement Gaps." *NASSP Bulletin*, 2005, 89(64), 76-89.

"The Communication Challenge of Standards-Based Reporting."

Phi Delta Kappan, 2004, 86(4), 326-329.

"Zero Alternatives." *Principal Leadership*, 2004, 5(2) 49-53.

"How Classroom Assessments Improve Learning."

Educational Leadership, 2003, 60(5) 6-11

"Computerized Gradebooks and the Myth of Objectivity."

Phi Delta Kappan, 2002, 83(10), 775-780.

"Grading Practices in Schools." In J. W. Guthrie (Ed.),

Encyclopedia of Education (2nd ed.). New York: Macmillan, 2002.

"Helping Standards Make the Grade."

Educational Leadership, 2001, 59(1), 20-27.

"High Percentages are Not the Same as High Standards."

Phi Delta Kappan, 2001, 82(7), 534-538.

"Grading Policies that Work Against Standards ... and How to Fix Them." *NASSP Bulletin*, 2000, 84(620), 20-29.

"Making Standards Work." *The School Administrator*, 1999, 56(9), 44.

"Making the Grade: What Benefits Students?"

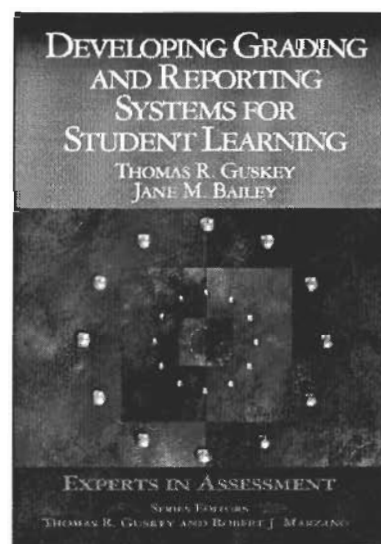
Educational Leadership, 1994, 52(2), 14-20.

Developing Grading and Reporting Systems for Student Learning

Thomas R. Guskey, Jane M. Bailey

EXPERTS IN ASSESSMENT™ SERIES

Teachers, parents, students, administrators, and community members all agree that we need better grading and reporting systems. Often, these systems are inadequate because they are part of a tradition that can go unexamined and unquestioned for years. Here is the first serious look at the issue, written to provide all those involved—especially teachers—with a coherent and thoughtful framework.



Guskey and Bailey offer four pillars of successful grading and reporting systems:

- Communication is the primary goal of grading and reporting
- Grading and reporting are integral parts of the instructional process
- Good reporting is based on good evidence
- Creating change in grading and reporting requires creating a multi-faceted reporting system

Written to help readers develop a deeper and more reflective understanding of the various aspects of the subject, Thomas Guskey and Jane Bailey's work brings organization and clarity to a murky and disagreement-filled topic.

Here is a practical and essential guide for teachers, administrators or anyone concerned with understanding and implementing best practices in grading and reporting systems.

2001, 232 pages, 7" x 10"

Paperback: **\$32.95**, D6624-0-8039-6854-X

Hardcover: **\$69.95**, D6624-0-8039-6853-1

Helping Educators Do Their Work Better

Table of Contents

Introduction / Defining the Impetus for Change / Exploring the History of Grading and Reporting / Laying a Foundation for Change / Building a Grading and Reporting System / Grading and Reporting Methods I / Grading and Reporting Methods II / Grading and Reporting for Students with Special Needs / Special Problems in Grading and Reporting / Model Reporting Forms / Guidelines for Developing Effective Reporting Systems

ORDER FORM

Priority Code: D6624

BILL TO (if different) Please attach original purchase order.

☐ Purchase Order # _____
Name: _____
Title: _____
Organization: _____
Address: _____
City: _____ State: _____
Zip Code: _____
Telephone: _____
Required: ☐ ☐ ☐ - ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Four EASY WAYS to order!

SHIP TO

Name: _____
Title: _____
Organization: _____
Address: _____
City: _____ State: _____
Zip Code: _____

**CALL
Toll Free**

(800) 818-7243

Monday-Friday: 6 am-5 pm PT

**Complete
and send
whole page.**

MAIL

2455 Teller Road
Thousand Oaks, CA 91320-2218

**FAX
Toll Free**
(800) 417-2466

ONLINE
CorwinPress.com



A SAGE Publications Company

Qty.	Book #	Title	Unit Price	Total Price
	D6624-0-8039-6854-X	Paperback	\$32.95	
	D6624-0-8039-6853-1	Hardcover	\$69.95	

☐ Please send your latest catalog FREE FREE

DISCOUNTS ARE AVAILABLE
for large quantity orders —
CALL (800) 818-7243
and ask for a sales manager.

Prices subject to change without notice.
Professional books may be tax-deductible.
Federal ID Number 77-0260369

Total Book Order

Sales Tax

Add appropriate sales tax in CA, MA, MD, NY
Add appropriate GST & HST in Canada

Shipping and Handling

\$3.50 for first book, \$1.00 for each additional book
Canada: \$10.00 for first book, \$2.00 each additional book

Total Amount Due \$

Remit in US dollars

All orders are shipped Ground Parcel.

For other shipping methods and cost, call (800) 818-7243

Payment Method

☐ Check # _____ Payable to Corwin Press

CREDIT CARD

☐ ☐ ☐ ☐

Credit Card #: _____

Expiration Date:

____/____
month/year

Signature: _____

In case we have questions...

Telephone: _____

Fax: _____

E-mail: _____

☐ Yes, you may e-mail other Corwin Press offers to me.
Your e-mail address will NOT be released to any third party.

HOW'S MY KID DOING?

A PARENT'S GUIDE TO GRADES, MARKS, AND REPORT CARDS



THOMAS R. GUSKEY

How's My Kid Doing?

A Parent's Guide to Grades, Marks, & Report Cards

By

Thomas R. Guskey

Over the years, Tom Guskey is the teacher from whom I have learned the most about the principles of effective communication. He has consistently analyzed and articulated our communication options with immense clarity. He's done it again, this time for parents.

**-Rick Stiggins, President
Assessment Training Institute, Oregon**

ABOUT THE BOOK:

Most parents want schools to provide honest, clear, and explicit information on how their child is doing – with specific suggestions for improvement. Unfortunately, most schools are providing “progress reports” that parents find vague, confusing, inconsistent, and delivered in unfamiliar formats. *How's My Kid Doing* helps parents make sense of their child's grades, test scores, and report cards by explaining the advantages and shortcoming of different reporting methods. It answers parents' most frequently asked questions about plus and minus grades, grading on the curve, standards, and narrative evaluations. And, it offers strategies for working with teachers and with children to improve the system. Most important, it illustrates how educators and parents can become true partners in a child's learning.

ABOUT THE AUTHOR:

THOMAS R. GUSKEY is professor of education at the University of Kentucky, Lexington. He is a frequent speaker at national education conferences, and a leading expert on the topics of grading, assessment, and professional development in education.

Jossey-Bass, A Wiley Co. • 10475 Crosspoint Drive, Indianapolis, IN 46256 • PH: 800.956.7739 • FX: 800.605.2665 • www.jpb.com

Please send me the following: *How's My Kid Doing? A Parent's Guide to Grades, Marks & Report Cards*
By Thomas R. Guskey
ISBN: 0-7879-6073-X \$24.95 Cloth pp. 176
Address _____

City _____ State _____ Zip _____

Payment Enclosed. All payments must be in U.S. dollars.
(Payable to Jossey-Bass)

☐ Check/Money order ☐ Visa ☐ Discover
☐ American Express ☐ MasterCard

Card no. _____

Exp. Date _____ Day Phone _____

Signature _____

Please add shipping charges:

Merchandise	Surface	2-Day	1-Day
First Book	\$5.00	10.50	17.50
Each additional book	\$3.00	3.00	4.00

For orders of 15 books or more call 800.956.7739

Prices listed are for domestic United States customers and are subject to change without notice. Availability subject to stock on hand.

Book Price \$ _____ Shipping \$ _____ Sales Tax \$ _____ Total \$ _____

Grading & Reporting Student Progress to Enhance Learning

With Tom Guskey

New assessments and standards demand better reporting systems that are meaningful to both parents and students and have the capability to impact learning. This program will feature a variety of ways to report student progress and alternative forms of parent conferences. Policies and practices that negatively impact students will be examined.

- ♦ Consider the importance of changing traditional assessment and grading practices.
- ♦ Examine the role of daily and culminating assessments as tools for learning.
- ♦ Identify the purposes for grading and the need to clearly state it.
- ♦ Explore various ways to report student learning, including report cards and student-led conferences.
- ♦ Design reporting systems to better communicate and involve parents in student learning.
- ♦ Consider the impact of negative policies and practices to be avoided.

Think about the purpose of grading. Don't use grades as weapons. They do not serve that purpose well and never will. We know that grading and reporting are not essential to the instructional process. Teachers teach and students learn in the absence of grades. You need to decide the purpose.

– Tom Guskey

Stock 903 - \$355, Includes two videotape programs, a guidebook for discussion, and an audiotape soundtrack.

Stock No.	Description	Qty	Unit Price	Total
903	Grading & Reporting Student Progress to Enhance Learning		\$355	
Shipping & handling 5% in US				
Grand Total				

Call Now!!

1 800 572-1153



The Video Journal of Education
8686 South 1300 East
Sandy, Utah 84094
FAX 888 566-6888 or 801 566-6885
<http://www.videojournal.com>

Name _____ Title _____
 School _____
 District _____
 Address _____
 City, State, Zip _____
 Phone () _____ FAX () _____
 PO# _____ or ☐ Visa ☐ MC ☐ AMEX
 Card# _____
 Exp. _____ Signature _____

***Questionnaires
and
Activities***

GRADING AND REPORTING QUESTIONNAIRE

© Thomas R. Guskey

Name (Optional) _____ Grade Level _____

Years of Teaching Experience _____ Subject(s) _____

Directions: *Please read each question carefully, think about your response, and answer each as honestly as you can.*

1. What do you believe are the major reasons we use report cards and assign grades to students' work?

- a. _____
- b. _____

2. Ideally, what purposes do you believe report cards or grades should serve?

- a. _____
- b. _____

3. Although classes certainly differ, on average, what percent of the students in your classes receive the following grades:

A _____ B _____ C _____ D _____ E or F _____

4. What would you consider an ideal distribution of grades (in percent) in your classes?

A _____ B _____ C _____ D _____ E or F _____

5. The current grading system in many schools uses the following combination of letter grades, percentages, and/or categories:

A	100% - 90%	Excellent	Exceptional
B	89% - 80%	Good	Proficient
C	79% - 70%	Average	Basic
D	69% - 60%	Poor	Below Basic
E or F	59% -	Failing	

If you could make any changes in this system, what would they be?

- a. _____
- b. _____

6. Is there an established, uniform grading policy in your school or district?

Yes _____ No _____ I don't know _____

How well would you say you understand those policies?

Not at all

Somewhat

Very well

1 ----- 2 ----- 3 ----- 4 ----- 5

7. Grades and other reporting systems serve a variety of purposes. Based on your beliefs, rank order the following purposes from **1** (Most important) to **6** (Least important).

- ___ Communicate information to parents about students' achievement and performance in school
- ___ Provide information to students for self-evaluation
- ___ Select, identify, or group students for certain educational programs (Honor classes, etc.)
- ___ Provide incentives for students to learn
- ___ Document students' performance to evaluate the effectiveness of school programs
- ___ Provide evidence of students' lack of effort or inappropriate responsibility

8. Teachers use a variety of elements in determining students' grades. Among those listed below, please indicate those that you use and about what percent (%) each contributes to students' grades.

- | | |
|-------------------------------------|--------------------------------|
| ___ Major examinations | ___ Oral presentations |
| ___ Major compositions | ___ Homework completion |
| ___ Unit tests | ___ Homework quality |
| ___ Class quizzes | ___ Class participation |
| ___ Reports or projects | ___ Work habits and neatness |
| ___ Student portfolios | ___ Effort put forth |
| ___ Exhibits of students' work | ___ Class attendance |
| ___ Laboratory projects | ___ Punctuality of assignments |
| ___ Students' notebooks or journals | ___ Class behavior or attitude |
| ___ Classroom observations | ___ Progress made |
| ___ Other (Describe) _____ | |
| ___ Other (Describe) _____ | |

9. What are the most positive aspects of report cards and the process of assigning grades?

10. What do you like least about report cards and the process of assigning grades?

Grading Formulae: What Grade Do Students Deserve?

© Thomas R. Guskey

The table below shows the performance of seven students over five instructional units. Also shown are the summary scores and grades for these students calculated by three different methods: (1) the simple arithmetic average of unit scores, (2) the median or middle score from the five units, and (3) the arithmetic average, deleting the lowest unit score in the group.

Consider, too, the following explanations for these score patterns:

Student 1 struggled in the early part of the marking period but continued to work hard, improved in each unit, and did excellently in unit 5.

Student 2 began with excellent performance in unit 1 but then lost motivation, declined steadily during the marking period, and received a failing mark for unit 5.

Student 3 performed steadily throughout the marking period, receiving three *B*'s and two *C*'s, all near the *B* – *C* cut-score.

Student 4 began the marking period poorly, failing the first two units, but with newfound interest performed excellently in units 3, 4, and 5.

Student 5 began the marking period excellently, but then lost interest and failed the last two units.

Student 6 skipped school (unexcused absence) during the first unit, but performed excellently in every other unit.

Student 7 performed excellently in the first four units, but was caught cheating on the assessment for unit 5, resulting in a score of zero for that unit.

Summary Grades Tallied by Three Different Methods

Student	Unit 1	Unit 2	Unit 3	Unit 4	Unit 5	Average Score	Grade	Median Score	Grade	Deleting Lowest	Grade
1	59	69	79	89	99	79.0	C	79.0	C	84.0	B
2	99	89	79	69	59	79.0	C	79.0	C	84.0	B
3	77	80	80	78	80	79.0	C	80.0	B	79.5	C
4	49	49	98	99	100	79.0	C	98.0	A	86.5	B
5	100	99	98	49	49	79.0	C	98.0	A	86.5	B
6	0	98	98	99	100	79.0	C	98.0	A	98.8	A
7	100	99	98	98	0	79.0	C	98.0	A	98.8	A

Grading standards: 90% – 100% = A
 80% – 89% = B
 70% – 79% = C
 60% – 69% = D
 – 59% = F

Questions: Which grading method is best? Which is fairest?

What grade does each student deserve?

GRADING AND REPORTING STUDENT LEARNING

From: **Guskey, T. R., & Bailey, J. M. (2001).** *Developing Grading and Reporting Systems for Student Learning*. Thousand Oaks, CA: Corwin Press.

Guskey, T. R. (2002). *How's My Kid Doing: A Parent's Guide to Grades, Marks, & Report Cards*. San Francisco, CA: Jossey-Bass.

General Conclusions From the Research:

1. Grading and reporting are *not* essential to the instructional process.
2. Grading and reporting serve a variety of purposes, but no one method serves all purposes well.
3. Grading and reporting will always involve some degree of subjectivity.
4. Mathematic precision does *not* yield fairer or more objective grading.
5. Grades have some value as a reward, but no value as a punishment.
6. Grading and reporting should always be done in reference to learning criteria, never "on the curve."
7. Three general types of learning criteria are used in grading and reporting:
 - a. *Product* criteria
 - b. *Process* criteria
 - c. *Progress* criteria
8. Report cards are but *one way* to communicate with parents.

Guidelines for Better Practice:

1. Begin with a clear statement of purpose and specific learning goals.
 - a. Why are grading and reporting done?
 - b. For whom is the information intended?
 - c. What are the desired results?
2. Ensure that grading and reporting methods provide accurate and understandable descriptions students learning.
 - a. *More* a challenge in clear thinking and effective communication
 - b. *Less* an exercise in quantifying achievement
3. Use grading and reporting methods to enhance teaching and learning.
 - a. Facilitate communication between teachers, students, parents, and others.
 - b. Ensure that efforts to help students are consistent and harmonious
4. Alleviate questionable practices:
 - a. Example 1: Averaging to obtain a student's grade or mark.
 - b. Example 2: Assigning a 'zero' to work that is late, missed, or neglected.
 - c. Example 3: Taking credit away from students for behavioral infractions.

For additional information contact:

Thomas R. Guskey, Georgetown College, 400 East College Street, Georgetown, KY 40324
(Phone: 502-863-7010 E-mail: Guskey @ georgetowncollege.edu)

Copies of Slides

Grading and Reporting Student Learning

Thomas R. Guskey

Systemic Change

- ✓ Change is a Highly Complex Process
- ✓ Professional Development is Essential

Change is a
Prerequisite for
Improvement!

Grading and Reporting Questionnaire

Guiding Questions

1. What are the major reasons we use report cards and assign grades to students' work?
2. Ideally, what purposes should report cards or grades serve?
3. What elements should teachers use in determining students' grades?
(For example, major assessments, compositions, homework, attendance, class participation, etc.)

Purposes of Grading

1. Communicate the Achievement Status of Students to Their Parents and Others
2. Provide Information for Student Self-Evaluation
3. Select, Identify, or Group Students for Certain Educational Programs
4. Provide Incentives for Students to Learn
5. Document Students' Performance to Evaluate the Effectiveness of Instructional Programs
6. Provide Evidence of Students' Lack of Effort or Inappropriate Responsibility

Grading Elements

- ✓ Major Exams or Compositions
- ✓ Class Quizzes
- ✓ Reports or Projects
- ✓ Student Portfolios
- ✓ Exhibits of Students' Work
- ✓ Laboratory Projects
- ✓ Students' Notebooks or Journals
- ✓ Classroom Observations
- ✓ Oral Presentations
- ✓ Homework Completion
- ✓ Homework Quality
- ✓ Class Participation
- ✓ Work Habits and Neatness
- ✓ Effort Put Forth
- ✓ Class Attendance
- ✓ Punctuality of Assignments
- ✓ Class Behavior or Attitude
- ✓ Progress Made

Guiding Questions

1. Do you want a more general report that would look similar for several grade levels, or a more specific report that would look different at each grade level?
2. What product, process, and progress criteria should be reported at each level?
3. Within each subject area, how many standards will be reported? What are they?
4. How many levels of performance will be reported for each standard? How will those be labeled?
5. Will teachers' comments be encouraged and included? What form will these take? How will they be recorded?
6. How should things be arranged on the report? What format will be used? What information should be included?
7. What will parents be expected to do with this information?
8. What policies need to accompany these new reporting procedures (i.e. the use of zeros; absences; punctuality of assignments; make-up work; behavioral infractions; homework; final examinations; etc.)
9. When should the perspectives and input of parents and/or students be sought?

General Conclusions from the Research on Grading

#1 Grading and Reporting are *NOT* Essential to the Instructional Process

- ✓ Teachers can teach without grades.
- ✓ Students can and do learn without grades.

Checking *is* Essential !

- ➔ Checking is Diagnostic
 - Teacher is an *Advocate*
- ➔ Grading is Evaluative
 - Teacher is a *Judge*

#2 No One Method of Grading and Reporting Serves *All* Purposes Well !

Purposes of Grading

1. Communicate the Achievement Status of Students to Their Parents and Others
2. Provide Information for Student Self-Evaluation
3. Select, Identify, or Group Students for Certain Educational Programs
4. Provide Incentives for Students to Learn
5. Document Students' Performance to Evaluate the Effectiveness of Instructional Programs
6. Provide Evidence of Students' Lack of Effort or Inappropriate Responsibility

Architecture:

Form *Follows* Function.

Education:

Method *Follows* Purpose!

Solution:

Multiple Purposes Require a

***Multi-Faceted,
Comprehensive
Reporting System!***

Critical Factors in Determining Purposes

1. *What* information do we want to communicate?
2. *Who* is the primary audience?
3. *How* would we like that information to be used?

Letter Grades

➔ **Advantages:**

1. Brief Description of Adequacy
2. Generally Understood

➔ **Disadvantages:**

1. Require the Abstraction of Lots of Information
2. Cut-offs are Arbitrary
3. Easily Misinterpreted

Percentage Grades

➔ **Advantages:**

1. Provide Finer Discriminations
2. Increase Variation in Grades

➔ **Disadvantages:**

1. Require the Abstraction of Lots of Information
2. Increased Number of Arbitrary Cut-offs
3. Greater Influence of Subjectivity

Standards-Based (Checklist of Skills)

→ Advantages:

1. Clear Description of Achievement
2. Useful for Diagnosis and Prescription

→ Disadvantages:

1. Often Too Complicated for Parents to Understand
2. Seldom Communicate the Appropriateness of Progress

Steps in Developing Standards-Based Grading

1. **Identify the major learning goals or standards** that students will be expected to achieve at each grade level or in each course of study.
2. **Establish performance indicators** for the learning goals or standards.
3. **Determine graduated levels of performance** (benchmarks) for assessing each goal or standard.
4. **Develop reporting forms** that communicate teachers' judgments of students' learning progress and culminating achievement in relation to the learning goals or standards.

Challenges in Determining Graduated Levels of Student Performance

1. Levels of Understanding / Quality

Modest	Beginning	Novice	Unsatisfactory
Intermediate	Progressing	Apprentice	Needs Improvement
Proficient	Adequate	Proficient	Satisfactory
Superior	Exemplary	Distinguished	Outstanding

2. Level of Mastery / Proficiency

Below Basic	Below Standard	Pre-Emergent	Incomplete
Basic	Approaching Standard	Emerging	Limited
Proficient	Meets Standard	Acquiring	Partial
Advanced	Exceeds Standard	Extending	Thorough

3. Frequency of Display

Rarely	Never
Occasionally	Seldom
Frequently	Usually
Consistently	Always

4. Degree of Effectiveness

Ineffective	Poor
Moderately Effective	Acceptable
Highly Effective	Excellent

5. Evidence of Accomplishment

Little or No Evidence
Partial Evidence
Sufficient Evidence
Extensive Evidence

Guidelines for Reporting on Standards

1. **Avoid Comparative Language.**
"Below Average," "Average," and "Superior" communicates standing among classmates, not progress on standards.
2. **Provide Examples Based on Student Work.**
Show precisely what each level of performance means, based on models of excellence.
3. **Distinguish "Levels of Understanding" and "Frequency of Display."**
"Quality" is not the same as "Rate of Occurrence."
4. **Be Consistent.**
Use similar terms across school levels, assessments, instructional materials, and reporting forms.

Steps in Developing Standards-Based Grading

1. *The process is more a challenge in effective communication than simply documenting achievement.*
2. *Interpretation is the key element in effective communication.*
3. *Consistency is essential to accurate interpretation.*
4. *Development involves a series of compromises.*

Narratives

→ Advantages:

1. Clear Description of Progress and Achievement
2. Useful for Diagnosis and Prescription

→ Disadvantages:

1. Extremely Time-Consuming for Teachers to Develop
2. May Not Communicate Appropriateness of Progress
3. Comments Often Become Standardized

**Methods can be
Combined to
Enhance their
Communicative
Value !**

***Grades with Comments are
Better than Grades Alone!***

Grade	Standard Comment
A	Excellent ! Keep it up.
B	Good work. Keep at it.
C	Perhaps try to do still better?
D	Let's bring this up.
F	Let's raise this grade !

From: Page, E. B. (1958). Teacher comments and student performance:
A seventy-four classroom experiment in school motivation.
Journal of Educational Psychology, 49, 173-181.

Solution:

- 1. *Determine* the Primary Purpose of each Grading and Reporting Tool.**
- 2. *Select or Develop* the Most Appropriate Method for Each Tool.**
- 3. Develop a *Multi-Faceted, Comprehensive Reporting System!***

**#3 Grading and
Reporting Will
Always Involve
Some Degree of
*Subjectivity !***

**In General,
Reporting is *More*
Subjective:**

- ✓ The More *Detailed* the Reporting Method.
- ✓ The More *Analytic* the Reporting Process.
- ✓ The More 'Effort' is Considered.
- ✓ The More 'Behavior' Influences Judgments.

**However, More
Detailed and *Analytic*
Reports are Better
*Learning Tools !***

Challenge:

To Balance Reporting Needs with Instructional Purposes

#4 Mathematic Precision Does *NOT* Yield *Fairer* or *More Objective* Grading!

Student Achievement Profiles:

Student 1 struggled in the early part of the marking period but continued to work hard, improved in each unit, and did excellently in unit 5.

Student 2 began with excellent performance in unit 1 but then lost motivation, declined steadily during the marking period, and received a failing mark for unit 5.

Student 3 performed steadily throughout the marking period, receiving three B's and two C's, all near the B – C cut-score.

Student 4 began the marking period poorly, failing the first two units, but with newfound interest performed excellently in units 3, 4, and 5.

Student 5 began the marking period excellently, but then lost interest and failed the last two units.

Student 6 skipped school (unexcused absence) during the first unit, but performed excellently in every other unit.

Student 7 performed excellently in the first four units, but was caught cheating on the assessment for unit 5, resulting in a score of zero for that unit.

Grading Formulae

Student	Unit 1	Unit 2	Unit 3	Unit 4	Unit 5	Average Score	Grade	Median Score	Grade	Deleting Lowest	Grade
1	59	69	79	89	99	79.0	C	79.0	C	84.0	B
2	99	89	79	69	59	79.0	C	79.0	C	84.0	B
3	77	80	80	78	80	79.0	C	80.0	B	79.5	C
4	49	49	98	99	100	79.0	C	98.0	A	86.5	B
5	100	99	98	49	49	79.0	C	98.0	A	86.5	B
6	0	98	98	99	100	79.0	C	98.0	A	98.8	A
7	100	99	98	98	0	79.0	C	98.0	A	98.8	A

Questionable Practices:

- ✓ Averaging to Obtain a Course Grade
- ✓ Giving Zeros for Work Missed or Work Turned in Late
- ✓ Taking Credit Away from Students For Infractions

Alternatives to *Averaging* Inconsistent Evidence on Student Learning:

- ✓ Give priority to the most recent evidence.
- ✓ Give priority to the most comprehensive evidence.
- ✓ Give priority to evidence related to the most important learning goals or standards.

Alternatives to Giving *Zeros* :

- ✓ **Assign “I” or “Incomplete” Grades.**
Include specific and immediate consequences.
- ✓ **Report Behavioral Aspects Separately.**
Separate “Product” (Achievement) from “Process” and “Progress.”
- ✓ **Change Grading Scales.**
Use Integers (A=4, B=3, C=2, ...) instead of Percentages.

**Grading requires
Thoughtful and
Informed
Professional Judgment!**

**#5 Grades have Some
Value as *Rewards*,
but *NO* Value as
Punishments !**

***Message:*
Do Not Use Grades
as *Weapons* !**

**#6 Grading and Reporting
should *Always* be done
in reference to
Learning Criteria,
Never “On The Curve”**

Grading Criteria

1. *Product* Criteria
2. *Process* Criteria
3. *Progress* Criteria

#7 Grade Distributions Reflect *Both*:

- ✓ 1. Students' Level of Performance
- ✓ 2. The Quality of the Teaching

#8 Report Cards are but *One Way* of Communicating with Parents !

Forms of Reporting to Parents Include:

- | | |
|-------------------------------------|-------------------------------------|
| ✓ Report Cards | ✓ Personal Letters |
| ✓ Notes with Report Cards | ✓ Homework |
| ✓ Standardized Assessment Reports | ✓ Evaluated Assignments or Projects |
| ✓ Weekly / Monthly Progress Reports | ✓ Portfolios or Exhibits |
| ✓ Phone Calls | ✓ School Web Pages |
| ✓ School Open Houses | ✓ Homework Hotlines |
| ✓ Newsletters | ✓ Parent-Teacher Conferences |
| | ✓ Student-Led Conferences |

*For More Information on **Student-Led Conferences**, contact:*

Jane M. Bailey
Director of Teaching & Learning
Petoskey Public Schools
1130 Howard Street
Petoskey, MI 49770

Phone: 231/348-2352
E-mail: bailey.jm.m@petoskeyschools.org

In Reporting to Parents:

1. Include *Positive Comments*.
2. Describe *Specific Learning Goals or Expectations* (Include Samples of the Student's Work).
3. Provide *Specific Suggestions* on What Parents Can Do To Help.
4. Stress *Parents' Role as Partners* in the Learning Process.

#9 High Percentages are *NOT* the same as *High Standards!*

Think about your days as a student. Then, *describe* your most memorable...

- ✓ Negative grading experience.
- ✓ Positive grading experience.

Guidelines for *Better* Practice

#1 Begin with a *Clear Statement of Purpose*

- ✓ Why Grading and Reporting Are Done?
- ✓ For Whom the Information is Intended?
- ✓ What are the Desired Results?

#2 Provide *Accurate and Understandable* Descriptions of Student Learning

- ✓ *More* a Challenge in Effective Communication
- ✓ *Less* an Exercise in Quantifying Achievement

#3 Use Grading and Reporting to *Enhance* Teaching and Learning

- ✓ Facilitate Communication Between Teachers, Parents, and Students
- ✓ Ensure Efforts to Help Students are Harmonious

For Help or Additional Information:

Thomas R. Guskey
Georgetown College
400 East College Street
Georgetown, KY 40324

Phone: 502-863-7010

E-mail: Guskey @ georgetowncollege.edu

Readings

Making the Grade: What Benefits Students?

Thomas R. Guskey

Although the debate over grading and reporting practices continues, today we know which practices benefit students and encourage learning.

Charged with leading a committee that would revise his school's grading and reporting system, Warren Middleton described his work this way:

The Committee On Grading was called upon to study grading procedures. At first, the task of investigating the literature seemed to be a rather hopeless one. What a mass and a mess it all was! Could order be brought out of such chaos? Could points of agreement among American educators concerning the perplexing grading problem actually be discovered? It was with considerable misgiving and trepidation that the work was finally begun.



© Jeffrey High/Image Productions

Few educators today would consider the difficulties encountered by Middleton and his colleagues to be particularly surprising. In fact, most probably would sympathize with his lament. What they might find surprising, however, is that this report from the Committee on Grading was published in 1933!

The issues of grading and reporting on student learning have perplexed educators for the better part of this century. Yet despite all the debate and the multitude of studies, coming up with prescriptions for best practice seems as challenging today as it was for Middleton and his colleagues more than 60 years ago.

Points of Agreement

Although the debate over grading and reporting continues, today we know better which practices benefit students and encourage learning. Given the multitude of studies—and their often incongruous results—researchers do appear to agree on the following points:

1. Grading and reporting aren't essential to instruction. Teachers don't need grades or reporting forms to teach well. Further, students don't need them to learn (Frisbie and Waltman 1992).

Teachers do need to check regularly on how students are doing, what they've learned, and what problems or difficulties they've experienced. But grading and reporting are different from checking; they involve judging the adequacy of students' performance at a specific time. Typically, teachers use checking to diagnose and prescribe and use grading to evaluate and describe (Bloom et al. 1981).

When teachers do both checking and grading, they become advocates as well as judges—roles that aren't necessarily compatible (Bishop 1992). Finding a meaningful compromise between these dual roles makes many teachers uncomfortable, especially those with a child-centered orientation (Barnes 1985).



© Jeffrey High/Image Productions

2. *No one method of grading and reporting serves all purposes well.* Grading enables teachers to communicate the achievements of students to parents and others, provide incentives to learn, and provide information that students can use for self-evaluation. In addition, schools use grades to identify or group students for particular educational paths or programs and to evaluate a program's effectiveness (Feldmesser 1971, Frisbie and Waltman 1992). Unfortunately, many schools attempt to address all of these purposes with a single method and end up achieving none very well (Austin and McCann 1992).

Letter grades, for example, briefly describe learning progress and give some idea of its adequacy (Payne 1974). Their use, however, requires abstracting a great deal of information into a single symbol (Stiggins 1994). In addition, the cut-off between grade categories is always arbitrary and

Teachers don't need grades or reporting forms to teach well. Further, students don't need them to learn.

difficult to justify. If scores for a grade of *B* range from 80 to 89, students at both ends of that range receive the same grade, even though their scores differ by nine points. But the student with a score of 79—a one-point difference—receives a grade of *C*.

The more detailed methods also have their drawbacks. Narratives and checklists of learning outcomes offer specific information for documenting progress, but good narratives take time to prepare, and—not surprisingly—as teachers complete more narratives, their comments become increasingly standardized. From the parents' standpoint, checklists of learning outcomes often appear too complicated to understand. In addition, checklists seldom

communicate the appropriateness of students' progress in relation to expectations for their level (Afflerbach and Sammons 1991).

Because one method won't adequately serve all purposes, schools must identify their primary purpose

for grading and select or develop the most appropriate approach (Cangelosi 1990). This process often involves the difficult task of seeking consensus among several constituencies.

3. *Regardless of the method used, grading and reporting remain inherently subjective.* In fact, the more detailed the reporting method and the more analytic the process, the more likely subjectivity will influence results (Ornstein 1994). That's why, for example, holistic scoring procedures tend to have greater reliability than analytic procedures.

Subjectivity in this process, however, isn't always bad. Because teachers know their students, understand various dimensions of students'

work, and have clear notions of the progress made, their subjective perceptions may yield very accurate descriptions of what students have learned (Brookhart 1993, O'Donnell and Woolfolk 1991).

When subjectivity translates into bias, however, negative consequences can result. Teachers' perceptions of students' behavior can significantly influence their judgments of scholastic performance (Hills 1991). Students with behavior problems often have no chance to receive a high grade because their infractions overshadow their performance. These effects are especially pronounced in judgments of boys (Bennett et al. 1993). Even the neatness of students' handwriting can significantly affect a teacher's judgment (Sweedler-Brown 1992).

Training programs can help teachers identify and reduce these negative effects and lead to greater consistency in judgments (Afflerbach and Sammons 1991). Unfortunately, few teachers receive adequate training in grading or reporting as part of their preservice experiences (Boothroyd and McMorris 1992). Also, few school districts provide adequate guidance to ensure consistency in teachers' grading or reporting practices (Austin and McCann 1992).

4. Grades have some value as rewards, but no value as punishments. Although educators would undoubtedly prefer that motivation to learn be entirely intrinsic, the existence of grades and other reporting methods are important factors in determining how much effort students put forth (Chastain 1990, Ebel 1979). Most students view high grades as positive recognition of their success, and some work hard to avoid the consequences of low grades (Feldmesser 1971).



© Jeffrey High/Image Productions

At the same time, no studies support the use of low grades as punishments. Instead of prompting greater effort, low grades usually cause students to withdraw from learning. To protect their self-image, many students regard the low grade as irrelevant and meaningless. Other students may blame themselves for the low mark, but feel helpless to improve (Selby and Murphy 1992).

Sadly, some teachers consider grades or reporting forms their

learning criteria, never on the curve. Using the normal probability curve as a basis for assigning grades typically yields greater consistency in grade distributions from one teacher to the next. The practice, however, is detrimental to teaching and learning.

Grading on the curve pits students against one another in a competition for the few rewards (high grades) distributed by the teacher. Under these conditions, students readily see that helping others will threaten their own

chances for success (Johnson et al. 1979, Johnson et al. 1980). Learning becomes a game of

The more detailed the reporting method and the more analytic the process, the more likely subjectivity will influence results.

“weapon of last resort.” In their view, students who don't comply with requests suffer the consequences of the greatest punishment a teacher can bestow: a failing grade. Such practices have no educational value and, in the long run, adversely affect students, teachers, and the relationship they share. Rather than attempting to punish students with a low mark, teachers can better motivate students by regarding their work as incomplete and requiring additional effort.

5. Grading and reporting should always be done in reference to

winners and losers—with most students falling into the latter category (Johnson and Johnson 1989). In addition, modern research has shown that the seemingly direct relationship between aptitude or intelligence and school achievement depends upon instructional conditions, not a probability curve.

When the instructional quality is high and well matched to students' learning needs, the magnitude of this relationship diminishes drastically and approaches zero (Bloom 1976). Moreover, the fairness and equity of grading on the curve is a myth.

Learning Criteria

When grading and reporting relate to learning criteria, teachers have a clearer picture of what students have learned. Students and teachers alike generally prefer this approach because it seems fairer (Kovas 1993). The types of learning criteria usually used for grading and reporting fall into three categories:

■ *Product criteria* are favored by advocates of performance-based approaches to teaching and learning. These educators believe grading and reporting should communicate a summative evaluation of student achievement (Cangelosi 1990). In other words, they focus on what students know and are able to do at that time. Teachers who use product criteria often base their grades or reports exclusively on final examination scores, overall assessments, or other culminating demonstrations of learning.

■ *Process criteria* are emphasized by educators who believe product criteria don't provide a complete picture of student learning. From their perspective, grading and reporting should reflect not just the final results but also *how* students got there. Teachers who consider effort or work habits when reporting on student learning are using process criteria. So are teachers who take into consideration classroom quizzes, homework, class participation, or attendance.

■ *Progress criteria*, often referred to as "improvement scoring" and "learning gain," consider how much students have gained from their learning experiences. Teachers who use progress criteria look at *how far* students have come rather than where they are. As a result, scoring criteria may become highly individualized.

Teachers who base their grading and reporting procedures on learning criteria typically use some combination of the three types (Frary et al. 1993; Nava and Loyd 1992; Stiggins

et al. 1989). Most researchers and measurement specialists, on the other hand, recommend using product criteria exclusively. They point out that the more process and progress criteria come into play, the more subjective and biased grades become (Ornstein 1994). How can a teacher know, for example, how difficult a task was for students or how hard they worked to complete it? If these criteria are included at all, most experts recommend they be reported separately (Stiggins 1994).

Practical Guidelines

Despite years of research, there's no evidence to indicate that one grading or reporting method works best under all conditions, in all circumstances. But in developing practices that seek to be fair, equitable, and useful to students, parents, and teachers, educators can rely on two guidelines:

■ *Provide accurate and understandable descriptions of learning.* Regardless of the method or form used, grading and reporting should communicate effectively what students have learned, what they can do, and

whether their learning status is in line with expectations for that level. More than an exercise in quantifying achievement, grading and reporting must be seen as a challenge in clear thinking and effective communication (Stiggins 1994).

■ *Use grading and reporting methods to enhance, not hinder, teaching and learning.* A clear, easily understood reporting form facilitates communication between teachers and parents. When both parties speak the same language, joint efforts to help students are likely to succeed. But developing such an equitable and understandable system will require the elimination of long-time practices such as averaging and assigning a zero to work that's late, missed, or neglected.

Averaging falls far short of providing an accurate description of what students have learned. For example, students often say, "I have to get a B on the final to pass this course." Such a comment illustrates the inappropriateness of averaging. If a final examination is truly comprehensive and students' scores accu-



© Jeffrey High/Image Productions

A Look Back at Grading Practices

Although student assessment has been a part of teaching and learning for centuries, grading is a relatively recent phenomenon. The ancient Greeks used assessments as formative, not evaluative, tools. Students demonstrated, usually orally, what they had learned, giving teachers a clear indication of which topics required more work or instruction.

In the United States, grading and reporting were virtually unknown before 1850. Back then, most schools grouped students of all ages and backgrounds together with one teacher. Few students went beyond the elementary education offered in these one-room schoolhouses. As the country grew—and as legislators passed compulsory attendance laws—the number and diversity of students increased. Schools began to group students in grades according to their age, and to try new ideas about curriculum and teaching methods. Here's a brief timeline of significant dates in the history of grading:

Late 1800s: Schools begin to issue progress evaluations. Teachers simply write down the skills that students have mastered; once students complete the requirements for one level, they can move to the next level.

Early 1900s: The number of public high schools in the United States increases dramatically. While elementary teachers continue using written descriptions to document student learning, high school teachers introduce percentages as a way to certify students' accomplishments in specific subject areas. Few educators question the gradual shift to percentage grading, which seems a natural by-product of the increased demands on high school teachers.

1912: Starch and Elliott publish a study that challenges percentage grades as reliable measures of student achievement. They base their findings on grades assigned to two papers

written for a first-year English class in high school. Of the 142 teachers grading on a 0 to 100 scale, 15 percent give one paper a failing mark; 12 percent give the same paper a score of 90 or more. The other paper receives scores ranging from 50 to 97. Neatness, spelling, and punctuation influenced the scoring of many teachers, while others considered how well the paper communicated its message.

1913: Responding to critics—who argue that good writing is, by nature, a highly subjective judgment—Starch and Elliott repeat their study but use geometry papers. Even greater variations occur, with scores on one paper ranging from 28 to 95. Some teachers deducted points only for wrong answers, but others took neatness, form, and spelling into account.

1918: Teachers turn to grading scales with fewer and larger categories. One three-point scale, for example, uses the categories of Excellent, Average, and Poor. Another has five categories (Excellent, Good, Average, Poor, and Failing) with the corresponding letters of A, B, C, D, and F (Johnson 1918, Rugg 1918).

1930s: Grading on the curve becomes increasingly popular as educators seek to minimize the subjective nature of scoring. This method ranks students according to some measure of their performance or proficiency. The top percentage receives an A, the next percentage receives a B, and so on (Corey 1930). Some advocates (Davis 1930) even specify the precise percentage of students to be assigned each grade, such as 6–22–44–22–6.

Grading on the curve seems fair and equitable, given research suggesting that students' scores on tests of innate intelligence approximate a normal probability curve (Middleton 1933).

As the debate over grading and reporting intensifies, a number of

schools abolish formal grades altogether (Chapman and Ashbaugh 1925) and return to using verbal descriptions of student achievement. Others advocate pass-fail systems that distinguish only between acceptable and failing work (Good 1937). Still others advocate a "mastery approach": Once students have mastered a skill or content, they move to other areas of study (Heck 1938, Hill 1935).

1958: Ellis Page investigates how student learning is affected by grades and teachers' comments. In a now classic study, 74 secondary school teachers administer a test, and assign a numerical score and letter grade of A, B, C, D, or F to each student's paper. Next, teachers randomly divide the tests into three groups. Papers in the first group receive only the numerical score and letter grade. The second group, in addition to the score and grade, receive these standard comments: *A—Excellent! B—Good work. Keep at it. C—Perhaps try to do still better? D—Let's bring this up. F—Let's raise this grade!* For the third group, teachers mark the score and letter grade, and write individualized comments.

Page evaluates the effects of the comments by considering students' scores on the next test they take. Results show that students in the second group achieved significantly higher scores than those who received only a score and grade. The students who received individualized comments did even better. Page concludes that grades can have a beneficial effect on student learning, but only when accompanied by specific or individualized comments from the teacher.

—Thomas R. Guskey

Source: H. Kirschenbaum, S. B. Simon, and R. W. Napier, (1971), *Wad-ja-get? The Grading Game in American Education*, (New York: Hart).

rately reflect what they've learned, why should a B level of performance translate to a D for the course grade?

Any single measure of learning can be unreliable. Consequently, most researchers recommend using several indicators in determining students' grades or marks—and most teachers

concur (Natriello 1987). Nevertheless, the key question remains, "What information provides the most accurate depiction of students' learning at this time?" In nearly all cases, the answer is "the most current information." If students demonstrate that past assessment information doesn't accu-

rately reflect their learning, new information must take its place. By continuing to rely on past assessment data, the grades can be misleading about a student's learning (Stiggins 1994).

Similarly, assigning a score of zero to work that is late, missed, or neglected doesn't accurately depict

learning. Is the teacher certain the student has learned absolutely nothing, or is the zero assigned to punish students for not displaying appropriate responsibility (Canady and Hotchkiss 1989, Stiggins and Duke 1991)?

Further, a zero has a profound effect when combined with the practice of averaging. Students who receive a single zero have little chance of success because such an extreme score skews the average. That is why, for example, Olympic events such as gymnastics and ice skating eliminate the highest and lowest scores; otherwise, one judge could control the entire competition simply by giving extreme scores. An alternative is to use the median score rather than the average (Wright 1994), but use of the most current information remains the most defensible option.

Meeting the Challenge

The issues of grading and reporting on student learning continue to challenge educators today, just as they challenged Middleton and his colleagues in 1933. But today we know more than ever before about the complexities involved and how certain practices can influence teaching and learning.

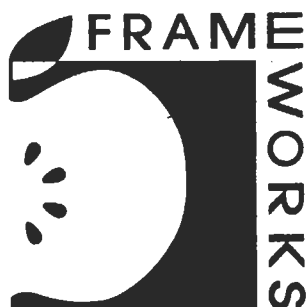
What do educators need to develop grading and reporting practices that provide quality information about student learning? Nothing less than clear thinking, careful planning, excellent communication skills, and an overriding concern for the well being of students. Combining these skills with our current knowledge on effective practice will surely result in more efficient and more effective reporting. ■

References

- Afflerbach, P., and R. B. Sammons. (1991). "Report Cards in Literacy Evaluation: Teachers' Training, Practices, and Values." Paper presented at the annual meeting of the National Reading Conference, Palm Springs, Calif.
- Austin, S., and R. McCann. (1992). "'Here's Another Arbitrary Grade for your Collection': A Statewide Study of Grading Policies." Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Barnes, S. (1985). "A Study of Classroom Pupil Evaluation: The Missing Link in Teacher Education." *Journal of Teacher Education* 36, 4: 46-49.
- Bennett, R. E., R. L. Gottesman, D. A. Rock, and F. Cerullo. (1993). "Influence of Behavior Perceptions and Gender on Teachers' Judgments of Students' Academic Skill." *Journal of Educational Psychology*, 85: 347-356.
- Bishop, J. H. (1992). "Why U.S. Students Need Incentives to Learn." *Educational Leadership* 49, 6: 15-18.
- Bloom, B. S. (1976). *Human Characteristics and School Learning*. New York: McGraw-Hill.
- Bloom, B. S., G. F. Madaus, and J. T. Hastings. (1981). *Evaluation to Improve Learning*. New York: McGraw-Hill.
- Boothroyd, R. A., and R. F. McMorris. (1992). "What Do Teachers Know About Testing and How Did They Find Out?" Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco.
- Brookhart, S. M. (1993). "Teachers' Grading Practices: Meaning and Values." *Journal of Educational Measurement* 30, 2: 123-142.
- Canady, R. L., and P. R. Hotchkiss. (1989). "It's a Good Score! Just a Bad Grade." *Phi Delta Kappan* 71: 68-71.
- Cangelosi, J. S. (1990). "Grading and Reporting Student Achievement." In *Designing Tests for Evaluating Student Achievement*, pp. 196-213. New York: Longman.
- Chapman, H. B., and E. J. Ashbaugh. (October 7, 1925). "Report Cards in American Cities." *Educational Research Bulletin* 4: 289-310.
- Chastain, K. (1990). Characteristics of Graded and Ungraded Compositions." *Modern Language Journal*, 74, 1: 10-14.
- Corey, S. M. (1930). "Use of the Normal Curve as a Basis for Assigning Grades in Small Classes." *School and Society* 31: 514-516.
- Davis, J. D. W. (1930). "Effect of the 6-22-44-22-6 Normal Curve System on Failures and Grade Values." *Journal of Educational Psychology* 22: 636-640.
- Ebel, R. L. (1979). *Essentials of Educational Measurement* (3rd ed.). Englewood Cliffs, N.J.: Prentice Hall.
- Feldmesser, R. A. (1971). "The Positive Functions of Grades." Paper presented at the annual meeting of the American Educational Research Association, New York.
- Frary, R. B., L. H. Cross, and L. J. Weber. (1993). "Testing and Grading Practices and Opinions of Secondary Teachers of Academic Subjects: Implications for Instruction in Measurement." *Educational Measurement: Issues and Practices* 12, 3: 23-30.
- Frisbie, D. A., and K. K. Waltman. (1992). "Developing a Personal Grading Plan." *Educational Measurement: Issues and Practices* 11, 3: 35-42.
- Good, W. (1937). "Should Grades Be Abolished?" *Education Digest* 2, 4: 7-9.
- Heck, A. O. (1938). "Contributions of Research to Classification, Promotion, Marking and Certification." Reported in *The Science Movement in Education (Part II), Twenty-Seventh Yearbook of the National Society for the Study of Education*. Chicago: University of Chicago Press.
- Hill, G. E. (1935). "The Report Card in Present Practice." *Education Methods* 15, 3: 115-131.
- Hills, J. R. (1991). "Apathy Concerning Grading and Testing." *Phi Delta Kappan* 72, 2: 540-545.
- Johnson, D. W., and R. T. Johnson. (1989). *Cooperation and Competition: Theory and Research*. Endina, Minn.: Interaction.
- Johnson, D. W., L. Skon, and R. T. Johnson. (1980). "Effects of Cooperative, Competitive, and Individualistic Conditions on Children's Problem-Solving Performance." *American Educational Research Journal* 17, 1: 83-93.
- Johnson, R. H. (1918). "Educational Research and Statistics: The Coefficient Marking System." *School and Society* 7, 181: 714-116.
- Johnson, R. T., D. W. Johnson, and M. Tauer. (1979). "The Effects of Cooperative, Competitive, and Individualistic Goal Structures on Students' Attitudes and Achievement." *Journal of Psychology* 102: 191-198.
- Kovas, M. A. (1993). "Make Your Grading Motivating: Keys to Performance-Based Evaluation." *Quill and Scroll* 68, 1: 10-11.
- Middleton, W. (1933). "Some General Trends in Grading Procedure." *Education* 54, 1: 5-10.
- Natriello, G. (1987). "The Impact of Evalu-

- ation Processes On Students." *Educational Psychologists* 22: 155-175.
- Nava, F. J. G., and B. H. Loyd. (1992). "An Investigation of Achievement and Nonachievement Criteria in Elementary and Secondary School Grading." Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- O'Donnell, A., and A. E. Woolfolk. (1991). "Elementary and Secondary Teachers' Beliefs About Testing and Grading." Paper presented at the annual meeting of the American Psychological Association, San Francisco.
- Ornstein, A. C. (1994). "Grading Practices and Policies: An Overview and Some Suggestions." *NASSP Bulletin* 78, 559: 55-64.
- Page, E. B. (1958). "Teacher Comments and Student Performance: A Seventy-Four Classroom Experiment in School Motivation." *Journal of Educational Psychology* 49: 173-181.
- Payne, D. A. (1974). *The Assessment of Learning*. Lexington, Mass.: Heath.
- Rugg, H. O. (1918). "Teachers' Marks and the Reconstruction of the Marking System." *Elementary School Journal* 18, 9: 701-719.
- Selby, D., and S. Murphy. (1992). "Graded or Degraded: Perceptions of Letter-Grading for Mainstreamed Learning-Disabled Students." *British Columbia Journal of Special Education* 16, 1: 92-104.
- Starch, D., and E. C. Elliott. (1912). "Reliability of the Grading of High School Work in English." *School Review* 20: 442-457.
- Starch, D., and E. C. Elliott. (1913). "Reliability of the Grading of High School Work in Mathematics." *School Review* 21: 254-259.
- Stewart, L. G., and M. A. White. (1976). "Teacher Comments, Letter Grades, and Student Performance." *Journal of Educational Psychology* 68, 4: 488-500.
- Stiggins, R. J. (1994). "Communicating with Report Card Grades." In *Student-Centered Classroom Assessment*, pp. 363-396. New York: Macmillan.
- Stiggins, R. J., and D. L. Duke. (1991). "District Grading Policies and Their Potential Impact on At-risk Students." Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- Stiggins, R. J., D. A. Frisbie, and P. A. Griswold. (1989). "Inside High School Grading Practices: Building a Research Agenda." *Educational Measurement: Issues and Practice* 8, 2: 5-14.
- Sweedler-Brown, C. O. (1992). "The Effect of Training on the Appearance Bias of Holistic Essay Graders." *Journal of Research and Development in Education* 26, 1: 24-29.
- Wright, R. G. (1994). "Success for All: The Median Is the Key." *Phi Delta Kappan* 75, 9: 723-725.

Thomas R. Guskey is Professor of Education Policy Studies and Evaluation, College of Education, University of Kentucky, Lexington, KY 40506.



A HOLISTIC LITERACY LEARNING

STAFF
DEVELOPMENT
PROGRAM
FOR GRADES K - 8

Authored by
*Jan Turbill, Andrea Butler
and Brian Cambourne
with Gail Langton*

FRAMEWORKS

- provides one time training of in-district facilitators.
- offers long-term support.
- uses the latest research in staff development, language and natural learning.
- includes an ongoing examination of holistic theory and the implications for classroom practice.
- has the flexibility to meet the needs of your district.
- complements Reading Recovery and Chapter 1 teachers moving toward a holistic approach.

1995 FACILITATOR TRAINING SITES

California
Canada
Colorado
Florida
New York

North Carolina
Ohio
Pennsylvania
South Carolina

Texas
Virginia
Washington
Wisconsin

Carol Riley
Office of Staff Development
Wayne-Finger Lakes BOCES
3501 County Road 20
Stanley, NY 14561

Telephone:
(800) 935-2623 or
(716) 526-4669

FAX:
(716) 526-4653

Receive free registrations.
Call for details.

Specialized add-on
modules available
without further training.

NEW
REDUCED PRICE
CALL
FOR DETAILS

(8 0 0) 9 3 5 - 2 6 2 3

Helping Standards Make the GRADE

When reporting on student work, educators need a clear, comprehensive grading system that shows how students are measuring up to standards.

Thomas R. Guskey

The issue of grading looms on the horizon for standards-based education. With standards and assessments now in place, educators face the daunting task of how best to grade and report student learning in terms of those standards. Most educators recognize the inadequacies of their current grading and reporting methods (Marzano, 2000). Few, however, have found alternatives that satisfy the diverse needs of students, parents, teachers, school administrators, and community members.

Standards don't lessen the responsibility of educators to evaluate the performance of students and to report the results. Nevertheless, the focus on standards poses unique challenges in grading and reporting. What are those challenges, and how can educators develop standards-based grading and reports that are accurate, honest, and fair?

Criterion-Referenced Standards

The first challenge is moving from norm-referenced to criterion-referenced grading standards. *Norm-referenced* standards compare each student's performance to that of other students in the group or class. Teachers first rank students on some measure of their achievement or performance. They assign a set percentage of top-ranked students (usually 10 to 20 percent) the highest grade, a second set percentage (perhaps 20 to 30 percent) the second highest grade, and so on. The percentages typically correspond to an approximation of the bell-shaped, normal probability curve, hence the expression "grading on the curve." Most adults experienced this type of grading during their school days.

Criterion-referenced standards, in contrast, compare each student's performance to clearly stated performance descriptions that differentiate levels of quality. Teachers judge students' performance by what each student does, regardless of how well or poorly their classmates perform.

Using the normal probability curve as a basis for assigning grades yields highly consistent grade distributions from one teacher to the next. All teachers' classes have essentially the same percentages of As, Bs, and Cs. But the consequences for students are overwhelmingly negative. Learning becomes highly competitive because students must compete against one another for the few high grades that the teacher distributes. Under these conditions, students see that helping others threatens their own chances for success. Because students do not achieve high grades by performing well, but rather by doing better than their classmates, learning becomes a game of winners and losers, and because teachers keep the number of rewards arbitrarily small, most students must be losers (Haladyna, 1999; Johnson & Johnson, 1989). Strong evidence shows that "grading on the curve" is detrimental to relationships—both among students and among teachers and students (Krumboltz & Yeh, 1996).

In a standards-based system, grading and reporting must be criterion-referenced. Teachers at all levels must identify what they want their students to learn and be able to do and what

A second challenge is to differentiate the types of grading criteria that teachers will use. Although teachers and students generally consider criterion-referenced grading to be more fair and equitable (Kovas, 1993), the specific grading criteria that teachers use may be very diverse. We can classify these criteria into three broad categories: *product*, *process*, and *progress* (Guskey, 1996).

Standards don't lessen the responsibility of educators to evaluate the performance of students and to report the results.

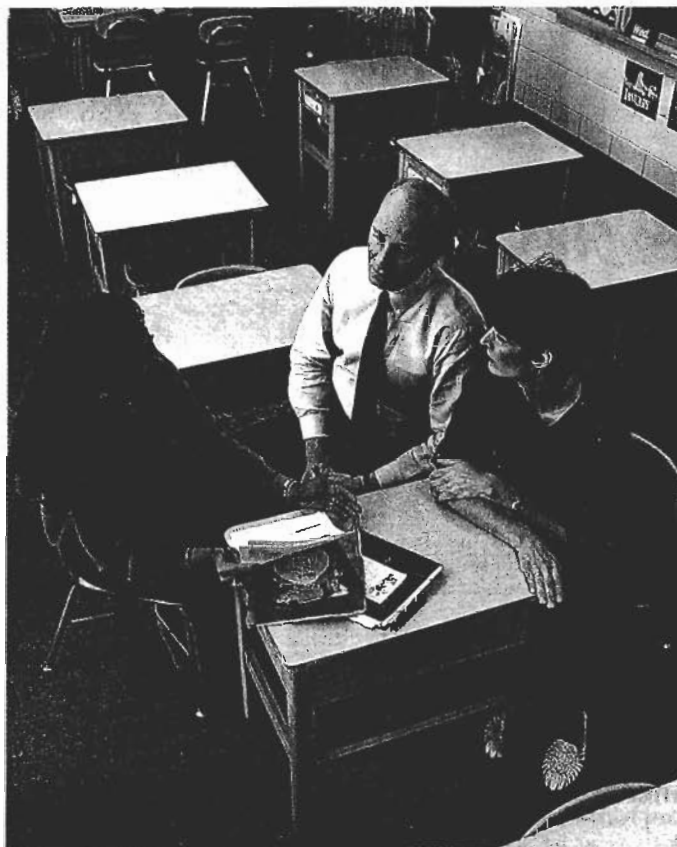
Process criteria relate not to the final results, but to how students got there. Educators who believe that product criteria do not provide a complete picture of student learning generally favor process criteria. For example, teachers who consider student effort, class behavior, or work habits are using process criteria. So are those who count



daily work, regular classroom quizzes, homework, class participation, punctuality of assignments, or attendance in determining students' grades.

Progress criteria relate to how much students actually gain from their learning experiences. Other terms include learning gain, improvement grading, value-added grading, and educational growth. Teachers who use progress criteria typically look at how far students have come rather than where students are. Others attempt to judge students' progress in terms of their "learning potential." As a result, progress grading criteria are often highly individualized among students.

Because they are concerned about student motivation, self-esteem, and the social consequences of grading, few teachers today use product criteria solely in determining grades. Instead, most base their grading on some combination of criteria, especially when a student receives only a single grade in a subject area (Brookhart, 1993; Frary, Cross, & Weber, 1993). The majority of teachers also vary the criteria they use from student to student, taking into account individual circumstances (Truog & Friedman, 1996). Although teachers do so in an effort to be fair, the result is often a hodgepodge grade that includes elements of achievement, effort, and improvement (Brookhart, 1991). Interpreting the grade or report thus becomes difficult for parents, administrators, community members, and even the students (Friedman & Frisbie, 1995). An A, for example, may mean that the student knew what the teacher expected before instruction began (product), didn't learn as well as expected but tried very hard (process), or simply made significant improvement (progress).



displeasure with indifference, deception, or disruption (Tomlinson, 1992).

A practical solution to this problem, and one that increasing numbers of teachers and schools are using, is to establish clear indicators of product, process, and progress, and then to report each separately (Stiggins, 2001; Wiggins, 1996). Teachers separate grades or marks for learning skills, effort, work habits, or progress from grades for achievement and performance. Parents generally prefer this approach because it gives them more detailed and prescriptive information. It also simplifies reporting for teachers because they no longer have to combine so many diverse types of information into a single grade. The key to success, however, rests in the

Measurement experts generally recommend using product criteria exclusively in determining students' grades. They point out that the more process and progress criteria come into play, the more subjective and biased grades are likely to be (O'Connor, 1999; Ornstein, 1994). How can a teacher know, for example, how difficult a task was for students or how hard they worked to complete it?

Many teachers, however, point out that if they use product criteria exclusively, some high-ability students receive high grades with little effort, whereas the hard work of less-talented students is seldom acknowledged. Others say that if teachers consider only product criteria, low-ability students and those who are disadvantaged—students who must work the hardest—have the least incentive to do so. These students find the relationship between high effort and low grades unacceptable and, as a result, often express their

clear specification of those indicators and the criteria to which they relate. This means that teachers must describe how they plan to evaluate students' achievement, effort, work habits, and progress, and then must communicate these plans directly to students, parents, and others.

Reporting Tools

A third challenge for standards-based education is clarifying the purpose of each reporting tool. Although report cards are the primary method, most schools today use a variety of reporting devices: weekly or monthly progress reports, open-house meetings, newsletters, evaluated projects or assignments, school Web pages, parent-teacher conferences, and student-led conferences (Guskey & Bailey, 2001). Each reporting tool must fulfill a specific purpose, which requires considering three vital aspects of communication:

- What information do we want to communicate?
- Who is the primary audience for that information?
- How would we like that information to be used?

Many educators make the mistake of choosing their reporting tools first, without giving careful attention to the purpose. For example, some charge headlong into developing a standards-based report card without first addressing core questions about why they are doing it. Their efforts often encounter unexpected resistance and rarely bring positive results. Both parents and teachers perceive the change as a newfangled fad that presents no real advantage over traditional reporting methods. As a result, the majority of these efforts become short-lived experiments and are abandoned after a few troubled years of implementation.

Efforts that begin by clarifying the purpose, however, make intentions clear from the start. If, for instance, the purpose of the report card is to communicate to parents the achievement status of students, then parents must understand the information on the report card and know how to use it. This means that educators should include parents on report card committees and give their input careful consideration. This not only helps mobilize everyone in the reporting process, it also keeps efforts on track. The famous adage that guides architecture also applies to grading and reporting: *Form follows function*. Once the purpose or function is clear, teachers can address more easily questions regarding form or method (Guskey & Bailey, 2001).

Developing a Reporting Form

The fourth challenge for standards-based education is developing the centerpiece of a standards-based reporting system: the report card. This typically involves a four-step process. First, teams of educators identify the major learning goals or

standards that students are expected to achieve at each grade level or course of study. Second, educators establish performance indicators for those learning goals or standards. In other words, educators decide what evidence best illustrates students' attainment of each goal or standard. Third, they determine graduated levels of quality for assessing student performance. This step involves identifying incremental levels of attainment, sometimes referred to as benchmarks, as students progress toward the learning goals or standards (Andrade, 2000; Wiggins & McTighe, 1998). Finally, educators, often in collaboration with parents, develop a reporting form that communicates teachers' judgments of students' progress and achievement in relation to the learning goals or standards.

are too broad or general, however, make it hard to identify students' unique strengths and weaknesses. Most state-level standards, for example, tend to be broad and need to be broken down or "unpacked" into homogeneous categories or topics (Marzano, 1999). For grading and reporting purposes, educators must seek a balance. The standards must be broad enough to allow for efficient communication of student learning, yet specific enough to be useful (see Gronlund, 2000; Marzano & Kendall, 1995; Wiggins & McTighe, 1998).

Another issue is the differentiation of standards across marking periods or grade levels. Most schools using standards-based grading develop reporting forms that are based on

Many parents initially respond to a standards-based reporting form with, "This is great. But tell me, how is my child doing *really*?"

Identifying Reporting Standards

Identifying the specific learning goals or standards on which to base grades is probably the most important, but also the most challenging, aspect of standards-based grading. These learning goals or standards should stipulate precisely what students should know and be able to do as a result of their learning experiences. In earlier times, we might have referred to cognitive skills, learning competencies, or performance outcomes (Guskey, 1999). Teachers frequently list these learning goals in their lesson plans, make note of them on assignments and performance tasks, and include them in monthly or weekly progress reports that go home to parents.

A crucial consideration in identifying learning goals or standards is determining the degree of specificity. Standards that are too specific make reporting forms cumbersome to use and difficult to understand. Standards that

grade-level learning goals or standards. Each standard has one level of complexity set for each grade that students are expected to meet before the end of the academic year. Most parents, however, are accustomed to grading systems in which learning standards become increasingly complex with each marking period. If the standard states "Students will write clearly and effectively," for example, many parents believe that their children should do this each marking period, not simply move toward doing so by the end of the academic year. This is especially true of parents who encourage their children to attain the highest mark possible in all subject areas every marking period.

To educators using such forms, students who receive 1 or 2 on a 4-point grading scale during the first or second marking period are making appropriate progress and are on track for their grade level. For parents,

however, a report card filled with 1s and 2s, when the highest mark is a 4, causes great concern. They think that their children are failing. Although including a statement on the reporting form, such as "Marks indicate progress toward end-of-the-year learning standards," is helpful, it may not alleviate parents' concerns.

Facilitating Interpretation

Many parents initially respond to a standards-based reporting form with, "This is great. But tell me, how is my child doing *really*?" Or they ask, "How is my child doing compared to the other children in the class?" They ask these questions because they don't know how to interpret the information. Further, most parents had

comparative, norm-based reporting systems when they were in school and are more familiar with reports that compare students to their classmates. Above all, parents want to make sense of the reporting form. Their fear is that their children will reach the end of the school year and won't have made sufficient progress to be promoted to the next grade.

Example of a Double-Mark, Standards-Based Reporting Form

Elementary Progress Report

Reading	1st	2nd	3rd	4th
Understands and uses different skills and strategies	1+	2++		
Understands the meaning of what is read	1++	2+		
Reads different materials for a variety of purposes	1-	2-		
Reading level		1++	2+	
Work habits	S	S		
Writing	1st	2nd	3rd	4th
Writes clearly and effectively	1+	2++		
Understands and uses the steps in the writing process	1++	2++		
Writes in a variety of forms for different audiences and purposes	1+	2-		
Analyzes and evaluates the effectiveness of written work	N	1+		
Understands and uses the conventions of writing: punctuation, capitalization, spelling, and legibility	1-	2-		
Work habits	S	S		
Communication	1st	2nd	3rd	4th
Uses listening and observational skills to gain understanding	1+	2-		
Communicates ideas clearly and effectively (formal communication)	1-	2+		
Uses communication strategies and skills to work effectively with others (informal communication)	N	1+		
Work habits	U	S		

This report is based on grade-level standards established for each subject area. The ratings indicate your student's progress in relation to the year-end standard.

Evaluation Marks

4 = Exceptional
3 = Meets standard
2 = Approaches standard
1 = Beginning standard
N = Not applicable

Level Expectation Marks

++ = Advanced
+ = On level
- = Below level

Social Learning Skills & Effort Marks

E = Exceptional
S = Satisfactory
U = Unsatisfactory

To ensure more accurate interpretations, several schools use a two-part marking system with their standards-based reporting form (see example). Every marking period, each student receives two marks for each standard. The first mark indicates the student's level of progress with regard to the standard—a 1, 2, 3, or 4, indicating *beginning*, *progressing*, *proficient*, or *exceptional*. The second mark indicates the relation of that level of progress to established expectations at this point in the school year. For example, a ++ might indicate *advanced for grade-level expectations*, a + might indicate *on target or meeting grade-level expectations*, and a - would indicate *below grade-level expectations or needs improvement*.

The advantage of this two-part marking system is that it helps parents make sense of the reporting form each marking period. It also helps alleviate their concerns about what seem like low grades and lets them know whether their children are progressing at an appropriate rate. Further, it helps parents take a standards-based perspective in viewing their children's performances. Their question is no longer "Where is my child in comparison to his or her classmates?" but "Where is my child in relation to the grade-level learning goals and expectations?"

The one drawback of the two-part marking system is that expectations must take into account individual differences in students' development of cognitive skills. Because students in any classroom differ in age and cognitive development, some might not meet the specified criteria during a particular marking period—even though they will likely do so before the end of the year. This is especially common in kinder-



garten and the early primary grades, when students tend to vary widely in their entry-level skills but can make rapid learning progress (Shuster, Lemma, Lynch, & Nadeau, 1996). Educators must take these developmental differences into consideration and must explain them to parents.

Choosing Performance-Level Descriptors

Standards-based reporting forms that use numerical grading scales also require a key or legend that explains the meaning of each numeral. These descriptors help parents and others understand what each numeral means.

A common set of descriptors matches performance levels 1, 2, 3, and 4 with the achievement labels *beginning*, *progressing*, *proficient*, and *exceptional*. If the standards reflect behavioral aspects of students' performance, then teachers more commonly use such descriptors as *seldom*, *sometimes*, *usually*, and *consistently/independently*. These labels are preferable to *above average*, *average*, and *below average*, which reflect norm-referenced comparisons rather than criterion-

referenced standards.

Such achievement descriptors as *exceptional* or *advanced* are also preferable to *exceeds standard* or *extending* to designate the highest level of performance. Educators can usually articulate specific performance criteria for an *exceptional* or *advanced* level of achievement or performance. *Exceeds standard* or *extending*, however, are much less precise and may leave students and parents wondering just what they need to do to exceed or extend. Descriptors should be clear, concise, and directly interpretable.

Many reporting forms include a fifth level of *not applicable* or *not evaluated* to designate standards that have not yet been addressed or were not assessed during that particular marking period. Including these labels is preferable to leaving the marking spaces blank because parents often interpret a blank space as an item that the teacher missed or neglected.

Maintaining Consistency

A final challenge is consistency. To communicate with parents, most schools and school districts involved in

courses of study. Because of these curricular differences, standards-based reporting forms at the middle and secondary levels must vary from student to student. The marks need to relate to each student's achievement and performance in his or her particular courses or academic program. Although advances in technology, such as computerized reporting forms, allow educators to provide such individualized reports, relatively few middle and high school educators have taken up the challenge.

The standards must be broad enough to allow for efficient communication of student learning, yet specific enough to be useful.

New Standards for Grading

As educators clarify student learning goals and standards, the advantages of standards-based grading become increasingly evident. Although it makes reporting forms more detailed and complex, most parents value the richness of the information when the reports are expressed in terms that they can understand and use. Reporting forms that use a two-part marking system show particular promise—but such a system may require additional explanation to parents. Teachers must also set expectations for learning progress not just at the grade level, but also for each marking period.

Successfully implementing standards-based grading and reporting demands a close working relationship among teachers, parents, and school and district leaders. To accurately interpret the reporting form, parents need to know precisely what the standards mean and how to make sense of the various levels of achievement or performance in relation to those standards. Educators must ensure, therefore, that parents are familiar with the language and terminology. Only when all groups understand what grades mean and how they

are used to improve student learning will we realize the true value of a standards-based approach to education. ■

References

- Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, 57(5), 13-18.
- Brookhart, S. M. (1991). Grading practices and validity. *Educational Measurement: Issues and Practice*, 10(1), 35-36.
- Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement*, 30(2), 123-142.
- Frary, R. B., Cross, L. H., & Weber, L. J. (1993). Testing and grading practices and opinions of secondary teachers of academic subjects: Implications for instruction in measurement. *Educational Measurement: Issues and Practice*, 12(3), 23-30.
- Friedman, S. J., & Frisbie, D. A. (1995). The influence of report cards on the validity of grades reported to parents. *Educational and Psychological Measurement*, 55(1), 5-26.
- Gronlund, N. E. (2000). *How to write and use instructional objectives* (6th ed.). Upper Saddle River, NJ: Merrill.
- Guskey, T. R. (1996). Reporting on student learning: Lessons from the past—Prescriptions for the future. In T. R. Guskey (Ed.), *Communicating student learning: 1996 Yearbook of the Association for Supervision and Curriculum Development* (pp. 13-24). Alexandria, VA: ASCD.
- Guskey, T. R. (1999). Making standards work. *The School Administrator*, 56(9), 44.
- Guskey, T. R., & Bailey, J. M. (2001). *Developing grading and reporting systems for student learning*. Thousand Oaks, CA: Corwin Press.
- Haladyna, T. M. (1999). *A complete guide to student grading*. Boston: Allyn & Bacon.
- Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Endina, MN: Interaction.
- Kendall, J. S., & Marzano, R. J. (1995). *The systematic identification and articulation of content standards and benchmarks: Update*. Aurora, CO: McREL.
- Kovas, M. A. (1993). Make your grading motivating: Keys to performance-based evaluation. *Quill and Scroll*, 68(1), 10-11.
- Krumboltz, J. D., & Yeh, C. J. (1996). Competitive grading sabotages good teaching. *Phi Delta Kappan*, 78(4), 324-326.
- Marzano, R. J. (1999). Building curriculum and assessment around standards. *The High School Magazine*, 6(5), 14-19.
- Marzano, R. J. (2000). *Transforming classroom grading*. Alexandria, VA: ASCD.
- O'Connor, K. (1999). *How to grade for learning*. Arlington Heights, IL: Skyline.
- Ornstein, A. C. (1994). Grading practices and policies: An overview and some suggestions. *NASSP Bulletin*, 78(559), 55-64.
- Shuster, C., Lemma, P., Lynch, T., & Nadeau, K. (1996). *A study of kindergarten and 1st grade report cards: What are young children expected to learn?* Paper presented at the annual meeting of the American Educational Research Association, New York.
- Stiggins, R. J. (2001). *Student-involved classroom assessment* (3rd ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.
- Tomlinson, T. (1992). *Hard work and high expectations: Motivating students to learn*. Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.
- Truog, A. L., & Friedman, S. J. (1996). *Evaluating high school teachers' written grading policies from a measurement perspective*. Paper presented at the annual meeting of the National Council on Measurement in Education, New York.
- Wiggins, G. (1996). Honesty and fairness: Toward better grading and reporting. In T. R. Guskey (Ed.), *Communicating student learning: 1996 Yearbook of the Association for Supervision and Curriculum Development* (pp. 141-176). Alexandria, VA: ASCD.
- Wiggins, G., & McTighe, J. (1998). *Understanding by design*. Alexandria, VA: ASCD.

Thomas R. Guskey is Professor of Educational Policy Studies and Evaluation, College of Education, University of Kentucky, Lexington, KY 40506; guskey@pop.uky.edu.

Making High School Grades Meaningful

Most teachers base students' grades on more than one factor. The difficulty is figuring out how to weight and combine the different pieces that go into the final mark. Mr. Guskey suggests a system that not only avoids those problems but gives a better overall picture of a student's performance than the traditional single letter grade.

BY THOMAS R. GUSKEY

MICHAEL AND Sheila attend the same high school and take many of the same classes. Michael is an exceptionally bright but obstinate student. He consistently gets high grades on classroom quizzes and tests, even though he rarely completes homework assignments and is often tardy. His compositions and reports show keen insight and present thoughtful analyses of critical issues but are usually turned in two or three days late. Because of his missing homework assignments and lack of punctuality, Michael receives C's in most of his classes, and his grade-point average lands him in the middle of his high school class rankings. But Michael scores at the highest level on the state

THOMAS R. GUSKEY is a professor in the College of Education, University of Kentucky, Lexington. ©2006, Thomas R. Guskey.



accountability assessment and qualifies for an honors diploma.

Sheila, on the other hand, is an extremely dedicated and hard-working student. She completes every homework assignment, takes advantage of extra-credit options in all of her classes, and regularly attends special study sessions held by her teachers. Yet, despite her efforts, Sheila often performs poorly on classroom quizzes and tests. Her compositions and reports are well organized and turned in on time but rarely demonstrate more than a surface understanding of critical issues. Sheila also receives C's in most of her classes and has a class ranking very similar to Michael's. But because she scores at a low level on the state accountability assessment, Sheila is at risk of receiving an alternative diploma.

A rare situation, you say? Unlikely or even impossible? Ask any high school teacher today and most will tell you that they know students very much like Michael and Sheila. Many will admit that they currently have similar students in their classes. While Michael and Sheila may not be typical high school students, they also are not unusual.

How is it possible for students with such different levels of demonstrated knowledge and skill to receive essentially the same grades in their high school classes? How can they have roughly the same grade-point average and class ranking? What does this tell us about the meaning of high school grades and the students who receive those grades? And, most important, what does this tell us about the grading policies and practices of many high school teachers?

HODGEPODGE GRADING

Many educators contend that the problem lies in the accountability assessments. They believe that the discrepancy between high school course grades and scores on state accountability assessments demonstrates the inadequacy and invalidity of the assessment results.¹ Indeed, these narrow once-a-year assessments may not reveal the true scope or depth of students' knowledge and skills. On the other hand, policy makers argue that teachers are the source of the problem. They think the mismatch between grades and scores on accountability assessments stems from bias and subjectivity in teachers' grading practices.² There is ample evidence that most teachers receive little training in effective grading and that unintentional bias often influences teachers' grade assignments.³ However, a more likely explanation lies in the nature of grading itself and in the

challenges teachers face in assigning grades that offer a fair and accurate picture of students' achievement and performance.

High school teachers today draw from many different sources of evidence in determining students' grades, and studies show that teachers differ in the procedures they use to combine or summarize that evidence.⁴ Some of the major sources of evidence teachers use include:

- | | |
|---------------------------------|---|
| • Major exams or compositions | • Homework completion |
| • Class quizzes | • Homework quality |
| • Reports or projects | • Class participation |
| • Student portfolios | • Work habits and neatness |
| • Exhibits of student work | • Effort |
| • Laboratory projects | • Attendance |
| • Student notebooks or journals | • Punctuality of assignment submissions |
| • Classroom observations | • Class behavior or attitude |
| • Oral presentations | • Progress made |

When asked *which* of these sources of evidence they consider in determining students' grades, some portion of teachers will report using each one of the elements on the list. When asked *how many* of these sources of evidence they include, however, responses vary widely. Some teachers base grades on as few as two or three elements, while others incorporate evidence from as many as 15 or 16 — and this is true even among teachers who teach in the same school.

Two factors seem to account for this variation. First is a lack of clarity about the purpose of grading. Decisions about what evidence to use in determining students' grades are extremely difficult to make when the purpose of grading is unclear. Different sources of evidence vary in their appropriateness and validity depending on the identified purpose.

A second reason for the variation is the format used to report grades. Most high school reporting forms allow only a single grade to be assigned to students for each course or subject area. This compels teachers to distill all of these diverse sources of evidence into a single symbol. The result is a "hodgepodge grade" that includes elements of achievement, attitude, effort, and behavior.⁵ Even when teachers clarify the weighting strategies they use to combine these elements and employ computerized grading programs to ensure accuracy in their computations, the final grade remains a confusing amalgamation that is impossible to interpret and

rarely presents a true picture of a student's proficiency.⁶

To make high school grades more meaningful, we need to address both of these factors. First, we must clarify our purpose in grading. Second, we must decide what evidence best serves that purpose and how best to communicate a summary of that evidence to parents and others.

CLARIFYING PURPOSES AND CRITERIA

When asked to identify the purpose of grading, most high school teachers indicate that grades should describe how well students have achieved the learning goals established for a course. In other words, grades should reflect students' performance based on specific learning criteria. Teachers and students alike prefer this approach because they consider it both fair and equitable.⁷ But, as described earlier, teachers use widely varying criteria to determine students' grades. In most cases, these can be grouped into three broad categories: *product*, *process*, and *progress* criteria.

Product criteria are favored by advocates of standards-based or performance-based approaches to teaching and learning. These educators believe the primary purpose of grading is to communicate a summative evaluation of student achievement and performance.⁸ In other words, they seek to assess *what* students know and are able to do at a particular point in time. Teachers who use product criteria typically base grades exclusively on final examination scores, final reports or projects, overall assessments, and other culminating demonstrations of learning.

Process criteria are emphasized by educators who believe product criteria do not provide a complete picture of student learning. From their perspective, grades should reflect not only the final results but also *how* students got there. Teachers who consider effort or work habits when assigning grades are using process criteria, as are teachers who factor regular classroom quizzes, homework, punctuality of assignments, class participation, or attendance into grade calculations.

Progress criteria are used by educators who believe that the most important aspect of grading is how much students have gained from their learning experiences. Other names for progress criteria include "learning gain," "improvement scoring," "value-added learning," and "educational growth." Some educators draw distinctions between progress, which they measure backward from a final performance standard or goal, and growth, which is measured forward from the place a

student begins on a learning continuum.⁹ However, when achievement is judged using well-defined learning standards that include graduated levels of performance, progress and growth criteria can be considered synonymous.

Teachers who use progress criteria typically look at how much improvement students have made over a specified period of time, rather than just where they are at any one point. As a result, the scoring criteria used in determining student grades may be highly individualized. Most of the current research evidence on the use of progress criteria in grading comes from studies of individualized instruction and special education programs.¹⁰

Because of concerns about student motivation, self-esteem, and the social consequences of grades, few teachers use only product criteria in determining grades. Instead, most routinely base their grading procedures on some combination of all three types of evidence.¹¹ Many also vary their grading criteria from student to student, taking into account individual circumstances.¹² Although teachers defend this practice on the basis of fairness, it seriously blurs the meaning of any grade. Interpreting grades thus becomes exceptionally challenging, not only for parents but also for administrators, community members, and even the students themselves.¹³ A grade of A, for example, may mean that the student knew what was intended before instruction began (product), did not learn as well as expected but tried very hard (process), or simply made significant improvement (progress).

CONFLICTING SOLUTIONS

Recognizing these interpretation problems, most researchers and measurement specialists recommend the exclusive use of product criteria in determining students' grades. They point out that the more process and progress criteria come into play, the more subjective and biased grades become.¹⁴ How can a teacher know, for example, how difficult a task was for students or how hard they worked to complete it?

Many teachers point out, however, that if they use only product criteria in determining grades, some high-ability students will receive high grades with little effort, while the hard work of less-talented students will go unacknowledged. Consider, for example, two students enrolled in the same physical education class. The first is a well-coordinated athlete who can easily perform any task the teacher asks and so typically does not put forth serious effort. The second student is strug-

gling with a weight problem but consistently tries hard, exerts extraordinary effort, and also displays exceptional sportsmanship and cooperation. Nevertheless, this student is unable to perform at the same level as the athlete. Few teachers would consider it fair to use only product criteria in determining the grades of these two students.¹⁵

Teachers also emphasize that, if only product criteria are considered, low-ability students and those who are disadvantaged — the students who must work hardest — have the least incentive to do so. These students find the relationship between high effort and low grades frustrating and often express their frustration with indifference, deception, or disruption.¹⁶

A MEANINGFUL ALTERNATIVE

An increasing number of teachers and schools have adopted a practical solution to the problems associated with incorporating these different learning criteria into student grades: they report separate grades or marks on each set of criteria. In other words, after establishing explicit indicators of product, process, and progress criteria, teachers assign a separate grade to each. In this way grades or marks for learning skills, effort, work habits, and learning progress are kept distinct from assessments of achievement and performance.¹⁷ The intent is to provide a better, more accurate, and much more comprehensive picture of what students accomplish in school.

While high school teachers in the United States are

just beginning to catch on to the idea of separate grades for product, process, and progress criteria, many Canadian educators have used the practice for years.¹⁸ Each marking period teachers assign students an “achievement” grade based on the students’ performance on projects, assessments, and other demonstrations of learning. Often expressed as a letter grade or percentage (A = advanced, B = proficient, C = basic, D = needs improvement, F = unsatisfactory), this “achievement” grade represents the teacher’s judgment of the student’s level of performance or accomplishment relative to explicit learning goals established for the course. Computations of grade-point averages and class ranks are based solely on these “achievement” or product grades.

In addition, teachers also assign separate grades or marks for homework, class participation, punctuality of assignment submissions, effort, learning progress, and the like. Because these factors usually relate to specific student behaviors, most teachers record numerical marks for each (4 = consistently, 3 = usually, 2 = sometimes, and 1 = rarely). To clarify a mark’s meaning, teachers identify specific behavioral indicators for these factors and for the levels of performance in each. For example, the indicators for a “homework” mark might include:

4 = All homework assignments completed and turned in on time.

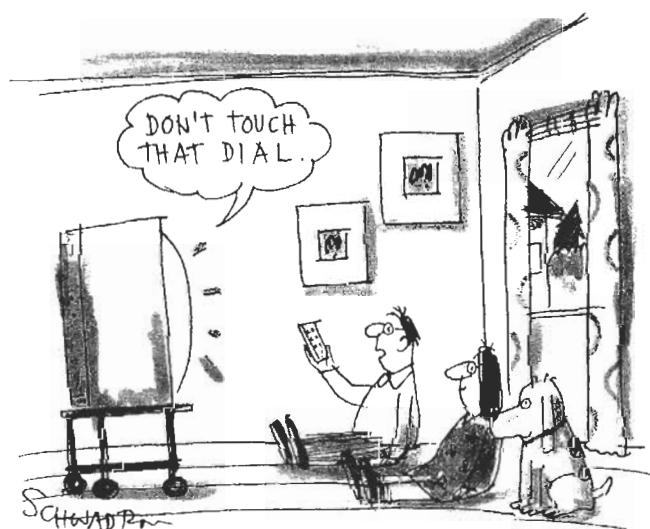
3 = Only one or two missing or incomplete homework assignments.

2 = Three to five missing or incomplete homework assignments.

1 = Numerous missing or incomplete homework assignments.

Teachers sometimes question the need for this level of specificity. Upon reflection, however, most discover that by including homework assignments as part of an overall grade for students, they already face this challenge. When determining an overall grade, teachers must decide how much credit to give students for completing homework assignments or how much to take away for assignments that were turned in late or not at all. Similarly, when reporting a separate grade for homework, teachers must ensure that students understand the various performance levels so that they know what the mark signifies and what must be done to improve.

Often teachers presume that reporting multiple grades will increase their grading workload. But those who use the procedure claim that it actually makes grading easier and less work. Teachers gather the same evidence on student learning that they did when calculating an over-



“What’s a dial?”

all grade but no longer worry about how to weight or combine that evidence. As a result, they avoid irresolvable arguments about the appropriateness or fairness of various weighting strategies.

Reporting separate grades for product, process, and progress criteria also makes grading more meaningful. If a parent questions the teacher about a product grade, for example, the teacher simply points to the various

The key to success in reporting multiple grades rests on the clear specification of indicators related to product, process, and progress criteria.

process indicators and suggests, "Perhaps if your child completed homework assignments and participated more in class, the 'achievement' grade would be higher." Parents favor the practice because it provides a more comprehensive profile of their child's performance in school. Employers and college admission officers also like systems of separate grades because they offer more detailed information on students' accomplishments. With all grades reported on the transcript, a college admissions office can distinguish between the student who earned high achievement grades with relatively little effort and the one who earned equally high grades through diligence and hard work. The transcript thus becomes a more robust document, presenting a better and more discerning portrait of students' high school experiences.¹⁹

Schools would still have the information needed to compute grade-point averages and class rankings, if such computations are still deemed important. Now, however, those averages and rankings would be untainted by undefined aspects of process and progress. As such, they would represent a more valid and appropriate measure of achievement and performance. Furthermore, to the extent that classroom assessments and state accountability assessments are based on the same standards for learning, the relationship between product grades and accountability assessment results would likely be much higher.

The key to success in reporting multiple grades, however, rests on the clear specification of indicators related to product, process, and progress criteria. Teachers must be able to describe exactly how they plan to evaluate students' achievement, attitude, effort, behavior, and

progress. Then they must clearly communicate these criteria to students, parents, and others.

CONCLUSION

The relationship between high school grades and students' performance on state accountability assessments will never be perfect. Grades are derived from courses that can vary significantly across schools and classrooms. In contrast, state accountability assessments typically are designed to measure proficiency based on a set of common standards for student learning. As such, the developers of these types of assessments purposefully avoid content that may be unique to particular learners or learning situations. Furthermore, course grades normally reflect a much broader range of knowledge and skills than can be measured by limited accountability assessments with restricted modes of student response.²⁰ Nevertheless, concerns about honesty and fairness compel us to reduce the mismatch between these two important measures of student knowledge and skill.

Developing meaningful, reasonable, and equitable grading policies and practices will continue to challenge high school educators. The challenge remains all the more daunting, however, if we continue to use reporting forms that require teachers to combine so many diverse sources of evidence into a single grade. Distinguishing specific "product" criteria on which to base an "achievement" grade allows teachers to offer a better and more precise description of students' academic achievement and performance. To the extent that "process" criteria related to homework, class participation, attitude, effort, responsibility, behavior, and other nonacademic factors remain important, they too can be reported. But they should be reported separately. Adopting this approach will clarify the meaning of grades and greatly enhance their communicative value.

1. See, for example, Robert T. Brennan et al., "The Relative Equitability of High-Stakes Testing Versus Teacher-Assigned Grades: An Analysis of the Massachusetts Comprehensive Assessment System (MCAS)," *Harvard Educational Review*, Summer 2001, pp. 173-216.

2. Randy Elliot Bennett et al., "Influence of Behavior Perceptions and Gender on Teachers' Judgments of Students' Academic Skill," *Journal of Educational Psychology*, June 1993, pp. 347-56; John R. Hills, "Apathy Concerning Grading and Testing," *Phi Delta Kappan*, March 1991, pp. 540-45; and Clyde A. Wiles, "Investigating Gender Bias in the Evaluations of Middle School Teachers of Mathematics," *School Science and Mathematics*, October 1992, pp. 295-98.

3. Susan Austin and Richard McCann, "Here's Another Arbitrary Grade for Your Collection: A Statewide Study of Grading Policies," paper presented at the annual meeting of the American Educational Research Association, San Francisco, 1992; Roger A. Boothroyd and Robert F. McMorris, "What Do Teachers Know About Testing and How Did They

Find Out?," paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, 1992; and Richard J. Stiggins, "Inside High School Grading Practices: Building a Research Agenda," *Educational Measurement: Issues and Practice*, Summer 1989, pp. 5-14.

4. Gregory J. Cizek, Shawn M. Fitzgerald, and Robert E. Racho, "Teachers' Assessment Practices: Preparation, Isolation, and the Kitchen Sink," *Educational Assessment*, vol. 3, 1995/1996, pp. 159-79; and James H. McMillan, "Secondary Teachers' Classroom Assessment and Grading Practices," *Educational Measurement: Issues and Practice*, March 2001, pp. 20-32.

5. Susan M. Brookhart, "Grading Practices and Validity," *Educational Measurement: Issues and Practice*, March 1991, pp. 35-36; and Lawrence H. Cross and Robert B. Frary, "Hodgepodge Grading: Endorsed by Students and Teachers Alike," paper presented at the annual meeting of the National Council on Measurement in Education, New York, 1996.

6. See David T. Conley, "Who Is Proficient?: The Relationship Between Proficiency Scores and Grades," paper presented at the annual meeting of the American Educational Research Association, New Orleans, 2000; and Thomas R. Guskey, "Computerized Gradebooks and the Myth of Objectivity," *Phi Delta Kappan*, June 2002, pp. 775-80.

7. Marcia A. Kovas, "Make Your Grading Motivating: Keys to Performance Based Evaluation," *Quill and Scroll*, October/November 1993, pp. 10-11.

8. Stephen J. Friedman, "Grading Teachers' Grading Policies," *NASSP Bulletin*, April 1998, pp. 77-83; and Ken O'Connor, *How to Grade for Learning: Linking Grades to Standards*, 2nd ed. (Arlington Heights, Ill.: SkyLight, 2002).

9. See, for example, Grant Wiggins, "Honesty and Fairness: Toward Better Grading and Reporting," in Thomas R. Guskey, ed., *Communicating Student Learning: 1996 Yearbook of the Association for Supervision and Curriculum Development* (Alexandria, Va.: Association for Supervision and Curriculum Development, 1996), pp. 141-76.

10. Warren W. Esty and Anne R. Teppo, "Grade Assignment Based on Progressive Improvement," *Mathematics Teacher*, November 1992, pp. 616-18; and Russell Gersten, Sharon Vaughn, and Susan U. Brengelmann, "Grading and Academic Feedback for Special Education Students and Students with Learning Difficulties," in Guskey, *Communicating Student Learning*, pp. 47-57.

11. Susan M. Brookhart, "Teachers' Grading Practices: Meaning and Values," *Journal of Educational Measurement*, Summer 1993, pp. 123-142; Cross and Frary, op. cit.; Stephen J. Friedman and Merlin Manley, "Improving High School Grading Practices: Experts vs. Practitioners," *NASSP Bulletin*, vol. 76, 1992, pp. 100-104; Fe' Josefa G. Nava and Brenda H. Loyd, "An Investigation of Achievement and Nonachievement Criteria in Elementary and Secondary School Grading," paper presented at the annual meeting of the American Educational Research Association, San Francisco, 1992.

12. Stephen J. Friedman and Anthony J. Troug, "Evaluation of High School Teachers' Written Grading Policies," *ERS Spectrum*, Summer 1999, pp. 34-42; and Gary Natriello, Carolyn J. Riehl, and Aaron M. Pallas, *Between the Rock of Standards and the Hard Place of Accommodation: Evaluation Practices of Teachers in High Schools Serving Disadvantaged Students* (Baltimore: Center for Research on Effective Schooling for Disadvan-

THE LOS ANGELES COUNTY OFFICE OF EDUCATION PRESENTS PARENT EXPECTATIONS SUPPORT ACHIEVEMENT (PESA)

Facilitator training for parent workshop leaders

Help parents prepare their children for success.

Become a Certified PESA Facilitator and lead parent workshops at your school!

PESA fulfills the requirement of providing parent involvement activities to improve student academic achievement and school performance for the federal reform legislation of the No Child Left Behind Act of 2001 (Title I, Sec. 1118, Parent Involvement).

PESA facilitator workshops are available in English, Spanish, Chinese, Korean, and Armenian languages upon request.

English PESA Facilitator Trainings are scheduled for:

May 15-16, 2006 – Detroit, MI	Jan. 25-26, 2007 – San Bernardino, CA
Aug. 1-2, 2006 – San Diego, CA	Feb. 6-7, 2007 – Miami, FL
Sep. 20-21, 2006 – Bakersfield, CA	Feb. 27-28, 2007 – Houston, TX
Oct. 17-18, 2006 – Virginia Beach, VA	Mar. 13-14, 2007 – Anaheim, CA
Nov. 7-8, 2006 – Chicago, IL	May 10-11, 2007 – Ventura, CA
Nov. 28-29, 2006 – San Francisco, CA	May 15-16, 2007 – Oklahoma City, OK

- The \$300 registration fee includes the 2-day training, PESA Facilitator Manual, instructional video, interaction wall chart, and refreshments.

- Please call (800) 566-6651 for a Registration Form with locations.

Schedule a PESA Facilitator Training at your site and receive a discount on registration fees.

To request a registration form or additional information regarding the TESA or PESA programs, please call (800) 566-6651.



Look for the TESA training schedule on page 687 of this issue.

E-mail: tesa_pes@lacoe.edu Website: <http://streamer.lacoe.edu/PESA>



tagged Students, Johns Hopkins University, 1994).

13. Stephen J. Friedman and David A. Frisbie, "The Influence of Report Cards on the Validity of Grades Reported to Parents," *Educational and Psychological Measurement*, February 1995, pp. 5-26; Kristie K. Waltman and David A. Frisbie, "Parents' Understanding of Their Children's Report Card Grades," *Applied Measurement in Education*, vol. 7, 1994, pp. 223-40; and Warren W. Willingham, Judith M. Pollack, and Charles Lewis, "Grades and Test Scores: Accounting for Observed Differences," *Journal of Educational Measurement*, Spring 2002, pp. 1-37.

14. Allan C. Ornstein, "Grading Practices and Policies: An Overview and Some Suggestions," *NASSP Bulletin*, April 1994, pp. 55-64.

15. See Thomas R. Guskey and Jane M. Bailey, *Developing Grading and Reporting Systems for Student Learning* (Thousand Oaks, Calif.: Corwin Press, 2001).

16. Tommy Tomlinson, *Hard Work and High Expectations: Motivating Students to Learn* (Washington, D.C.: Office of Educational Research and Improvement, U.S. Department of Education, 1992).

17. Thomas R. Guskey, *How's My Kid Doing? A Parent's Guide to Grades, Marks, and Report Cards* (San Francisco: Jossey Bass, 2002); and Richard J. Stiggins, "Communicating with Report Cards," in idem, *Student-Involvement Assessment for Learning*, 4th ed. (Upper Saddle River, N.J.: Pearson Education, 2005), pp. 275-319.

18. Jane Bailey and Jay McTighe, "Reporting Achievement at the Secondary Level: What and How," in Guskey, *Communicating Student Learning*, pp. 119-40.

19. Clifford Adelman, *Answers in the Tool Box: Academic Intensity, Attendance Patterns, and Bachelor's Degree Attainment* (Washington, D.C.: Office of Educational Research and Improvement, U.S. Department of Education, 1999).

20. Willingham, Pollack, and Lewis, op. cit.

K

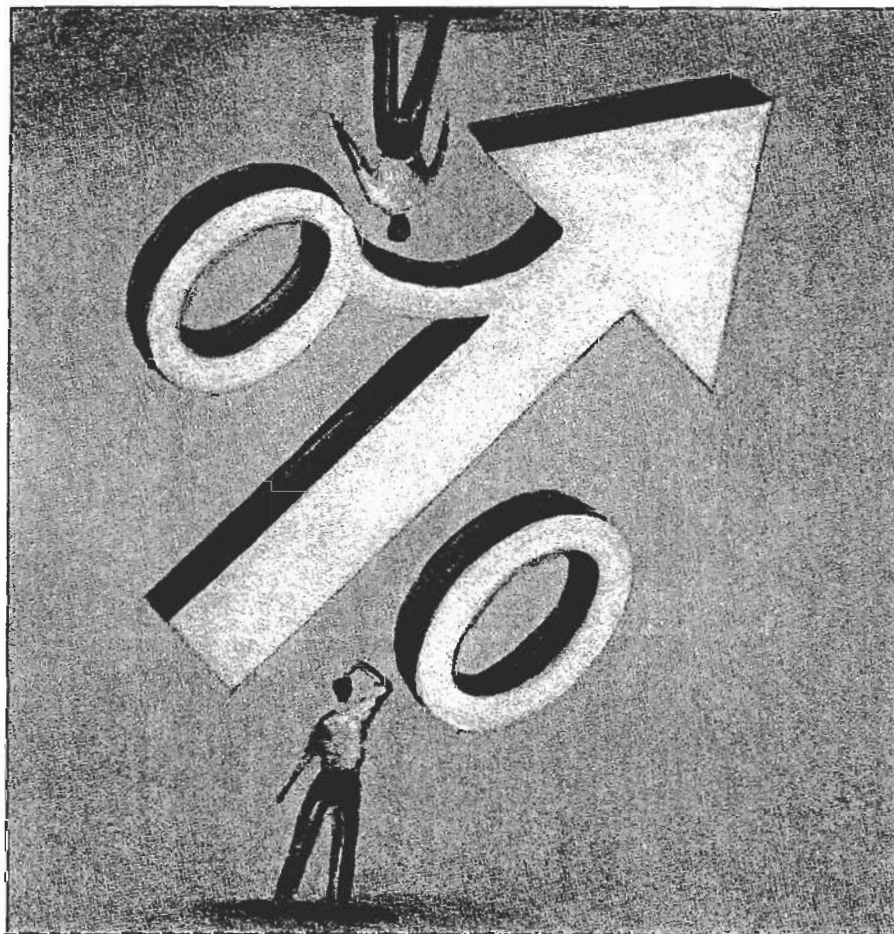
High Percentages Are *Not* The Same as High Standards

Mr. Guskey reminds us that, even when complex statistical formulas are used in setting cutoff scores, their mathematical precision is not a substitute for sound professional judgment.

BY THOMAS R. GUSKEY

HOW TO set appropriate cutoff scores for student performance on state assessments and other high-stakes examinations is a widely debated issue in education today. Typically these debates focus on what percentage of items students should be expected to answer correctly in order to have their performance judged "proficient" or "competent." On the Texas Assessment of Academic Skills (TAAS), for example, students must answer 70% correct in order to attain a passing score. This debate often extends to the classroom level, where teachers set cutoff scores for different grades. What percentage correct should students be expected to attain, for instance, to receive a grade of A or a grade of B, and so on?

THOMAS R. GUSKEY is a professor of Educational Policy Studies and Evaluation, College of Education, University of Kentucky, Lexington.



Both policy makers and teachers generally assume that higher cutoff percentages mean more rigorous standards and higher expectations for student performance. A cutoff of 80% correct for proficiency in mathematics, for instance, is considered more rigorous than a 70% correct cutoff for proficiency in language arts. Similarly, the teacher who sets 95% correct as the cutoff for a grade of A is considered to be more demanding and to have higher standards than the teacher who uses a cut-

off of only 90% or 92% correct for an A. This reasoning leads to the belief that raising the percentage for a cutoff is one way to raise both the standards and the expectations we set for student performance.

Unfortunately, it isn't quite that simple. Setting cutoff percentages for assessments and for grades is an arbitrary decision that says little about the standards or the expectations set for students' learning. A much more important consideration is the difficulty of the tasks students are asked

to perform or the cognitive complexity of the questions they are required to answer.

The cutoff percentage representing an excellent level of performance on an extremely challenging task or a very difficult set of questions might be quite different from the cutoff percentage considered excellent on a relatively simple task. This does not imply that the challenge is determined strictly by how well other students perform (i.e., norm-referenced). Rather, it means that tasks or items designed to assess a given learning goal (i.e., criterion-referenced) can vary widely in their intricacy and cognitive complexity.

Suppose, for example, we were interested in assessing students' basic knowledge about the Presidents of the United States. We could ask an open-ended, constructed-response question (also known as a "short-answer" or "completion" item):

Who was the 17th President of the United States?

Fewer than 10% of students are able to answer this question correctly. Its high level of difficulty is actually rather odd because most people know that Abraham Lincoln was the 16th President, and they know that the name of the President who succeeded him was Johnson. Putting these two pieces of information together, however, proves quite difficult for the vast majority.

We might then consider framing the same question as a multiple-choice, selected-response item. For example:

Who was the 17th President of the United States?

- A. Abraham Lincoln
- B. Andrew Johnson
- C. Ulysses S. Grant
- D. Millard Fillmore

This remains a fairly difficult item for most students. Because of the multiple-choice format, however, about 30% are now able to answer correctly. Of course, if all students simply chose an answer at random, the limited-response, multiple-choice format would allow 25% to select the correct response.

Suppose we next adjust the possible responses, making the distinctions a bit more obvious:

Who was the 17th President of the United States?

- A. George Washington
- B. Andrew Johnson
- C. Jimmy Carter
- D. Bill Clinton

Now identifying the correct response is much easier, and about 60% of students are able to answer correctly. We could probably assume that those who are still unable to identify the correct response have very limited knowledge of U.S. Presidents.

Of course, we could make a final adjustment to the possible responses in order to make the item easier still:

Who was the 17th President of the United States?

- A. The War of 1812
- B. Andrew Johnson
- C. The Louisiana Purchase
- D. A Crazy Day for Sally

About 90% of students are able to answer this item correctly. Those who don't are usually drawn to the response "A Crazy Day for Sally" because they recognize it as the one response that doesn't belong with the others.

Some might argue that knowing who was the 17th President of the United States is a rather trivial learning outcome — and that might be true. The point is that, while each of these items assesses the same learning objective, same goal, or same achievement target, each varies greatly in its difficulty.

Suppose that there were four assessments designed to measure students' subject-area proficiency or their achievement in a high school course. Assessment 1 consisted of items of the first type described above; assessment 2 consisted of items of the second type, and so on. Those four assessment devices would present vastly different challenges to students, and the scores students attained on such assessments would undoubtedly reflect those differences. Would it be fair to set the same "proficiency" cutoff percentage for each of those four assessments?

The Challenge of Setting Appropriate Cutoffs

Focusing on a percentage correct as a



"Basically, what you're saying is I get a box of chocolate chip cookies, and the sixth-grade class gets a field trip to Tuscany?"

cutoff is seductive but very misleading because tests and assessments vary widely in how they are designed. Some assessments include items that are so challenging that students who answer a low percentage of items correctly still do very well.

Take the Graduate Record Examinations (GRE), for example, a series of tests used to determine admission to graduate schools. Individuals who answer only 50% of the questions correctly on the GRE physics test perform better than more than 70% of those who take the test (already a highly self-selected group). For the GRE mathematics test, 50% correct would outperform approximately 60% of the individuals who take the test. And among those who take the GRE literature test, only about half get 50% correct.² In most classrooms, of course, students who answer only 50% correct would receive a failing grade.

Should we conclude from this information that prospective graduate students in physics, mathematics, and literature are a bunch of "failures"? Of course not. Without careful examination of the questions or tasks students are asked to address,

cutoff percentages are just not that meaningful.

Researchers suggest that an appropriate approach to setting cutoffs must combine teachers' judgments of the importance of the concepts addressed and consideration of the cognitive processing skills required by the items or tasks.³ Using this type of cutoff or grade-assignment procedure shifts teachers' thinking so that grades on classroom assessments and other demonstrations of learning reflect the quality of student thinking instead of simply the number of points attained. It incorporates the value the teacher places on successful performance and the teacher's perception of the level of thinking that students must use to answer a question or perform a task.

Sadly, this ideal is seldom realized. Rarely does such thought and consideration go into setting the cutoffs for students' performance or the grades they receive. Even in high-stakes assessment situations in which the consequences for students can be quite serious, this level of deliberative judgment is uncommon.

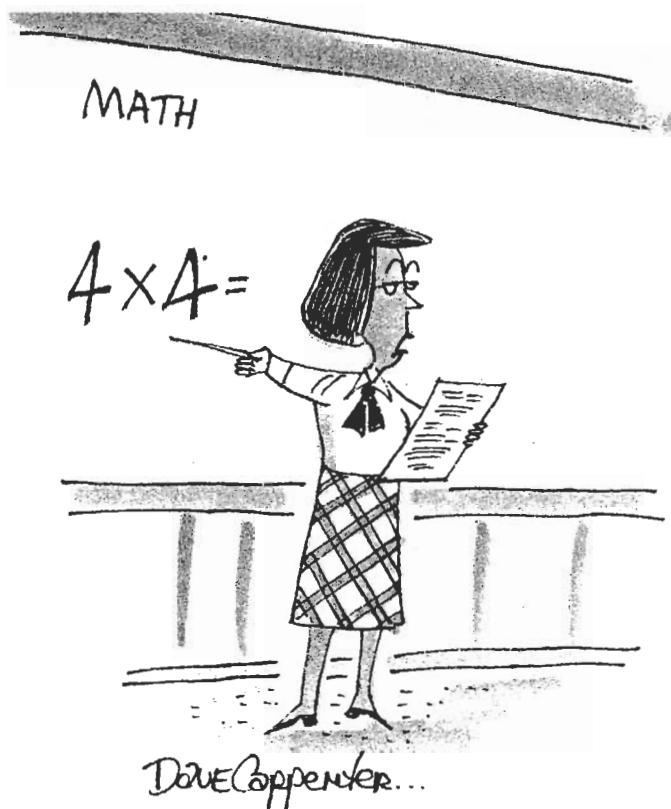
Making matters even more complicated is the fact that the challenge or diffi-

culty of an assessment task is also directly related to the quality of the teaching. Students who are taught well and provided ample opportunities to practice and demonstrate what they have learned are likely to find well-aligned performance tasks or assessment questions much easier than students who are taught poorly and given few practice opportunities. Hence, a 90% cutoff might be relatively easy to meet for students who are taught well, while a 70% cutoff might prove exceptionally difficult for those students who experience poor-quality teaching.

Conclusion

The point of this discussion is not that cutoff percentages are unimportant. They are a vital and necessary consideration in any assessment of student learning. However, setting cutoffs is a more complex process than most policy makers and educators anticipate, and it is typically much more arbitrary than most imagine.⁴

What we must keep in mind is that, even when complex statistical formulas are used in setting cutoffs, their mathematical precision is not a substitute for sound professional judgment. Raising standards or increasing expectations for students' learning is not accomplished simply by raising the cutoff percentages for performance levels or different grade categories. Raising standards requires thoughtful examination of the tasks students are asked to complete and the questions they are asked to answer in order to demonstrate their learning. It might also involve taking into account the quality of the teaching students experienced prior to the assessment. Only when such judgment becomes a regular part of the assessment process will we be able to make accurate and valid decisions about the quality of students' performance.



"No, 'off-road vehicle' is incorrect."

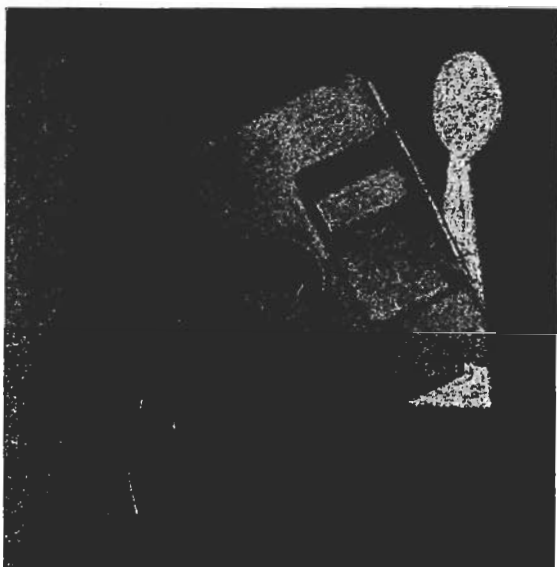
1. This item was developed by Professor Jeffrey Smith of Rutgers University.

2. Drew H. Gitomer and Mari A. Pearlman, "Are Teacher Licensing Tests Too Easy? Are Standards Too Low?," *ETS Developments*, vol. 45, 1999, pp. 4-5.

3. Anthony J. Nitko and Boleslaw Niemierko, "Qualitative Letter Grade Standards for Teacher-Made Summative Classroom Assessments," paper presented at the annual meeting of the American Educational Research Association, Atlanta, 1993.

4. Thomas R. Guskey and Jane M. Bailey, *Developing Grading and Reporting Systems for Student Learning* (Thousand Oaks, Calif.: Corwin, 2001). **K**

Computerized Gradebooks And the Myth Of Objectivity



Computerized grading programs and electronic gradebooks can be useful tools. But in the end, Mr. Guskey reminds us, teachers must still decide what grade offers the most accurate and fairest

description of each student's achievement and level of performance.

BY THOMAS R. GUSKEY

IF YOU ASK middle school or high school teachers today how they determine their students' grades, the first thing most of them will do is open a computerized grading program. They'll show you the vast array of data they keep on each student and explain how they weigh the different pieces of information. At the end of the marking period, they combine these various measures and, with the help of the computer, calculate a summary score to the one-hundred-thousandth of a decimal point. The computer then converts this summary score into the letter grade that is printed on a report card and sent home to parents. Many teachers will also go on to describe the fairness and objectivity of this process, pointing out how the mathematical precision of the computer makes it easy for them to explain and to defend their grading policies to students, to parents, and to administrators.

But do computerized gradebooks really make grading fairer and more objective? Or have the technical capabilities of these programs seduced teachers and school leaders into a false sense of confidence in the

accuracy and validity of the grades they assign?

COMPUTERIZED GRADEBOOKS

Computerized grading programs and electronic gradebooks rank among the best-selling computer software available to educators today. They appeal to teachers primarily because they simplify record-keeping. The spreadsheet formats and database management systems

*THOMAS R. GUSKEY is a professor in the College of Education, University of Kentucky, Lexington. This article is based on material drawn from *Developing Grading and Reporting Systems for Student Learning*, by Thomas R. Guskey and Jane M. Bailey (Corwin Press, 2001). © 2002, Thomas R. Guskey.*

TABLE 1.
Summary Grades Tallied by Three Different Methods

Student	Unit 1	Unit 2	Unit 3	Unit 4	Unit 5	Average Score	Grade	Median Score	Grade	Deleting Lowest	Grade
1	59	69	79	89	99	79	C	79	C	84	B
2	99	89	79	69	59	79	C	79	C	84	B
3	77	80	80	78	80	79	C	80	B	79.5	C
4	49	49	98	99	100	79	C	98	A	86.5	B
5	100	99	98	49	49	79	C	98	A	86.5	B
6	0	98	98	99	100	79	C	98	A	98.8	A
7	100	99	98	98	0	79	C	98	A	98.8	A

Grading Scale: 90%-100%=A, 80%-89%=B, 70%-79%=C, 60%-69%=D, 59% or lower=F.

included in these programs make it easy for teachers to enter and tally precisely large amounts of numerical information.¹ Thus they are suited particularly well to the point-based grading systems of middle school and high school teachers, who often record numerical data on the performance of more than 100 students each week.

Most computerized grading programs also present educators with a wide range of options. Some simply help teachers to keep more detailed records on students' learning progress.² Others allow teachers to present summaries of their students' achievement and performance in a variety of different formats, including computer displays, online reports, and even digital portfolios. Still other programs actually perform grading tasks. The simplest of these scan, mark, and analyze assessments composed of true/false, matching, and multiple-choice items. More recently, however, exciting advances have been made in the use of computers to evaluate and grade students' essays, compositions, and other writing samples.³

For all their advantages, however, computerized grading programs also have their shortcomings. Perhaps the most serious is that they lead the educators who use them to believe that mathematical precision necessarily brings greater objectivity and enhanced fairness to grading. Many teachers assume that, so long as the mathematical calculations are correct and all students are treated the same, then the grades assigned are accurate and just. But numerical precision is not the same as evaluative fairness, honesty, or truth. While computerized grading programs and electronic gradebooks may greatly simplify record-keeping, they do not lessen the challenge involved in assigning grades that accurately and honestly reflect students' level of performance.

MATHEMATICAL PRECISION VERSUS VALID GRADES

Consider, for example, the data in Table 1. The scores on the left side of the table reflect the performance of seven students over five instructional units. The scores on the right represent summary scores for these students calculated by three different methods. The first method is the simple arithmetic average of the unit scores, with all units receiving equal weight. The second is the median or middle score from the five units.⁴ Because the median is positional rather than proportional, it's not influenced by extreme scores, as is an average. The third method is also an arithmetic average, but with the lowest unit score in the group deleted. This method is based on the assumption that no one, including students, performs at a peak level all the time.⁵ These are the three tallying methods most frequently used by teachers and most commonly employed in computerized grading programs and electronic gradebooks.

Consider, too, the following explanations for these score patterns:

- Student 1 struggled in the early part of the marking period but continued to work hard, improved in each unit, and performed excellently in unit 5.
- Student 2 began with excellent performance in unit 1 but then lost motivation, declined steadily during the marking period, and received a failing mark for unit 5.
- Student 3 performed steadily throughout the marking period, receiving three B's and two C's, both near the cutoff between B and C.
- Student 4 began the marking period poorly and failed the first two units but, with newfound interest, performed excellently in units 3, 4, and 5.

- Student 5 began the marking period excellently but then lost interest and failed the last two units.

- Student 6 skipped school (an unexcused absence) during the first unit but performed excellently in every other unit.

- Student 7 performed excellently in the first four units but was caught cheating on the assessment for unit 5 and received a score of zero for that unit.

As is evident from Table 1, all three of these tallying methods are mathematically precise. Yet each one yields a very different pattern of grades for these seven students. If you use the simple arithmetic average, all seven students would receive the same grade of C. If you use the median, there would be just two C's, one B, and four A's. And if you use an arithmetic average with the lowest score deleted, there would be just one C, four B's, and two A's. Note, too, that the one student who would receive a grade of C using this third method had unit grades of just two C's and three B's. More important, not one student would receive the same grade across all three methods. In fact, two students (Student 4 and Student 5) could receive a grade of A, B, or C, depending on the tallying method you use.

The teacher responsible for assigning grades to the performance of these seven students has to answer a number of difficult questions. For example, which of these three methods is fairest? Which method provides the most accurate summary of each student's achievement and level of performance? Do all seven students deserve the same grade, as using the arithmetic average suggests, or are there defensible reasons to justify different grades for certain students? And if there are reasons to justify different grades, can these reasons be clearly specified? Can they be fairly and equitably applied to the performance of all students? Can these reasons be clearly communicated to students before instruction begins? Would it be fair to apply them if they were not communicated to students?

The nature of the assessment information from which these scores are derived could make matters even more tangled. It might make a difference, for example, if the content of each unit assessment was cumulative. In other words, the assessment for unit 2 contained material from units 1 and 2, and the unit 5 assessment included material from all five previous units. And if it did, would this make these grading decisions any easier, or would it further complicate summary calculations?

What should be evident in this example is that the use of computerized grading programs won't solve these

complex grading problems. Although such programs can simplify numerical record-keeping, the mathematical precision they offer does not make the grading process any more objective or any fairer. Calculating a summary score to the one-hundred-thousandth of a decimal point doesn't yield a more accurate depiction of students' achievement and level of performance. Each teacher still must decide what information goes into the calculation, what weight will be attached to each source of information, and what method will be used to tally and summarize that information.

This example also illustrates several questionable grading practices that computerized grading programs typically ignore. Although not new and certainly not inherent in the use of technology in grading, the potentially harmful effects of these practices make it imperative that educators carefully examine their impact and consider other alternatives. Three such practices include 1) averaging scores to determine a grade, 2) the use of zeroes, and 3) taking credit away from students or lowering their grade because of behavioral infractions.

AVERAGING SCORES TO DETERMINE A GRADE

If a mark or grade is supposed to represent an accurate description of how well students have learned, as most experts on grading agree it should,⁶ then the practice of averaging generally falls far short. For example, how often have you heard students lament, "I have to get an A on the final exam in order to pass this course"? But does this situation really make sense, or does it illustrate the inappropriateness of averaging? If a final examination or summative performance truly represents a comprehensive assessment of what students have learned, how can an A level of performance there translate to a C or D for the course grade? Similarly, if a final grade is to reflect what students have learned and can do at the end of the course, can averaging scores from past assessments with measures of current performance be considered appropriate?

Educators generally recognize learning as a progressive and incremental process. Most also agree that students should have multiple opportunities to demonstrate their learning. But is it fair to consider all these learning trials in determining students' grades? If at any time in the instructional process students demonstrate that they've learned the concepts well and mastered the intended learning goals, doesn't that make all previous information on their learning of those con-



cepts inaccurate and invalid? Why then should such information be “averaged in” when determining students’ grades?

Because any single measure of learning can be unreliable, most researchers recommend using several indicators to determine students’ marks or grades.⁷ Nevertheless, teachers must continually ask themselves, “What information provides the most accurate depiction of students’ learning at this time?” In nearly all cases, the answer is “the most current information.” If students demonstrate that past assessment results no longer accurately reflect their learning, that information must be discarded and replaced by the new information. Continuing to rely on past assessment data miscommunicates students’ achievement. Can you imagine, for example, the karate teacher suggesting that a student who starts with a white belt but then progresses to earn a black belt actually deserves a gray belt?

Averaging can also have detrimental effects on student motivation. Suppose, for example, that a student does poorly on one or two major assessments administered early in the marking period, as was the case with Student 4 and Student 6 in Table 1. Knowing that those scores will be “averaged in” as part of the final grade, what motivation do these students have to do well on other assessments? Even if they perform at the highest level from that time on, the practice of averaging gives them virtually no chance of attaining a high grade.

And consider this extreme but true occurrence. A high school student I know experienced the death of

a beloved family member during the first marking period of his senior year. The trauma of that experience proved exceptionally difficult for this young man. As a result, he neglected his schoolwork completely and received failing grades in all his courses. But then, with help from counselors, family and community members, and his teachers, he recovered emotionally, re-dedicated himself to his schooling, and with diligent effort attained A’s in all his courses during the remaining three marking periods of the school year. Because of his school’s policy of averaging, however, his final course grades were all C’s. Did those C’s accurately reflect what he had learned? Did they represent what he had accomplished? Did they adequately describe his achievement or level of performance? Was this fair?

Recognizing that single measures of student learning can be flawed or unreliable, most teachers use multiple sources of information when assigning marks or grades. But simply combining all such measures and calculating an average is rarely appropriate or fair. Some educators argue that the median or middle score provides a more appropriate measure,⁸ but that practice, too, can be problematic.

To provide an accurate summary of students’ performance, teachers must begin by looking for consistency in the evidence gathered. If that evidence is consistent across several indicators, then deciding what grade to assign is relatively straightforward. This would be the case, for example, for students who obtained very similar scores on a class project, on two summative examinations, and on an oral report. But even these cases get complicated when scores consistently fall near the cutoff between two grades. Note, for example, the scores of Student 3 in Table 1.

If the evidence of student achievement is inconsistent, then teachers must look deeper and search for the reasons why.⁹ They also have to face the difficult challenge of deciding what evidence or combination of evidence represents the truest and most appropriate summary of students’ achievement and performance. In such cases, three general guidelines can be recommended.¹⁰

First, the most recent evidence should always be given priority or greater weight. Because grades are usually meant to represent students’ current achievement status or level of performance, the most accurate evidence is generally the evidence collected most recently. Therefore, scores from assessments at the end of the marking period are typically more representa-

tive of what students have learned than those collected at the beginning.

A second strategy is to give priority or greater weight to the most comprehensive forms of evidence. If certain sources of evidence represent cumulative summaries of the knowledge and skills students have acquired, then these should hold the greatest weight in determining students' grades. Exceptions to this approach might be necessary, however, for students who suffer inordinate test or performance anxiety. Such students typically do remarkably well on assignments, quizzes, and class discussions, but then "freeze" during larger assessments or performances. In these cases, teachers may have to consider other means of gathering evidence, such as orally questioning those students or providing some other means for them to demonstrate their learning, in order to get a more valid representation of what they can do.

A third approach would be to "rank order" the evidence gathered in terms of its importance to the learning goals or standards of the course. Those sources of evidence that relate to the most important goals or standards should then be given priority. For example, teachers might attach greater importance to students' scores on a project that required them to synthesize and apply what they had learned than they might give to the scores students attained on assessments designed to tap basic knowledge and comprehension of course content.

Whatever strategy teachers choose, they must be sure to apply that strategy consistently. Although exceptions to accommodate unusual or extenuating circumstances are always permissible, fairness in grading dictates that teachers inform students about their grading policies and practices in advance and then faithfully and consistently apply those policies.

THE USE OF ZEROES

Few teachers believe that grades should be used to punish students for their lack of effort or for demonstrating inadequate responsibility. At the same time, however, many teachers assign zeroes to student work that is missed, neglected, or turned in late.¹¹ Obviously, if grades are to represent how well students have learned, then the practice of assigning zeroes for "administrative or behavioral" reasons clearly misses the mark.

Zeroes have an even more profound effect if combined with the practice of averaging. Students who receive a single zero have little chance of success because

such an extreme score so drastically skews the average. (Note, for example, the scores of Student 6 and Student 7 in Table 1.) For this reason, in scoring Olympic events like gymnastics and diving, the highest and lowest judges' scores are always eliminated before the averaging takes place. If they were not, a single judge could control the results of an entire competition simply by giving extreme scores.

Some teachers defend the practice of assigning zeroes by arguing that they cannot give students credit for work that is incomplete or not turned in — and that's certainly true. But there are far better ways to motivate and encourage students to complete assignments than by assigning them zeroes, especially considering the overwhelmingly negative effects.

One alternative approach is to assign an "incomplete" and then require students to do additional work to bring their performance up to an acceptable level. Students who miss an assignment or neglect a project deadline, for example, might be required to attend after-school study sessions or special Saturday school programs in order to complete their work. In other words, these students are not "let off the hook" with a zero. Instead, students learn that they have responsibilities in school and that their actions have specific consequences. In addition, it helps to make the grade a more accurate reflection of what the students have actually learned.

LOWERING GRADES BECAUSE OF BEHAVIOR

Another typical grading practice with detrimental effects is lowering students' grades because of behavioral infractions. Some teachers lower students' grades for classroom disruptions and similar forms of misconduct. Other teachers consider tardiness or class attendance in determining students' grades and often reduce the grades of students who are late or who miss class sessions. Teachers also vary widely in how they handle such offenses as plagiarism, copying another student's work, and other forms of "cheating." But most teachers weigh such transgressions heavily when determining students' grades.

Student 6 and Student 7 in Table 1 offer excellent examples. Although Student 6 performed exceptionally well throughout most of the marking period, a zero due to an unexcused absence could severely affect his or her course grade. Student 7 performed excellently in four units but was then caught cheating on the assessment for unit 5 and received a zero. Most teachers

would undoubtedly consider this a fair response to Student 7's infraction. But when it comes to determining this student's course grade, the issues become thornier. Some teachers would look at the achievement history over the marking period, conclude that this incident was an exception, and assign the student a high grade. Others would reason that the high marks in earlier units could well have been attained through cheating as well, although the student didn't get caught. Hence, they would feel justified in assigning a lower grade.

The essential question the teacher must address in each of these cases is, "What is the purpose of grading?" If the purpose of grading is to present a summary judgment of students' achievement and level of performance, then to count these behavioral infractions in determining the grade clearly miscommunicates. Although such infractions cannot be ignored, it's clear that they are not part of the evidence that shows what these students have learned and are able to do.

A better strategy is to report these behavioral infractions separately and not include them as part of the course grade. For example, in a growing number of schools, reporting forms are designed to include indicators of students' class behaviors and work habits in addition to grades representing their achievement and level of performance.¹² In other words, teachers report "multiple grades" in each course, separating evidence of students' learning from information about their behavior and conduct.

Some educators might feel that reporting multiple grades makes both record-keeping and grading procedures overly complicated. But those who use this approach report that it actually simplifies grading. They collect no additional information from students and have eliminated the final step of having to combine these diverse sources of evidence. By separating the different aspects of students' performance in school, these teachers provide more specific information to parents and to students. In addition, they are able to identify more clearly students' strengths as well as areas in which improvement is needed.

Computerized grading programs and electronic gradebooks greatly simplify the record-keeping tasks teachers face. They allow teachers to collect and efficiently summarize large amounts of data on student learning. But the efficiency and mathematical precision of these programs does not make the grades they generate more accurate, honest, fair, or objective.

Grading requires careful planning, thoughtful judgment, a clear focus on purpose, excellent communication skills, and an overriding concern for the well-being of students — qualities that no computer possesses. Teachers at all levels must make carefully reasoned decisions about which components will be included in determining students' grades, how those components will be combined and summarized, and what format will be used to report the summaries. While computerized grading programs and electronic gradebooks can be useful tools, they do not relieve teachers of the professional responsibilities involved in making these crucial decisions. In the end, teachers must still decide what grade offers the most accurate and fairest description of each student's achievement and level of performance.

1. Joe Huber, "Gradebook Programs: Which Ones Make the Grade?," *Technology Connection*, vol. 4, no. 1, 1997, pp. 21-23; and Edward L. Vockell and Douglas J. Fiore, "Electronic Gradebooks: What Current Programs Can Do for Teachers," *Clearing House*, vol. 66, 1993, pp. 141-45.
2. Kenneth W. Eastwood, "Reporting Student Progress: One District's Attempt with Student Literacy," in Thomas R. Guskey, ed., *Communicating Student Learning: 1996 Yearbook of the Association for Supervision and Curriculum Development* (Alexandria, Va.: Association for Supervision and Curriculum Development, 1996), pp. 65-78.
3. Ellis B. Page and Nancy S. Petersen, "The Computer Moves into Essay Grading: Updating the Ancient Test," *Phi Delta Kappan*, March 1995, pp. 561-65; and William Wresch, "The Imminence of Grading Essays by Computer — 25 Years Later," *Computers and Composition*, vol. 10, no. 2, 1993, pp. 45-58.
4. Russell G. Wright, "Success for All: The Median Is the Key," *Phi Delta Kappan*, May 1994, pp. 723-25.
5. Robert Lynn Canady and Phyllis Riley Hotchkiss, "It's a Good Score! Just a Bad Grade," *Phi Delta Kappan*, September 1989, pp. 68-71.
6. Thomas M. Haladyna, *A Complete Guide to Student Grading* (Boston: Allyn & Bacon, 1999); Robert J. Marzano, *Transforming Classroom Grading* (Alexandria, Va.: Association for Supervision and Curriculum Development, 2000); Ken O'Connor, *How to Grade for Learning* (Arlington Heights, Ill.: Skylight, 1999); and Richard J. Stiggins, *Student-Involvement Classroom Assessment*, 3rd ed. (Upper Saddle River, N.J.: Merrill/Prentice-Hall, 2001), pp. 409-65.
7. Gary Natriello, "The Impact of Evaluation Processes on Students," *Educational Psychologist*, vol. 22, 1987, pp. 155-75.
8. Wright, op. cit.
9. Susan M. Brookhart, "Teaching About Communicating Assessment Results and Grading," *Educational Measurement: Issues and Practice*, vol. 18, 1999, pp. 5-13.
10. Thomas R. Guskey and Jane M. Bailey, *Developing Grading and Reporting Systems for Student Learning* (Thousand Oaks, Calif.: Corwin Press, 2001).
11. Canady and Hotchkiss, op. cit.; Richard J. Stiggins and Daniel L. Duke, "District Grading Policies and Their Potential Impact on At-Risk Students," paper presented at the annual meeting of the American Educational Research Association, Chicago, 1991.
12. Cathy Costello and Barry McKillop, "Dealing with Lates and Absences," *Orbit*, vol. 30, no. 4, 2000, pp. 43-46.

The Communication Challenge Of Standards-Based Reporting

As traditional reporting systems based on letter grades are replaced by standards-based reporting systems, parents are often left wondering how their child is doing in school. Mr. Guskey offers some suggestions for overcoming this communication challenge.

BY THOMAS R. GUSKEY

A CLASSIC comic from the “Hi & Lois” strip shows their son arriving home from school and proudly announcing, “My teacher gave me a ‘Super’ on my report.” “Wow!” exclaims Lois. “Is that the best you can get?” “No,” he replies. “‘Stupendous,’ ‘Outrageous,’ and ‘Magnificent’ are all better. ‘Super’ is just okay.”

Like all good humor, this comic strip strikes a familiar note with many readers, especially the parents of school-age children. It also highlights one of the greatest challenges educators face today: describing students’ level of academic performance in meaningful ways to parents and others.¹

Moving away from traditional reporting systems based on letter grades and toward standards-based reporting systems means that we must articulate clearly what we expect students to learn and be able to do. That *curriculum* challenge is generally met through the development of specific content and performance standards. While meeting this challenge has been difficult and the quality of the work wide-ranging, most states and school districts today have curricula that are based on standards. The *communication* challenge of issuing progress reports and report cards that describe students’ performance with regard to those standards,

however, remains before us. It’s also proving to be a more difficult challenge than most educators ever anticipated.

STANDARDS-BASED REPORT CARDS

Developing a standards-based report card is a multi-step process. First, the major learning goals or standards must be identified, and the specific performance cri-



THOMAS R. GUSKEY is a professor in the College of Education, University of Kentucky, Lexington. © 2004, Thomas R. Guskey.

teria for demonstrating mastery of those goals or standards have to be set. Next, graduated levels of performance — or benchmarks — for achieving each goal or standard must be established. This effort typically requires determining three or four identifiable steps in students' progress toward mastery of each standard. In addition, meaningful labels need to be attached to these levels or steps in order to describe students' progress to their parents, to other interested parties, and to the students themselves. This is where the communication challenge gets particularly tricky.

To discover what terminology educators currently use to convey different levels of progress in student learning, I recently collected the labels from standards-based report cards obtained from a non-random sample of school districts throughout the U.S. and Canada. I also gathered the labels used to denote different levels of student performance in a number of state assessment programs and several well-known standardized assessment programs. Two colleagues and I then grouped these labels into general categories based on our judgments of what aspects of performance they were intended to describe. (See Table 1.) While most of these judgments were easy enough, deciding whether a label pertained to a level of "Understanding/Quality" or a level of "Mastery/Proficiency" proved particularly troublesome and remains open to discussion.

Next we shared these labels with parents of school-age children in structured focus groups. We asked the parents to identify which labels made sense and which ones did not. Their responses were amazingly consistent, highly informative, and, in some cases, quite surprising.

PARENTS' INTERPRETATIONS

We found that parents generally interpreted the labels according to their personal experiences with grading and reporting. And since parents' experiences with grades tend to be restricted to letter grades, most parents immediately translated each label into a letter grade. So, for example, "Advanced" means "A," "Proficient" means "B," and so on. Regardless of the labels actually used, the meaning parents took away from them was based on what they be-

lieved they understood best, and, for most parents, that was letter grades.

By and large parents also interpreted the labels from a norm-referenced perspective. Again, probably as a result of their personal experiences in schools where grades were based on each student's relative standing among classmates, parents interpreted the labels similarly. So for many parents, "Basic" and "Intermediate" imply "average" or "in the middle of the class."

After explaining to parents that these labels were designed to communicate a student's learning progress with regard to specific learning goals or standards, rather than to designate a student's standing among classmates, we asked parents to identify the labels that seemed clearer or more meaningful. Most of the labels received mixed responses, with no particular set being clearly preferred. However, certain labels were singled out by parents as confusing or meaningless.

Parents were especially baffled by the labels "Pre-Emergent" and "Emerging." Several remarked jokingly that "Emerging" conveyed images of "a slimy creature coming out of a swamp." When we indicated that "Emerging" generally implies "Beginning," they responded, "If you mean 'Beginning,' why not just say 'Beginning'?"

Another label parents found puzzling was "Exceeds Standard." Labels such as "Advanced," "Exemplary," "Distinguished," and "Outstanding" all seemed to have

TABLE 1.
Indicators of Student Performance

1. Levels of Understanding/Quality

Modest	Beginning	Novice	Unsatisfactory
Intermediate	Progressing	Apprentice	Needs Improvement
Proficient	Adequate	Proficient	Satisfactory
Superior	Exemplary	Distinguished	Outstanding

2. Levels of Mastery/Proficiency

Below Basic	Below Standard	Pre-Emergent	Incomplete
Basic	Approaching Standard	Emerging	Limited
Proficient	Meets Standard	Acquiring	Partial
Advanced	Exceeds Standard	Extending	Thorough

3. Frequency of Display

Rarely	Never
Occasionally	Seldom
Frequently	Usually
Consistently	Always

4. Degree of Effectiveness

Ineffective	Poor
Moderately Effective	Acceptable
Highly Effective	Excellent

5. Evidence of Accomplishment

Little or No Evidence
Partial Evidence
Sufficient Evidence
Extensive Evidence

clearer meaning. Parents understood how specific expectations or criteria might be associated with these levels. But to many parents, "Exceeds Standard" was especially vague and imprecise. Several interpreted it as meaning something "more than what's expected," but they were unsure just what that might be.

MEETING THE COMMUNICATION CHALLENGE

To improve the usefulness and communicative value of standards-based report cards, we need to ensure that parents and others understand the information they include. We must also acknowledge that, if parents don't understand the information in the report card, it's not their fault. As communicators, it is our responsibility to make sure that our message is clear and comprehensible to those for whom it is intended. This is the essence of the communication challenge involved in developing a standards-based report card.

Therefore, in describing different levels of students' performance with regard to learning goals or standards, we must choose labels that are expressive, precise, and meaningful. The following four guidelines should help in that effort.

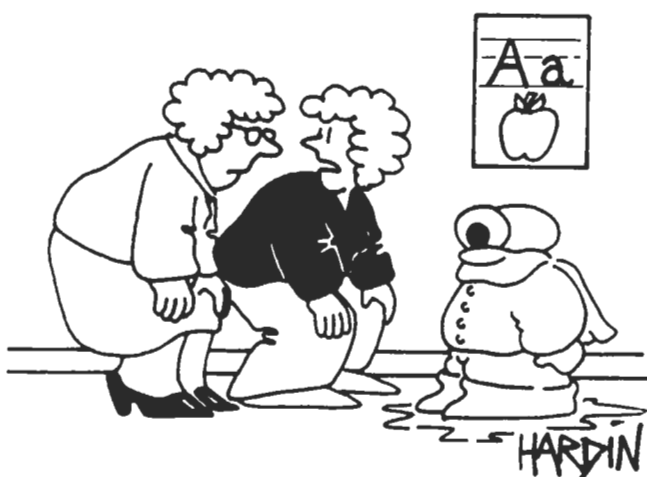
1. *Avoid comparative language.* Because parents so often interpret grades in terms of norm-referenced comparisons, in which a child's performance is judged relative to that of his or her classmates, adjusting to a standards-based, criterion-referenced system is particularly difficult. The transition is made all the more frustrating when educators use such comparative labels as "Below Average," "Average," and "Superior." The labels we use should always relate to clearly stated performance

indicators that communicate where students stand in reference to specific expectations for their learning. This helps parents change their perspective from "How is my child doing compared to other students in the class?" to "How is my child doing with regard to the learning expectations for this level?"

2. *Provide examples based on student work.* One of the best ways to promote understanding and to facilitate parents' transition from norm-referenced comparisons to standards-based reporting is to provide clear examples of student work at the various performance levels. Such examples enhance parents' knowledge of teachers' expectations. They also allow parents to become more discerning judges of their child's performance and then to better assist their child in making progress. This requires that school leaders provide time for teachers to engage in conversations about what is meant by "Proficient" and what examples of "Proficient" student work look like.

3. *Distinguish between "Levels of Understanding" and "Frequency of Display."* Parents get confused when educators use indicators that confound *what* students are able to do with *how often* they do it. The first implies "quality" to parents, while the second appears to signify "quantity" or "rate of occurrence." While "Frequency of Display" labels such as "Occasionally," "Frequently," and "Consistently" work well when describing students' work habits, study skills, or behavior in school, they often fall short when trying to explain to parents what students have learned and are able to do.

4. *Be consistent.* One reason so many parents translate labels into letter grades is that it provides a common basis for understanding and interpretation. This is particularly true in schools where one set of labels is used on the elementary report card, another set on the secondary report card, another set for state assessment results, and still another set for standardized assessment reports. No wonder parents who face this mish-mash ask, "Are 'Adequate' and 'Satisfactory' the same as 'Proficient'? Are they all equivalent to a 'B'?" Achieving consistency may prove difficult in schools bound to the use of labels incorporated in their state's assessment system. Still, by reducing the number of labels with which parents must contend, educators can facilitate parents' understanding and encourage greater parent involvement in education.



"It's Johnny Taylor's snowsuit, but I don't think he's in there."

Our knowledge of effective grading and reporting has grown tremendously in recent years, although little of that knowledge seems to be finding its way in-

to practice.² One theme that has emerged from this fund of new knowledge is that grading and reporting are less exercises in quantifying achievement than they are challenges in effective communication.³ Deciding what labels to use in describing students' level of performance with regard to standards is an essential first step in meeting that communication challenge.

Labels must be chosen to convey honest, meaningful, and useful information to parents and others in order to facilitate their understanding of educators' expectations for student learning. When parents and others recognize the intent of a standards-based report card and can make sense of the information it includes, they are better able to work with educators as partners in school improvement.⁴ Perhaps most important, a standards-based report card that uses clear and understand-

able labels helps break down the barriers between home and school and provides a basis for effective collaboration in efforts to help every student learn well.

1. Thomas R. Guskey, "Helping Standards Make the Grade," *Educational Leadership*, September 2001, pp. 20-27.

2. See, for example, Susan M. Brookhart, *Grading* (Upper Saddle River, N.J.: Merrill/Prentice-Hall, 2004); Thomas R. Guskey and Jane M. Bailey, *Developing Grading and Reporting Systems for Student Learning* (Thousand Oaks, Calif.: Corwin Press, 2001); Thomas H. Haladyna, *A Complete Guide to Student Grading* (Boston: Allyn & Bacon, 1999); and Robert J. Marzano, *Transforming Classroom Grading* (Alexandria, Va.: Association for Supervision and Curriculum Development, 2000).

3. Ken O'Connor, *How to Grade for Learning: Linking Grades to Standards*, 2nd ed. (Arlington Heights, Ill.: SkyLight, 2002); and Richard J. Stiggins, *Student-Involvement Classroom Assessment*, 3rd ed. (Upper Saddle River, N.J.: Merrill/Prentice-Hall, 2001), pp. 409-65.

4. Thomas R. Guskey, *How's My Kid Doing? A Parents' Guide to Grades, Marks, and Report Cards* (San Francisco: Jossey-Bass, 2002). **K**

THOUGHTS ON TEACHING

The Theory of Everything and Yada, Yada, Yada

THE FIRST thing I do every morning is check for education news. This morning my eyes were drawn immediately to the headline that read, "Trio Wins Prize for 'Theory of Everything.'" "Wow," I thought

BOBBY ANN STARNES writes and speaks on education issues. She lives in Loachapoka, Ala.

as I clicked on the link. "The theory of everything . . . that's a lot of stuff." I waited with anticipation as my dial-up connection loaded the story. Perhaps I would finally be able to understand some of the things that have puzzled me for a lifetime — like why we have a national holiday to celebrate Columbus' miscalculated route to Asia. Or how it is that, generation after generation, millions of people successfully conspire to perpetuate the Santa Claus hoax.

When the page finally loaded, a quick scan of the article left me disappointed. The trio's theory doesn't answer my questions. It explains things like asymptotic freedom, quarks, coins spinning on a table, and the way subatomic particles behave with gravity — all things I admit I've never spent a second wondering about.

Still, I really do admire their theory — or think I would if I understood anything

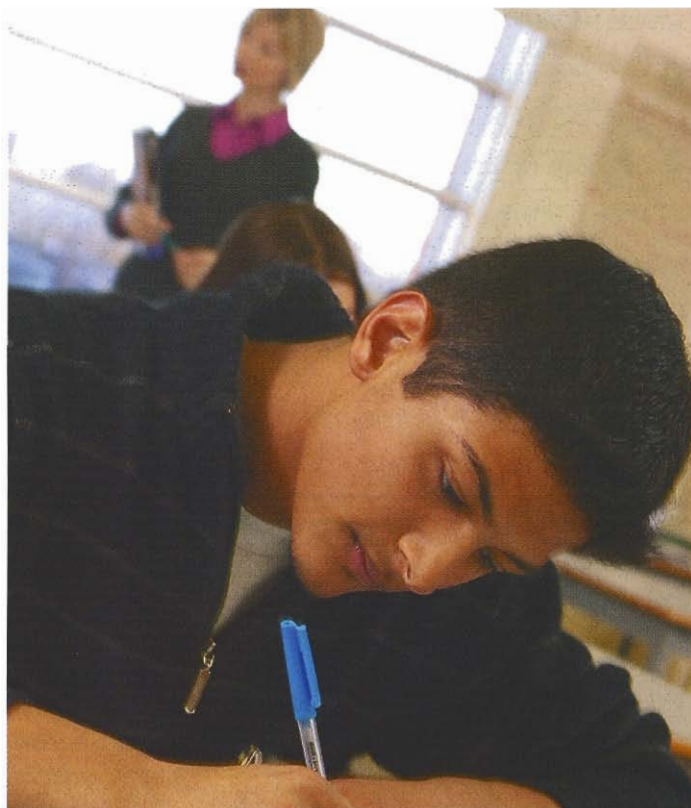
at all about it. Turns out the trio, David Gross, David Politzer, and Frank Wilczek, won the Nobel Prize for physics with an "outlandish" idea they began to explore more than 30 years ago. As a result of their tenacity, Finnish physicist Stig-Erik Starck reports that the scientists have "built a model of how the universe was born, how it works, and how it will ultimately die." These guys didn't just think outside the box. They invented a whole new box and *then* thought outside of it. One of the winners said his wife was putting the champagne on ice. I think most of us would agree that a Nobel Prize is champagne-worthy. But even more than that, it must be vindicating. It might even be worth going to a high school reunion. I bet the guys who ridiculed their idea as outlandish back in the Seventies are feeling pretty silly now.

I have a new theory, too. It isn't as big as theirs — but what *could* be? I think it would

Standards-Based Grading and Reporting

A Model for Special Education

Lee Ann Jung • Thomas R. Guskey



One of the most important functions of report cards and grades is to give families information on their children's progress in school. Families need to know their children's strengths and deficiencies, and interventions that can be undertaken at home to promote success. Recognizing the need for meaningful progress reporting, many schools have begun implementing "standards-based" grading and reporting practices (Guskey, 2001). Rather than reduce information on student learning to a single letter grade for each subject, standards-based grading allows teachers to report information on individual elements of learning. This level of detail is especially important to families of children with disabilities, for whom pivotal placement and intervention decisions

hinge on this information. The Individuals With Disabilities Education Act (IDEA) of 1997 and 2004 acknowledges this crucial need and requires that individualized education program (IEP) teams plan and document how progress will be monitored and communicated for students with disabilities (20 U.S.C. § 1414(d) (1) (A)). Despite this legal provision and widespread agreement on its importance, evidence indicates less compliance with progress monitoring than with any other IEP component (Etscheidt, 2006).

Challenges to Grading Students in Special Education

In recent years a marked increase has occurred in both the number of students with disabilities included in general

education classes as well as the amount of time they spend there (Handler, 2003). Although a wealth of research indicates the positive effects of including students with disabilities in general education classrooms (e.g., Baker, Wang, & Walberg, 1995; Carlberg & Kavale, 1980; Hunt, Farron-Davis, Beckstead, Curtis, & Goetz, 1994; Waldron, 1998), the process poses significant challenges to grading and reporting on the performance of students included in those general education classes. Is it best to report achievement on grade-level standards, for example, or should grades be adapted? Should the grades be based on achievement only, or on effort, progress, or some combination of all three? For students with disabilities who receive much of their education outside the general education classroom, the special education teacher typically assigns most grades, whereas the general education teachers determine grades for the few subject areas in which students are fully included. For students with disabilities who are fully included in the general education classroom, however, the division of grading responsibilities is less clear (Bursuck et al., 1996; Polloway et al., 1994).

A common strategy for grading students who are included involves the general education teacher's taking

responsibility for all areas on the regular report card and the special education teacher's taking responsibility for reporting on progress toward IEP goals. Although this approach seems logical, deciding the appropriate grade for a general education content area can be very difficult, particularly if performance in the content area is affected by the disability.

Take, for example, an eighth-grade student who is unable to demonstrate proficiency on the eighth-grade standards because of multiple, severe disabilities but has worked hard and progressed well toward IEP goals. On one hand, to fail such a student who has shown tremendous effort and progress clearly seems unfair. But on the other hand, giving passing marks to a student who has not yet met prescribed performance standards for that grade level also seems inappropriate. Further complicating this matter are the legal requirements of grading students with disabilities. Most notably, IEPs must "enable the child to achieve passing marks and advance from grade to grade" (*Board of Education v. Rowley*, 1982). Therefore, a failing grade for a student receiving special education services is considered an indicator that appropriate educational services were not provided.

Grading Adaptations

Although increasing numbers of students with disabilities are included in general education classrooms for greater portions of the day (Handler, 2003), little guidance or direction has come from the field of special education to help address the challenge of grading students in inclusive settings. Lacking explicit recommendations on grading, most general classroom teachers make individual, informal grading adaptations for such students (Polloway et al., 1994). To aid teachers in this adaptation process and to promote consistency, a variety of grading adaptations have been recommended over the years. Grading adaptations are procedures for individualizing a grading system for a student with disabilities (Silva, Munk, & Bursuck, 2005). Such adaptations generally fit within five categories: (a) con-

sidering progress on IEP goals; (b) measuring improvement over past performance; (c) prioritizing assignments or content differently; (d) including

Most general classroom teachers make individual, informal grading adaptations for such students.

indicators of behavior or effort in the grade; and (e) modifying the weights or scales for grading (Silva et al.).

For example, a student with a disability, if judged the same way as class peers, may have demonstrated C-level proficiency in social studies for the grading period. The teacher could implement a grading adaptation by giving extra points if the student surpassed IEP goals or exerted high effort. Theoretically, such adaptation provides encouragement and opportunities for success to students for whom grade-level standards may not be attainable. In reality, however, such adapted grades can lead such students to believe that their grades are not the result of what they do but who they are. This perception, in turn, may actually decrease their motivation (Ring & Reetz, 2000). Such grading adaptations also introduce issues of unfairness (Bursuck, Munk, & Olson, 1999). And furthermore, even with such adaptations, most students in special education continue to receive low passing grades, placing them at high risk for low self-esteem and dropping out of school (Donahue & Zigmond, 1990).

Implications of Standards-Based Grading

The shift to standards-based grading and reporting has further complicated grading students with disabilities who are included in general education classrooms. Although grading all students in special education on the basis of grade-level standards is inappropriate, most of the practices recommended to date are not well suited to a standards-based grading system. When teachers must

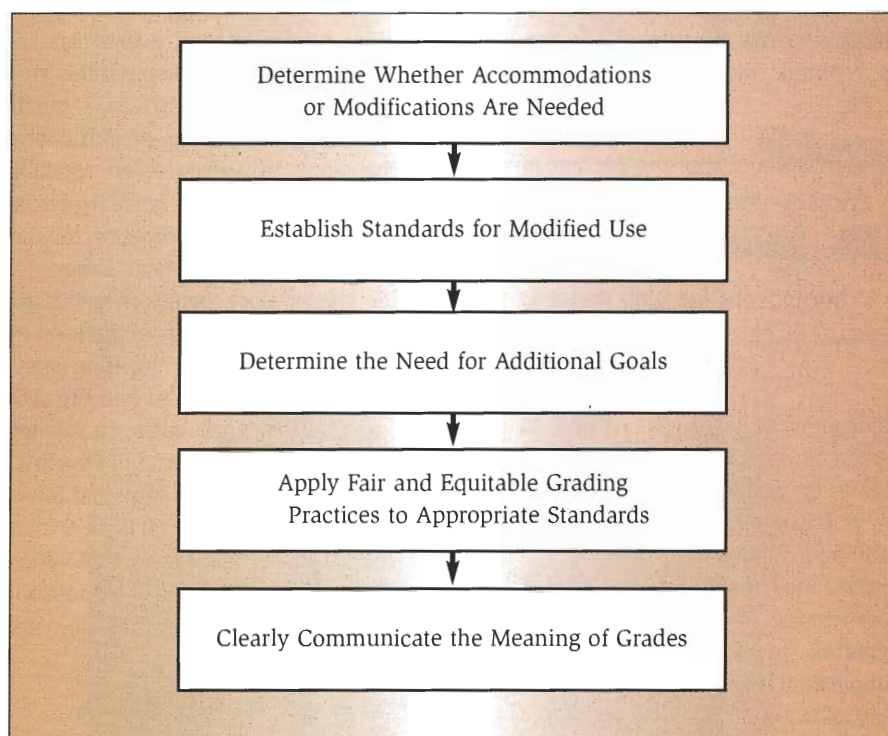
base their grades on specific learning standards, the meaning of the grade changes from a general overall assessment of learning (e.g., How did this student perform in science?) to a much more detailed description of a student's performance on a discrete set of skills (e.g., How well did the student master the ability to classify minerals on the basis of multiple physical criteria?). When the primary question addressed in assigning a grade shifts to the level of mastery of a particular learning standard, teachers are likely to find the task of grading students with disabilities much more troublesome (Thurlow, 2002). To provide meaningful and interpretable indicators of achievement that are useful for making accurate decisions about students in special education, more effective grading practices are sorely needed.

Setting a Solid Foundation

Before considering grading methods specific to students in special education, schools must have a high-quality grading and reporting system in place for all students. Thoughtful and well-reasoned grading policies can address many of the problems schools face with special education grading. One fundamental component of a high-quality grading and reporting system requires teachers to consider three distinct types of learning criteria:

- **Product criteria** relate to students' specific achievements or level of proficiency and are based on culminating demonstrations of learning, such as examinations, final reports, projects, or portfolios, and overall assessments of learning.
- **Process criteria** relate to students' effort, class behavior, or work habits. They also might include evidence from daily work, regular classroom quizzes, homework, class participation, or punctuality of assignments.
- **Progress criteria** relate to how much students gain from their learning experiences. Teachers who use progress criteria typically look at how far students have come rather than where students are (Guskey, 1996, 2006; Guskey & Jung, 2006).

Figure 1. Inclusive Grading Model



Most teachers base their grading on some combination of these three types of criteria (Brookhart, 1993; Frary, Cross, & Weber, 1993). The majority of teachers also vary the criteria they employ from student to student, taking into account individual circumstances (Truog & Friedman, 1996). Although teachers do so in an effort to be fair, the result is a “hodgepodge” grade (Brookhart, 1991; Cizek, Fitzgerald, & Rachor, 1996; McMillan, Myran, & Workman, 2002) that is difficult for parents to interpret (Friedman & Frisbie, 1995). An A, for example, may mean that the student knew what the teacher expected before instruction began (product), did not learn as well as expected but tried very hard (process), or made significant improvement (progress).

High-quality grading and reporting systems establish clear indicators of product, process, and progress criteria and then report each separately (Guskey, 1994; Stiggins, 2001; Wiggins, 1996). In other words, teachers separate grades or marks for achievement from those for homework, effort, work habits, or learning progress. Schools that have implemented such a system find it actually makes grading easier. No

Parents generally prefer this approach because it gives them more detailed and prescriptive information about their children’s learning.

more information needs to be gathered, and teachers can avoid debates about how best to combine diverse types of evidence into a single grade. Teachers also report that students take homework, effort, and other work habits more seriously when they are reported separately (Guskey, 2006). Parents generally prefer this approach because it gives them more detailed and prescriptive information about their children’s learning. For students in special education, it means that families not only receive specific feedback on their children’s achievement but also essential information on progress and effort that can be crucial to making intervention and placement decisions.

Inclusive Grading Model

Once a school has in place a high-quality grading and reporting system that separates product, process, and progress learning goals, educators can develop appropriate policies and practices for grading students with disabilities who are included in a standards-based learning environment. The 5-step Inclusive Grading Model presented in Figure 1 is designed to fit a standards-based grading and reporting system and meet legal requirements for reporting progress of students who have IEPs. The 5 steps of the model consist of the following:

1. Determine whether an accommodation or a modification is needed for each grade-level standard.
2. Establish the appropriate modified standard for each area requiring modification.
3. Outline any additional goals pertinent to the child’s academic success.
4. Apply equivalent grading practices to the appropriate standards.
5. Clearly communicate the grades’ meaning.

Let us consider each of these steps in detail.

Step 1: Determine Whether Accommodations or Modifications Are Needed

Each student who qualifies for special education must have an IEP that outlines a specific plan of individualized annual goals, along with instructional strategies and adaptations needed for the student to reach those goals. Each student’s IEP team meets at least once per year to discuss progress and to update the IEP. For most students who qualify for special education, adaptations are needed to give them access to the general education curriculum. By explicitly connecting adaptation needs with the general curriculum standards, IEP teams can set the stage for meaningful grading and reporting. Considering each grade-level standard individually, teams should decide whether no adaptations, *accommodations*, or *modifications* are needed. Adaptations that provide access to the general curriculum but do not fundamentally alter

the grade-level standard are known as *accommodations* (Freedman, 2005). For example, a high-school student who has a learning disability in the area of written expression may require an audio-tape of science lectures due to difficulty in taking notes. Because of the learning disability, this student may also need to be administered exams orally. Although the format for answering questions on exams is different in this instance, the content of the questions and the substance of responses remains the same. Therefore, achievement on the grade-level standard in science is what should be reported.

Some students receiving special education need curricular adaptations that are more substantial than accommodations. For those students, some or all of the grade-level standards may not be achievable during the academic year, and curricular *modifications* are needed. A modification is an adaptation to the curriculum that fundamentally alters the grade-level expectation (Freedman, 2005). For example, an IEP team may determine that a fourth-grade student who has a severe mathematics learning disability will not be able to achieve the fourth-grade mathematics standards that academic year. For this student, the mathematics curriculum will need to be modified to provide opportunities with mathematics content that are appropriate for the student's present level of development. These modifications would then be noted in the IEP.

Step 2: Establish Standards for Modified Areas

For the fourth-grade student in the foregoing example, communicating failure on the grade-level mathematics standards provides no meaningful information about that student's achievement or progress. Instead, the IEP team must determine a modified standard that this child will be able to achieve with appropriate special education services. Modified standards should be clearly linked with the grade-level standard and recorded on the IEP as an annual goal with short-term objectives. A child with mental retardation, for example, may not be ready to work on third-grade sci-

ence standards in mineral identification. The IEP team may choose to develop science standards on the skill of sorting and classifying that are fundamentally related to the third-grade science standards but also developmentally appropriate for the student. For areas requiring these types of modification, achievement on the modified standards is what should be graded and reported.

Step 3: Determine the Need for Additional Goals

For some students receiving special education, additional IEP goals may be pertinent to the student's development but extend beyond the general curriculum. A student with visual impairment, for example, may have orientation and mobility goals as a part of the IEP. For this student, being able to walk independently from the classroom to the lunchroom, to outside, and so forth, is important to being a part of the class. Although this goal may not be included within the structure of the regular report card, monitoring and reporting on this goal are important. Schools should continue to provide this information on a regular basis through a report card supplement so that families and others on the team are able to make decisions based on the child's progress and achievement (National Center on Secondary Education and Transition, 2005).

Step 4: Apply Fair and Equitable Grading Practices to Appropriate Standards

Once schools have a high-quality grading and reporting system in place that makes the purpose of grading clear and offers guidance on how to grade, IEP teams can apply grading practices appropriate for students with disabilities. For most students, including those in special education, the standards being measured are grade-level standards. In subject areas in which only accommodations are needed, students receiving special education should receive grades according to the same criteria as every other student in the class, with no penalty for accommodation unless otherwise noted on the IEP. A student who takes a history test orally, for example, should be graded on the basis of the content of his or her

responses. The grade should not be lowered because of the response format. However, it also should not be raised on the basis of effort, progress, or any other factor that is not a part of every other student's achievement grade.

Modified standards should be clearly linked with the grade-level standard and recorded on the IEP as an annual goal with short-term objectives.

For subject areas in which modified standards are used, grades should be based on the modified standard, not the grade-level standard. From the example above, the student who has mental retardation and is working toward a lower level sorting and classifying science standard should be assigned a grade based on that modified standard. Measuring and reporting progress on a standard the IEP team has already agreed to be unattainable would be meaningless and, arguably, illegal.

Step 5: Communicate the Meaning of the Grades

By providing information on students' specific achievements, separate from indicators of progress and effort, and then clearly communicating the meaning of each grade assigned, educators can offer families much better information about children's learning success. If some or all the grades for achievement are based on modified standards, then the reporting system must include additional information to ensure that families understand that their child's success is based on work appropriate for his or her development level, not the assigned grade level. Assigning grades on the basis of modified standards without communicating what was truly measured is no more meaningful or fair than giving failing grades on the basis of grade-level standards. Each grading period, schools might include on the report card a column in which special notations can be marked. Or a superscript letter or an asterisk could simply be

added to the grade or mark to indicate that it is based on modified standards. The accompanying footnote might then state, "Based on modified standards" and direct the reader to the standards on which the grade was based.

By law, however, the notation on the report card or transcript cannot, in any way, identify the student as receiving special education services. For example, the wording "modified standard" is a legal notation if modifications are available to all students, but "special education goals" and "IEP goals" are not. An accompanying report might include the student's IEP goals or a narrative describing the details of the IEP.

Conclusion

Educators at all levels desperately need clear and specific guidance in developing grading and reporting policies and practices for students with disabilities who are included in general education classes. They also need concise and meaningful data on the effects of such policies and practices. Although some grading adaptations have been studied in terms of their perceived fairness to teachers and students, additional evidence is needed to determine the effectiveness of various adaptations for grading the performance of students with special needs. For example, the following questions need to be addressed: Do families understand their children's progress? Can IEP teams use grades to make data-based decisions on the efficacy of interventions? Can schools use the information to determine whether a child has made adequate progress to advance to the next grade?

*Students with disabilities and
their families can have information
that they are able to interpret
accurately and use effectively.*

Separating product, process, and progress learning goals, and then situating achievement grades within the context of accommodations and modifications, offers a promising alternative to modified grading within a standards-

based environment. The IEP serves to document curricular accommodations and modifications for students who receive special education. After considering the accommodation and modification needs of students, IEP teams can determine for each content area whether students are to be held to grade-level standards or modified standards. If the team modifies particular standards they judge to be inappropriate for the student, then no further grading adaptations are needed. Achievement or product grades need not be adjusted by considering progress, effort, work habits, or other behaviors. Process and progress indicators remain an important part of grading and reporting but are kept separate from indicators of students' achievement of specific learning standards. By reporting product, progress, and process goals separately, educators can eliminate inaccurate grades based on an arbitrary mix of grading elements or on inappropriate standards. As a result, students with disabilities and their families can have information that they are able to interpret accurately and use effectively.

References

- Baker, E. T., Wang, M. C., & Walberg, H. J. (1995). The effects of inclusion on learning. *Educational Leadership*, 52(4), 33-35.
- Board of Education v. Rowley, 458 U.S. 176 (1982).
- Brookhart, S. M. (1991). Grading practices and validity. *Educational Measurement: Issues and Practice*, 10(1), 35-36.
- Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement*, 30(2), 123-142.
- Bursuck, W., Polloway, E. A., Plante, L., Epstein, M. H., Jayanthi, M., & McConeghy, J. (1996). Report card grading and adaptations: A national survey of classroom practices. *Exceptional Children*, 62, 301-318.
- Bursuck, W. D., Munk, D. D., & Olson, M. M. (1999). The fairness of report card grading adaptations: What do students with and without disabilities think? *Remedial and Special Education*, 20, 84-92.
- Carlberg, C., & Kavale, K. A. (1980). The efficacy of special versus regular class placement for exceptional children: A meta-analysis. *Journal of Special Education*, 14, 296-309.
- Cizek, G. J., Fitzgerald, S. M., & Rachor, R. E. (1996). Teachers' assessment practices: Preparation, isolation, and the kitchen sink. *Educational Assessment*, 3(2), 159-179.
- Donahue, K., & Zigmond, N. (1990). Academic grades of ninth-grade urban learning disabled students and low-achieving peers. *Exceptionality*, 1(1), 17-27.
- Etscheidt, S. K. (2006). Progress monitoring: Legal issues and recommendations for IEP teams. *TEACHING Exceptional Children*, 38(3), 56-60.
- Frary, R. B., Cross, L. H., & Weber, L. J. (1993). Testing and grading practices and opinions of secondary teachers of academic subjects: Implications for instruction in measurement. *Educational Measurement: Issues and Practice*, 12(3), 23-30.
- Freedman, M. K. (2005). *Student testing and the law: The requirements educators, parents, and officials should know*. Horsham, PA: LRP Publications.
- Friedman, S. J., & Frisbie, D. A. (1995). The influence of report cards on the validity of grades reported to parents. *Educational and Psychological Measurement*, 55(1), 5-26.
- Guskey, T. R. (1994). Making the grade: What benefits students. *Educational Leadership*, 52(2), 14-20.
- Guskey, T. R. (1996). Reporting on student learning: Lessons from the past—Prescriptions for the future. In T. R. Guskey (Ed.), *Communicating student learning: 1996 Yearbook of the Association for Supervision and Curriculum Development* (pp. 13-24). Alexandria, VA: Association for Supervision and Curriculum Development.
- Guskey, T. R. (2001). Helping standards make the grade. *Educational Leadership*, 59(1), 20-27.
- Guskey, T. R. (2006). Making high school grades meaningful. *Phi Delta Kappan*, 87(9), 670-675.
- Guskey, T. R., & Jung, L. A. (2006, Winter). The challenges of standards-based grading. *Leadership Compass*, 4(2), 6-10.
- Handler, B. R. (2003, April). *Special education practices: An evaluation of educational environmental placement trends since the regular education initiative*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Hunt, P., Farron-Davis, F., Beckstead, S., Curtis, D., & Goetz, L. (1994). Evaluating the effects of placement of students with severe disabilities in general education versus special classes. *Journal of the Association for Persons with Severe Handicaps*, 19, 200-214.
- Individuals With Disabilities Education Act, 20 U.S.C § 1400 to 1491 (1997).
- Individuals With Disabilities Education Improvement Act, 20 U.S.C § 1400 to 1482 (2004).

McMillan, J. H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *Journal of Educational Research*, 95(4), 203-213.

National Center on Secondary Education and Transition. (June, 2005). Key provisions on transition: IDEA 1997 compared to H.R. 1350 (IDEA 2004). Retrieved from <http://ncset.org/publications/related/ideatransition.asp>

Polloway, E. A., Epstein, M. H., Bursuck, W. D., Roderique, T. W., McConeghy, J. L., & Jayanthi, M. (1994). Classroom grading: A national survey of policies. *Remedial and Special Education*, 15, 162-170.

Ring, M. M., & Reetz, L. (2000). Modification effects on attribution of middle school students with learning disabilities. *Learning Disabilities Research & Practice*, 15, 34-42.

Silva, M., Munk, D. D., & Bursuck, W. D. (2005). Grading adaptations for students with disabilities. *Intervention in School and Clinic*, 41, 87-98.

Stiggins, R. J. (2001). Report cards. In *Student-involved classroom assessment* (3rd ed.; pp. 409-465). Upper Saddle River, NJ: Merrill/Prentice Hall.

Thurlow, M. L. (2002). Positive educational results for all students: The promise of standards-based reform. *Remedial and Special Education*, 23, 195-202.

Truog, A. L., & Friedman, S. J. (April, 1996). *Evaluating high school teachers' written grading policies from a measurement perspective*. Paper presented at the annual meeting of the National Council on Measurement in Education, New York.

Waldron, N. L. (1998). The effects of an inclusive school program on students with mild and severe learning disabilities. *Exceptional Children*, 64, 395-405.

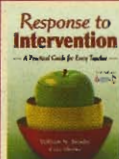
Wiggins, G. (1996). Honesty and fairness: Toward better grading and reporting. In T. R. Guskey (Ed.), *Communicating student learning: 1996 Yearbook of the Association for Supervision and Curriculum Development* (pp. 141-176). Alexandria, VA: Association for Supervision and Curriculum Development.

Lee Ann Jung (CEC KY Federation), Assistant Professor, Special Education and Rehabilitation Counseling; and **Thomas R. Guskey**, Professor, Educational Policy Studies and Evaluation, University of Kentucky, Lexington.

Address correspondence to Lee Ann Jung, 124 Taylor Education Building, University of Kentucky, Lexington, KY 40506 (e-mail: ljung@uky.edu).

TEACHING Exceptional Children, Vol. 40, No. 2, pp. 48-53.

Copyright 2007 CEC.



RTI

WORKSHOP

3 tiers 2 days
Washington, DC Metro Area

Council for Exceptional Children Regional Workshop

Response to Intervention: A Practical Guide for Every Teacher

1 BEST SELLING PUBLICATION
at CEC'S 2007 Annual Convention
RANKED #3 TOP ATTENDED WORKSHOP
at CEC'S 2007 Annual Convention

For educators who are new to the Response to Intervention (RTI) approach, this workshop offers an overview of key concepts and guidelines of RTI that benefit students in inclusive classrooms. Cara Shores, co-author of *Response to Intervention: A Practical Guide for Every Teacher*, demonstrates how general and special education teachers can use research-based intervention to individualize instruction, monitor student progress, and implement strategies to meet students' specific needs.

A true working workshop, the RTI 2-day, provides practical information designed to enable teachers and administrators to implement the RTI process effectively in their districts, schools, and classrooms. Be prepared to play an active role! Participants will have the opportunity to apply their knowledge as they work in teams to develop an RTI plan for a student from their school.

Earn 1.2 Continuing Education Units for each 2-day workshop!* **Presenter and Co-Author: Cara Shores

LEARNER OUTCOMES

Participants will be able to:

- Identify the major components of two models of RTI
- Identify resources for selection of research-based strategies and curriculum-based measurement
- Utilize the RTI process to identify learning problems based on student outcome data.

Washington, DC Metro Area

November 15-16

Sheraton Reston
11810 Sunrise Valley Drive
Reston, VA 20191
8:00 AM-4:00 PM

Registration \$399
CEC Member \$299
Hotel Room \$179 per day
703-620-9000



CEC extends its thanks to Corwin Press for supporting the 2007 RTI Regional Workshops.



Register now! Call 1-888-232-7733 or online at www.cec.sped.org

Are Zeros Your Ultimate Weapon?

By Thomas R. Guskey From *Principal Leadership*

GRADING is one of a teacher's greatest challenges and most important professional responsibilities. However, few teachers have any formal training in grading methods, and most teachers have limited knowledge about the effectiveness of various grading practices.

As a consequence, when teachers develop their grading policies, they typically reflect back on what they experienced as students and use strategies that they perceived to be fair, reasonable, and equitable. In other words, most teachers do what was done to them.

When teachers are asked why they grade or what purpose grading serves, their responses generally fall into these broad categories:

Thomas R. Guskey (guskey@uky.edu) is a Professor of Education, University of Kentucky, Lexington, KY 40506-0001. Condensed from *Principal Leadership*, 5 (October 2004), 49-53. Published by the National Association of Secondary School Principals (NASSP), 1904 Association Drive, Reston, VA 20191, from which related educational materials are available by contacting NASSP.

- To communicate the achievement status of students to parents and others. Grading and reporting provide parents and guardians with information about students' progress and allow them to be involved in the educational process.

- To provide information students can use for self-evaluation. Grading and reporting give students information about the adequacy of their academic performance.

- To select, identify, or group students for specific educational paths or programs. High grades are typically required for entry into advanced classes or honors programs; low grades are often the first indicator of learning problems that can result in a student's placement into a special needs program. In addition, grades are used as a criterion for admission to colleges and universities.

- To provide incentives for students to learn. Although some may debate the idea, there is extensive evidence that grades and other reporting methods are important factors in determining the amount of effort that students put forth and how seriously students regard a learning task.

- To evaluate the effectiveness of instructional programs. Grade distributions are often compared to judge the effectiveness of new programs or instructional techniques.

- To provide evidence of a student's lack of effort or inability to accept responsibility for inappropriate behavior. Grades and other reporting devices are frequently used to document unsuitable behavior by students, and some teachers threaten students with poor grades to encourage more acceptable behaviors.

Trying All, Achieving None

Although all of these may be legitimate, teachers seldom agree on which one is most important. As a result, teachers often attempt to address *all* these purposes with a single grading procedure or policy and usually end up achieving none very well.

Nearly all teachers do agree, however, that the least important is the sixth: *To provide evidence of students' lack of effort or inability to accept responsibility for inappropriate behavior.* But few teachers recognize that many of their grading practices serve precisely this purpose, most obviously when they assign zeros to students' work that is missed, neglected, or turned in late.

Many teachers see zeros as their ultimate grading weapon, using them to punish students for not making adequate effort or failing to show appropriate responsibility. Students get zeros for not meeting set deadlines, misbehaving in class, or refusing to heed the teacher's warnings. Some teachers recognize that assigning zeros punishes

Why Zeros Don't Work

students academically for behavioral infractions, but most believe such punishment is justified and deserved.

Teachers also use zeros as instruments of control. In most instances, teachers have little direct influence over students' most valued privileges or most feared punishments, unable to restrict their access to automobiles, computer games, television, or social activities. But they do control grades, which can indirectly influence these things. A low grade often prompts parents to enforce punishments more persuasive and compelling to students than those a teacher can enforce. The threat of a zero—and the resulting low grade—allows teachers to impose their will on students who otherwise might be indifferent to a teacher's demands.

Inaccurate and Unmotivating

The problems with assigning zeros, however, are numerous and significant. First, a zero is seldom an accurate reflection of what a student has learned or is able to do. Obviously, if the grade is to represent how well students have learned, mastered established learning standards, or achieved specified learning goals, then the practice of assigning zeros clearly misses the mark.

Second, the effect of assigning zeros is greatly magnified if combined with the common practice of averaging scores to attain students' overall course grades. Students readily see that receiving a single zero leaves them little chance for success or a high grade because such an extreme score drastically skews the average.

That is why, in scoring such Olympic events as gymnastics or diving, the highest and lowest scores from judges are always eliminated. If they were not, one judge could control the entire competition simply by giving extreme scores. A single zero has more influence on an average than any other score in the group.

Third, and perhaps most important, no studies support the use of zeros or low grades as effective punishments. Instead of prompting greater effort, zeros and the low grades they yield more often cause students to withdraw from learning. To protect their self-image, many regard their low mark or grade as irrelevant and meaningless. Other students may blame themselves for the low grade but often feel helpless to make improvements.

Frequently, teachers defend assigning zeros by arguing that they cannot give students credit for work that is incomplete or not turned in—and that is certainly true. But, considering these overwhelmingly negative effects, there are far better ways to motivate and encourage students to complete assignments in a timely manner than with zeros. ▶

Several schools have implemented the following alternatives and experienced great success as a result.

Assign "I" or "Incomplete" grades. One alternative to zeros is to assign an "I" or "Incomplete" grade with explicit requirements for completing the work. The consequence of receiving an "I" is usually required attendance at a special study session after school or a special Saturday class where students work to complete neglected assignments to a satisfactory level.

In other words, students are not let off the hook with a zero. Instead, they learn that they have specific responsibilities in school and that their actions have definite consequences. Not completing assigned work on time means that students must attend special after-school or Saturday sessions to complete the work, and no excuses are accepted. The consequence is direct, immediate, and academically sound.

Of course, such a policy requires additional funding for necessary support mechanisms. Teachers, volunteer parents, or older students must staff these after-school or Saturday sessions, which also require classroom space and possibly supplementary transportation. But schools with these policies generally save money in the long run.

When students realize their teachers are serious about school responsibilities, they also get serious about them. Because the policy's consequences and accompanying assistance are immediate, it helps students remedy learning or behavioral difficulties before they become major problems. As a result, less time and fewer resources will be needed for major remediation efforts in the future. Further, this policy is far more beneficial and fairer to students than simply assigning zeros because it makes a grade more accurately reflect what students have learned.

Report behavioral aspects separately. Another alternative to assigning zeros is to report behavioral aspects of students' performance separately. In many Canadian secondary schools, students receive multiple grades for each of their classes—both on the report card and on grade transcripts.

A main "achievement" grade is based on evidence of students' academic performance. This achievement grade might include results from major examinations, scores from compositions or reports, or portfolio or project assessments.

But on the basis of specific criteria, these teachers offer separate grades or marks for homework, punctuality of assignments, class participation, effort, and so on. These aspects of students' performance are typically labeled *learning skills*, *work habits*, or *academic behaviors*.

Reporting multiple grades on different aspects of student perfor-

mance may appear to create additional work for teachers. However, Canadian teachers who use this approach claim that it is easier and requires less work than assigning a single grade. These teachers gather the same evidence on student performance as other teachers. But by reporting multiple grades, they avoid the problems associated with combining many diverse sources of information into a single amalgamated grade. They are also spared from arguments about what "weight" to assign to each category or source of evidence.

Most important, the grades they assign are more meaningful. Calculations of Grade Point Average (GPA) and class rank, for example, more accurately reflect students' academic performance because they are based solely on "achievement" grades that are untainted by nonacademic, behavioral factors.

Change grading scales. One of the easiest and least objectionable ways to lessen the negative effects of zeros is to change grading scales. Schools using this approach shift from percentage grading scales where, for example, A = 90-100%, B = 80%-89%, etc., to whole number scales where A = 4, B = 3, C = 2, and so on.

In other words, although teachers can still assign zeros to student work that is missed, neglected, or turned in late, the effect of a zero is lessened because it is not so extreme. Although this approach ignores the problem of the grade not representing an accurate reflection of student learning, it does reduce the damage imposed by the extreme value of zero in a percentage grading system.

Teachers at all levels would undoubtedly prefer that students' motivation for learning be entirely intrinsic. Most recognize, however, that grades and other reporting methods are important factors in determining how much effort students put forth. Unfortunately, this recognition leads some teachers to use grades as weapons to punish students, even though the practice has no educational value and, in the long run, adversely affects students, teachers, and the relationship they share.

Developing honest and fair grading policies should begin with candid discussions about the purpose of grading and reporting. Teachers must consider what message they want to communicate through grading, who the primary audience for the message is, and what the intended goal of the communication is.

Once issues about purpose are resolved, decisions about the appropriateness of various grading policies—and the use of zeros—are much easier to address and resolve. If guided by reflections on the true purpose of grading, it is likely that teachers at all levels will abandon the use of zeros completely.

Inflation not the issue; focus on grades' purpose

By Thomas R. Guskey

Grade inflation deeply concerns many university officials today. Recent debates on grade inflation at the University of Kentucky and Eastern Kentucky University are typical of those raging at universities throughout the nation. Unfortunately, most of these debates focus on the wrong issues. As a result, the solutions proposed are misguided.

The problem with grade inflation is not simply that more students are receiving high grades. It is that we're not sure what those grades mean. Adding pluses and minuses to grades adds nothing to their meaning when the criteria by which grades are assigned remain ambiguous.

The question that needs to be addressed in these debates is, "What is the purpose of grading?" If, as some professors think, the purpose of grading is to discriminate among students, then we must maximize the differences between students in terms of their performance. Because it's difficult to distinguish among students if many do well, we must make the differences in their performance as great as possible. The best mechanism for maximizing these differences is poor teaching. When students are taught poorly, only those who are able to teach themselves learn well and receive high marks. The majority of students who need the help and assistance of their professors receive the low marks.

Maximizing differences among students is typically accompanied by "grading on the curve." This means that students are graded according to their relative standing among classmates. Grading on the curve makes it easy to adjust for grade differences between departments or classes by simply mandating that only a small percent of students, say the top 20 percent, receive the highest grade. But when students are graded on the curve, a high mark does not represent excellent performance, as some might think. It means only that the performance was somewhat better than that of others in the class, all of whom might have performed miserably.

Grading on the curve also makes learning highly competitive. Students must compete among themselves for the few scarce rewards (high grades) distributed by the professor. Under these conditions students avoid helping each other because doing so is detrimental to their chances for success. Getting a high grade doesn't mean performing excellently. It

means simply doing better than your classmates.

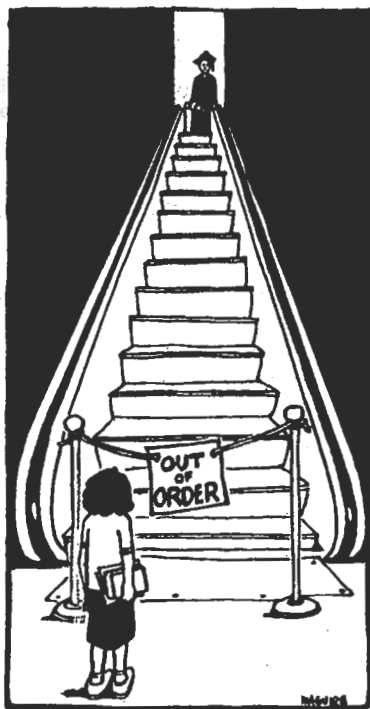
On the other hand, if the purpose of grading is to reflect how well students have learned, then we must follow different procedures. First, we must clarify what we want students to learn and be able to do. Second, we must identify clear criteria or standards by which their learning will be judged. That is, we must decide what evidence best represents what students have learned. Teaching then becomes an organized and purposeful effort designed to help all students meet those standards. The goal is to develop talent, not simply to select talent.

Grades that reflect well-defined learning standards have direct meaning. They describe what students have accomplished and the skills they have acquired. Grades based on learning standards also bring new significance to discussions of differences in grade distributions across departments and classes. Students' grades in some professors' classes may be higher because the standards are less rigorous. A comparison of related learning criteria would address this issue. It also may be, however, that some professors are simply better than others at helping their students meet rigorous and challenging learning standards. Such evidence would be invaluable in efforts to improve instructional quality.

The problem is that defining clear learning standards and deciding what evidence best reflects those standards is hard work. It takes lots of time, clear thinking, and dedicated effort. Professors don't always agree on what standards are appropriate or what evidence should be used to verify students' attainment of those standards. But isn't this precisely the kind of debate that should be going on in colleges and universities? And isn't this debate more likely to benefit students than one that revolves around how many categories to have in a grading system?

Adding more categories to the grading system provides only the illusion of accuracy in the absence of clearly specified learning criteria. It also does little to solve the real problems that underlie grade inflation. Until we precisely identify what students are expected to learn, articulate the criteria by which their learning will be judged and clearly communicate these criteria to students, grading will remain an arbitrary and highly subjective process that victimizes more students than it helps.

■
Thomas R. Guskey is a professor of education at the University of Kentucky.



BARRIE MAGUIRE