

Grades as Valid Measures of Academic Achievement of Classroom Learning

JAMES D. ALLEN

What is the purpose of grades? In this article I present one answer to this question from a perspective that many educators might see as somewhat radical or extreme. The perspective that I take is based on the fundamental educational psychology assessment principle of validity—the validity of what learning is being assessed and the validity of the communication of that assessment to others. I believe most teachers fail to give grades to students that are as valid as they should be. Because grading is something that has been done to each of us during our many years as students, it is hard to change the invalid “grading” schema that has become embedded in our minds. Now, as educators often required to grade students, and because of this embedded schema, we often grade students in invalid ways similar to how we were graded. Inadequate education in valid assessment and grading principles and practices is a reason many teachers continue to perpetuate invalid grading practices with students. Since educational testing and assessment is a major content knowledge area in educational psychology, the issues regarding assessment and grading that I address in this article could well be addressed in an educational psychology course. If our preservice and in-service teachers are going to learn appropriate assessment and grading practices then educational psychologists need to provide the relevant information in their classes.

The most fundamental measurement principle related to meaningful assessment and grading is the principle of validity (Gallagher 1998; Gredler 1999; Linn and Gronlund 2000; Stiggins 2001). Although there are many validity issues involved in classroom assessment that classroom teachers should consider, such as making sure the way they assess students corresponds to the type of academic learning behaviors being assessed (Ormrod 2000), the focus here is on the valid assessment and

communication of final class grades as summaries of students’ academic achievement of content knowledge of a subject. Validity addresses the accuracy of the assessment and grading procedures used by teachers (Gallagher 1998; Gredler 1999; Linn and Gronlund 2000). Do the assessment procedures and assignment of grades accurately reflect and communicate the academic achievement of the student? Validity is important because the sole purpose of grades is to accurately communicate to others the level of academic achievement that a student has obtained (Snowman and Biehler 2003). If the grades are not accurate measures of the student’s achievement, then they do not communicate the truth about the level of the student’s academic achievement. Unfortunately, as stated by Cizek, even as “grades continue to be relied upon to communicate important information about [academic] performance and progress . . . they probably don’t” (1996, 104).

Assigning grades to students is such a complex (and sometimes controversial) issue that some educators have proposed their abolition (Kohn 1999; Marzano 2000). Although I find this an interesting proposal, especially if one is trying to establish a classroom learning environment that is student-centered and encourages self-regulation and self-evaluation, the current reality for most teachers is that they are required to assign grades indicating students’ academic achievement in the subjects they teach. Therefore, grading should be as valid as possible. Not only is grading a major responsibility of classroom teachers, but it is also a practice with which they are often uncomfortable and that they find difficult (Barnes 1985; Lomax 1996; Thorndike 1997). The sources of the discomfort and difficulty for teachers regarding the grading of students seem to be threefold. First, the student activities that teachers think should constitute “academic achievement” and how to handle

James D. Allen is a professor of educational psychology in the Department of Educational and School Psychology at the College of Saint Rose in Albany, New York.

ancillary features of achievement such as students' efforts varies tremendously from teacher to teacher. Although ancillary information such as effort and attitude could be part of an overall student report, they should not be part of a grade that represents academic achievement (Tombari and Borich 1999). Second, teachers often seem to be unsettled regarding the communication function of grades, and they often try to communicate multiple pieces of information about students that can not possibly be contained within a single academic mark. This is an issue of making sure the grade is accurate as a valid communication to others. Third, because of the first two issues, many teachers assign grades that are invalid and not built on a solid principle of measurement (Cizek 1996; Marzano 2000). In addition, partially due to their long career as students experiencing invalid grading practices, as well as inadequate preservice and in-service education on assessment and grading, teachers continue to perpetuate invalid grading practices. Let us consider each of these points in greater depth.

Miscommunication and Confusing Purposes of Grades

Although students learn many things in the classroom, the primary objective is for students to learn academic content knowledge of a particular subject. In order for teachers to know if students are achieving this academic knowledge, they generally are required to not only assess students' knowledge in some way, but eventually summarize that assessment into a letter or numerical grade. This is known as "summative" evaluation. Hopefully, teachers are also gathering nongraded "formative" assessments of students to provide feedback to students as they learn, as well as considering how to motivate students to learn and encouraging them to be self-regulated learners. However, generally, teachers have to eventually place a grade on a grade sheet indicating what level of content knowledge a student has achieved in the subject listed. But why do we place a grade on a grade sheet, report card, or transcript? Why do we create a permanent written record of the grade? And why is the grade listed next to a name of an academic course such as English, U.S. History, Algebra, or Educational Psychology?

As illustrated by the title of the 1996 Yearbook of the Association for Supervision and Curriculum Development, *Communicating Student Learning* to interested parties is an important function of schools and teachers (Guskey 1996). Although there are various means to communicate student learning, currently a single report card grade for each academic subject is the most common and generally accepted system in middle and secondary schools (Bailey and McTighe 1996; Lake and Kafka 1996). Bailey and McTighe argue that as a communication system, "the primary purpose of secondary

level grades and reports [is] to communicate student achievement" so that informed decisions can be made about the student's future (1996, 120). Similarly, authors of major texts devoted to classroom assessment suggest that the major reason for assigning grades is to create a public record of a student's academic achievement that can accurately and effectively communicate to others the level of mastery of a subject a student has demonstrated (Airasian 2000; Gallagher 1998; Gredler 1999; Linn and Gronlund 2000; Nitko 2001; Oosterhof 2001; Stiggins 2001). Nitko points out that: "Grades. . . are used by students, parents, other teachers, guidance counselors, school officials, postsecondary educational institutions, and employers. Therefore [teachers] must assign grades with utmost care and maintain their validity" (2001, 365). However, according to Marzano, in contrast to teachers', students', parents', and community members' assumption that grades are valid "measures of student achievement . . . grades are so imprecise that they are almost meaningless" (2000, 1). Due to the wide variability in the criteria used in grading practices from teacher to teacher, the validity of student grades is unknown and they have limited value as guides for planning the academic and career futures of students (Thorndike 1997). Thus, if a single grade on a report card or transcript is to effectively communicate information to all these varied parties, then that single grade has to have some shared and accurate meaning (O'Connor 1995).

This lack of shared meaning seems to be found throughout our education system. A study by Baron (2000) shows that there is lack of coherence in the beliefs about grades held by parents and students and those held by the education community. Even in the same school, teachers often hold very different views about the purpose of grades and fail to communicate with their colleagues about their grading practices (Kain 1996). Grading practices by teachers rarely follow the measurement principles and grading practices recommended in measurement textbooks (Cross and Frary 1996; Frary, Cross, and Weber 1993). New teachers often work independently and are left to figure out their own grading policies, gradually adhering to the school's norms. There is a similar lack of coherence and communication among college teachers (Barnes, Bull, Campbell, and Perry 1998). Friedman and Frisbie (1995, 2000) make a particularly strong argument for making sure that report card grades accurately report information to parents about a student's academic progress and that teachers and administrators share a common understanding of what information a grade should communicate. They suggest that since grades become part of a students' permanent record, the purpose of these grades must be to communicate a valid summary of a student's academic achievement in the subject that is listed next to the grade on the record.

Grading systems used by teachers vary widely and unpredictably and often have low levels of validity due to the inclusion of nonacademic criteria used in the calculation of grades (Allen and Lambating 2001; Brookhart 1994; 2004; Frary, Cross, and Weber 1993; Olson 1989). Teachers have been found to make decisions about grades related to student effort in attempts to be "fair" in their grading practices (Barnes 1985). Studies have found that two out of three teachers believe that effort, student conduct, and attitude should influence final grades of students (Cross and Frary 1996; Frary, Cross, and Weber 1993). It has also been shown that grades are used as a motivational tool as well as to develop good study habits (Oosterhof 2001) and desirable classroom management behaviors (Allen 1983). Grades should not be a hodgepodge of factors such as student's level of effort, innate aptitude, compliance to rules, attendance, social behaviors, attitudes, or other nonachievement measures (Friedman and Frisbie 2000; Ornstein 1994). Although these factors may indirectly influence students' achievement of content knowledge, subjective—and often unknown to the teacher—factors such as these complicate the ability to interpret a grade since these factors may directly conflict with each other and distort the meaning of a grade measuring academic achievement (Cross and Frary 1996; Guskey 1994; Linn and Gronlund 2000; Nitko 2001; Stiggins 2001; Stumpo 1997). Nonacademic factors are often used as criteria for assigning grades because some teachers consider the consequences of grades more important than the value of clear communication of information and the interpretability of the grades (Brookhart 1993). It follows then that instead of the grade being a function of what a student has learned it has become a function of many variables. Simply put, it would appear that grades are often measures of how well a student lives up to the teacher's expectation of what a good student is rather than measuring the student's academic achievement in the subject matter objectives.

A grade can not be a teacher's "merged judgment"¹ of these factors, since as a single letter or numeric mark, the reported grade must communicate a single fact about the student if it is to be a valid or accurate source of information coherently shared between the reporter of the grade and the grade report's audience. How is the reader of a student's single grade on a transcript to know which factors are included and how much each unknown factor was weighed by the grader to determine the grade? Also, since many of these factors such as effort, motivation, and student attitude are subjective measures made by a teacher, their inclusion in a grade related to academic achievement increases the chance for the grade to be biased or unreliable, and thus invalid. The purpose of an academic report is to communicate the level of academic

achievement that a student has developed over a course of study. Therefore, the sole purpose of a grade on an academic report, if it is to be a valid source of information, is to communicate the academic achievement of the student. If other factors about the student are deemed important, such as a student's attitude, level of effort, or social behavior, then other appropriate forms of reporting these factors must be made available and used. If a multidimensional view of the student is desired, then a multidimensional system of reporting is required. Using a single grade as a summary of a teacher's "merged judgment" of a student leads to miscommunication, confusion, and a continuation of the lack of coherence among stakeholders about what a grade represents.

Since important decisions are often based on a student's grade, invalid grades may result in dire consequences for the student. Grades can open up or close down important learning opportunities for students (Jasmine 1999). With high grades, students get admitted to colleges and universities of their choice and receive scholarships and tuition assistance, since grades are a major selection criterion in the college admission process. The reverse is also true. It is very difficult for students to get admitted to some schools if their grades are not sufficiently high. Invalid grades that understate the student's knowledge may prevent a student with ability to pursue certain educational or career opportunities. Also, based on principles of attribution and social cognitive theories, if students receive grades lower than ones that accurately depict their true level of academic knowledge, it may lead students to believe they lack the ability to succeed academically and lower their sense of self-efficacy as well as their motivation to learn (Pintrich and Schunk 2002).

Grading and Lack of Professional Training

Cizek argues that the "lack of knowledge and interest in grading translates into a serious information breakdown in education" and that "reforming classroom assessment and grading practices will require educators' commitment to professional development, [and] classroom-relevant training programs" (1996, 103). Cizek's statement implies that an important area that needs to be addressed is the training of teachers in grading practices based on sound measurement principles relevant to their classroom lives.

This lack of knowledge about measurement theory and application to grading practices is a pervasive problem with preservice teacher training at the college level (Goodwin 2001; Schafer 1991; Stiggins 1991, 1999). One of the goals of a teacher education program should be to prepare preservice and in-service teachers to develop effective methods to assess students and to communicate clearly and accurately through their grading practices that assessment to others. However, very few teacher educa-

tion programs include measurement or assessment courses. Allen and Lambating (2001) found in a random sample of teacher education programs that less than one-third required an assessment course, and many of those that did were courses focused on "informal" assessments, or standardized assessment of students with special needs and not focused on classroom assessment and grading. Fewer than half of the fifty states require specific coursework on assessment for their initial certification of teachers (Lomax 1996; O'Sullivan and Chalnack 1991; Stiggins 1999).

Although assigning grades is probably the most important measurement decision that classroom teachers make, the coverage of grading in assessment textbooks is often not as fully developed as other measurement topics that are less relevant to teachers' day-to-day assessment practices (Airasian 1991; Lomax 1996). According to Stiggins (1999), how the concepts of "reliability" and "validity" are related to classroom grading practices is not addressed in the courses which introduce these terms to our preservice teachers. It is important to look at this issue because validity and reliability are considered the most fundamental principles related to measurement and therefore important to classroom assessment and grading (Gallagher 1998; Gredler 1999; Linn and Gronlund 2000).

Some argue that even when teachers are provided with some measurement instruction, they still use subjective value judgements when assigning grades (Brookhart 1993). Undergraduate teacher education majors, when asked about the criteria that should be used for their own grades, believe that "effort" is more important than amount of academic content learned (Placier 1995). One contributing factor may be that after sixteen years of obtaining grades based on factors other than academic achievement, teachers-in-training have a difficult time accepting theoretical principles that do not match with their personal experience. Many beliefs about school practices are well established before students enter college and often are resistant to change (Britzman 1986, 1991; Ginsberg and Clift 1990; Holt-Reynolds 1992; Pajares 1992; Richardson 1996). They form many of their perspectives about teaching from their years of observing teachers and their teaching practices (Lortie 1975). They have been recipients of hundreds of grades from their K-12 teachers and college professors before taking on the responsibility of assigning grades to their own students. Their perception regarding grades comes from their own long experience as students.

Brookhart (1998) suggests that classroom assessment and grading practices are at the center of effective management of classroom instruction and learning. Through the use of real classroom scenarios, preservice teachers need to be taught assessment strategies in relationship to instruction and not as decontextualized

measurement principles. As the past president of the American Educational Research Association, Lorrie Shepard has stated: "The transformation of assessment practices cannot be accomplished in separate tests and measurement courses, but rather should be a central concern in teaching methods courses" (2000, 4). In addition to instruction on how to assess and grade using sound principles of measurement, research suggests that preservice teachers need hands-on experience in grading students and how to work with cooperating teachers who assess and grade in ways different than those learned by the preservice teachers (Barnes 1985; Lomax 1996).

What the literature suggests is that educators at all levels make decisions when assigning grades that are not based on sound principles of validity that ensure the grade is a meaningful communication of a student's level of academic achievement. The literature also suggests that students in teacher education programs may be more influenced by the grading practices they have experienced as students in the past, as well as in their current courses taught by their education professors, than by what they learn about assessment and grading in their courses. Additionally, teachers in the field, as products of teacher education programs, seem to exhibit grading practices that confirm that they have not been influenced by measurement courses (Lambating and Allen 2002). This may be because they did not take any assessment courses, or because their long-held beliefs about grading were left unchallenged and the courses did not focus on assessment and grading issues related to measuring classroom learning.

Educational Implications and Conclusion

Concerns about the validity and reliability of grades for communicating meaningful information about students' academic progress have been raised for a long time (see Starch and Elliot 1912, 1913a, 1913b; Adams 1932). In addition, trying to help teachers to understand the purpose and effective functions of grades in the overall evaluation system has been addressed repeatedly in the literature (Airasian 2000; Brookhart 1993; Cross and Fray 1996; Gredler 1999; Guskey 1996; Linn and Gronlund 2000; Marzano 2000; O'Connor 1995; Stiggins 2001). However, there seems to be little progress being made in this area in actual classroom practice.

Two major thrusts need to occur in reforming grading practices. First, if factors such as effort, attitude, compliance, and behavior are to be noted about a student on a report card, then they should be reported with a separate mark and not figured in as part of a grade for academic achievement of content knowledge. However, as in most situations, if a teacher must summarize and communicate a student's classroom progress in an academic subject through a single report

card grade, then there must be a consensus that the grade represents the most accurate statement of the student's academic achievement, and only academic achievement. This is the essence of valid assessment. To include nonacademic criteria, such as the student's effort, compliance, attitude, or behavior, makes the grade impossible to interpret in any meaningful way. Perhaps, a simple way to reach this consensus is to teach ourselves and those we prepare to be teachers to reflect on the following question: "If I was given a student's transcript with a single letter grade listed next to the subject I teach, what would be the most logical interpretation I could make about what the grade represents about the student's knowledge of that academic subject?" Therefore, that is what I should try to have my grades communicate to whomever will read and interpret them in the future.

In order for teachers to act consistently in assigning valid grades based only on appropriate achievement criteria, a second major initiative needs to be undertaken to help teachers understand how to make good grading decisions. This initiative is best addressed through teacher education programs taking on the challenge to improve the assessment training of their students and improve their own grading practices. This entails several dimensions.

First, students' long-held beliefs about the purpose and use of grades need to be challenged by teacher educators. Students' beliefs and value systems related to grades need to be exposed and examined to help them understand the unscientific basis of their grading beliefs. Second, once these beliefs are exposed, instructors must provide students with the theoretical base for good assessment and grading practices as explicated by measurement experts that would replace students' naive notions of assessment and grading. This could be either through self-contained measurement courses taught in a relevant manner by educational psychologists, or integrated into methods courses through collaboration between educational psychology and teacher-education specialists. It would help if more teacher-education programs required adequate instruction on classroom assessment and grading practices. There also needs to be more effective and meaningful grading practices addressed in-depth in measurement textbooks. Third, teacher education students need to be provided with opportunities to encounter grading activities before they are placed into student teaching, in order to practice applying assessment principles and theory to classroom grading issues. Finally, during student teaching experiences, education majors must be given the opportunity, in conjunction with their cooperating teachers and the support of their college supervisors, to actually develop and implement a valid evaluation and grading plan. Schools of education need to work with school district teachers to help improve the

communication system for which grades function. Providing in-service "assessment and grading" workshops for practicing teachers, especially those operating as cooperating teachers, might help to establish a consensus of what is appropriate criteria to use for determining and assigning valid grades to indicate academic achievement.

One way to accomplish many of the above steps is through the use of case studies that focus on assessment and grading dilemmas often faced by real teachers. Discussion of case studies can help students to reflect on and expose their belief systems about grades and grading, and analyze them in relationship to educational psychology assessment principles such as validity. One example is the Sarah Hanover case which focuses on a grading dilemma a teacher must deal with when the question of the validity of a student's grade is raised by the student's parent (Silverman, Welty, and Lyon 1996).

However, the area that may be the most difficult to address is the change in the grading practices that teacher educators use in evaluating students. As long as preservice and in-service teachers take classes from education professors who base grading decisions on more than academic achievement, they will have little reason to either believe what we say or practice what we preach about assessment and grading. As teacher educators, we need to model sound grading practices in our own courses in which grades accurately communicate students' achievement of content knowledge learned in our courses, and not how hard they work or how often they attend our classes.

My intention in this article has been to suggest that by giving serious reflection to the meaning of the educational psychology measurement principle of validity, grading practices can improve and the grades we assign to students as teachers can be more accurate and educationally meaningful. We need to begin to break the cycle of invalid grading practices that prevail throughout the education system, and the only behaviors we as teachers can truly control are our own.

Key words: grading, education system, assessment

NOTE

1. The author has borrowed this phrase from an anonymous reviewer.

REFERENCES

- Adams, W. L. 1932. Why teachers say they fail pupils. *Educational Administration and Supervision* 18:594-600.
- Airasian, P. W. 1991. Perspectives on measurement instruction. *Educational Measurement: Issues and Practice* 10 (1): 13-16, 26.
- . 2000. *Assessment in the classroom: A concise approach*. 2nd ed. Boston: McGraw-Hill.
- Allen, J. D. 1983. Classroom management: Students' perspectives, goals and strategies. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada, April.

- Allen, J. D., and J. Lambating. 2001. Validity and reliability in assessment and grading: Perspectives of preservice and inservice teachers and teacher education professors. Paper presented at the annual meeting of the American Educational Research Association, Seattle, April.
- Bailey, J., and J. McTighe. 1996. Reporting achievement at the secondary level: What and how. In *Guskey 1996*, 119–40.
- Barnes, L. B., K. S. Bull, N. J. Campbell, and K. M. Perry. 1998. Discipline-related differences in teaching and grading philosophies among undergraduate teaching faculty. Paper presented at the annual meeting of the American Educational Research Association, San Diego, April.
- Barnes, S. 1985. A study of classroom pupil evaluation: The missing link in teacher education. *Journal of Teacher Education* 36 (4): 46–49.
- Baron, P. A. B. 2000. Consequential validity for high school grades: What is the meaning of grades for senders and receivers? Paper presented at the annual meeting of the American Educational Research Association, New Orleans, April.
- Britzman, D. P. 1986. Cultural myths in the making of a teacher: Biography and social structure in teacher education. *Harvard Educational Review* 56 (4): 442–56.
- . 1991. *Practice makes practice: A critical study of learning to teach*. New York: State University of New York Press.
- Brookhart, S. M. 1993. Teachers' grading practices: Meaning and values. *Journal of Educational Measurement* 30 (2): 123–42.
- . 1994. Teachers' grading: Practice and theory. *Applied Measurement in Education* 7 (4): 279–301.
- . 1998. Teaching about grading and communicating assessment results. Paper presented at the annual meeting of the National Council on Measurement in Education, San Diego, April, 1998.
- . 2004. *Grading*. Upper Saddle River, NJ: Pearson/Merrill/Prentice Hall.
- Cizek, G. J. 1996. Grades: The final frontier in assessment reform. *NASSP Bulletin* 80 (584): 103–10.
- Cross, L. H., and R. B. Frary. 1996. Hodgepodge grading: Endorsed by students and teachers alike. Paper presented at the annual meeting of the National Council on Measurement in Education, New York, April.
- 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 Practice* 12 (3): 2330.
- Friedman, S. J., and D. A. Frisbie. 1995. The influence of report cards on the validity of grades reported to parents. *Educational and Psychological Measurement* 55 (1): 5–26.
- . 2000. Making report cards measure up. *Education Digest* 65 (5): 45–50.
- Gallagher, J. D. 1998. *Classroom assessment for teachers*. Upper Saddle River, NJ: Merrill/Prentice Hall.
- Ginsburg, M. B., and R. T. Clift. 1990. The hidden curriculum of preservice teacher education. In *Handbook of research on teacher education*, ed. W. R. Houston, 450–65. New York: Macmillan.
- Goodwin, A. L. 2001. The case of one child: Making the shift from personal knowledge to informed practice. Paper presented at the annual meeting of the American Educational Research Association, Seattle, April.
- Gredler, M. E. 1999. *Classroom assessment and learning*. New York: Longman.
- Guskey, T. R. 1994. Making the grade: What benefits students? *Educational Leadership* 52 (2): 14–20.
- . 1996. *ASCD Yearbook, 1996: Communicating student learning*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Holt-Reynolds, D. 1992. Personal history-based beliefs as relevant prior knowledge in coursework: Can we practice what we preach? *American Educational Research Journal* 29 (2): 325–49.
- Jasmine, T. 1999. Grade distributions, grading procedures, and students' evaluations of instructors: A justice perspective. *Journal of Psychology* 133 (3): 263–71.
- Kain, D. L. 1996. Looking beneath the surface: Teacher collaboration through the lens of grading practices. *Teachers College Record* 97 (4): 569–87.
- Kohn, A. 1999. Grading is degrading. *Education Digest* 65 (1): 59–64.
- Lake, K., and K. Kafka. 1996. Reporting methods in grades K–8. In *Guskey 1996*, 90–118.
- Lambating, J., and J. D. Allen. 2002. How the multiple functions of grades influence their validity and value as measures of academic achievement. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, April.
- Linn, R. L., and N. E. Gronlund. 2000. *Measurement and assessment in teaching*. 8th ed. Englewood Cliffs, NJ: Merrill/Prentice Hall.
- Lomax, R. G. 1996. On becoming assessment literate: An initial look at preservice teachers' beliefs and practices. *Teacher Educator* 31 (4): 292–303.
- Lortie, D. 1975. *Schoolteacher: A sociological study*. Chicago: University of Chicago Press.
- Marzano, R. J. 2000. *Transforming classroom grading*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Nitko, A. J. 2001. *Educational assessment of students*. 3rd ed. Upper Saddle River, NJ: Merrill/Prentice Hall.
- O'Conner, K. 1995. Guidelines for grading that support learning and student success. *NASSP Bulletin* 79 (571): 91–101.
- Olson, G. H. 1989. On the validity of performance grades: The relationship between teacher-assigned grades and standard measures of subject matter acquisition. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, March.
- Oosterhof, A. 2001. *Classroom application of educational measurement*. Upper Saddle River, NJ: Prentice Hall.
- Ormrod, J. E. 2000. *Educational psychology: Developing learners*. 3rd ed. Upper Saddle River, NJ: Merrill/Prentice Hall.
- Ornstein, A. C. 1994. Grading practices and policies: An overview and some suggestions. *NASSP Bulletin* 78 (561): 55–64.
- O'Sullivan, R. G., and M. K. Chalnack. 1991. Measurement-related course work requirements for teacher certification and recertification. *Educational Measurement: Issues and Practice* 10 (1): 17–19, 23.
- Pajares, M. F. 1992. Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research* 62 (3): 307–32.
- Pintrich, P. R., and D. H. Schunk. 2002. *Motivation in education*. 2nd ed. Upper Saddle River, NJ: Merrill/Prentice Hall.
- Placier, M. 1995. "But I have to have an A": Probing the cultural meanings and ethical dilemmas of grades in teacher education. *Teacher Education Quarterly* 22 (1): 45–63.
- Richardson, V. 1996. The role of attitudes and beliefs in learning to teach. In *Handbook of research on teacher education*, 2nd ed., ed. J. Sikula, T. Buttery, and E. Guyton, 102–19. New York: Macmillan.
- Schafer, W. D. 1991. Essential assessment skills in professional education of teachers. *Educational Measurement: Issues and Practice* 10 (1): 3–6, 12.
- Shepard, L. A. 2000. The role of assessment in a learning culture. *Educational Researcher* 29 (7): 4–14.
- Silverman, R., W. M. Welty, and S. Lyon. 1996. *Case studies for teacher problem solving*. 2nd ed. New York: McGraw-Hill.
- Snowman, J., and R. F. Biehler. 2003. *Psychology applied to teaching*. 10th ed. Boston: Houghton Mifflin.
- Starch, D., and E. C. Elliot. 1912. Reliability of grading of high-school work in English. *School Review* 20:442–57.
- . 1913a. Reliability of grading work in mathematics. *School Review* 21:254–59.
- . 1913b. Reliability of grading work in history. *School Review* 20:676–81.
- Stiggins, R. J. 1991. Relevant classroom assessment training for teachers. *Educational Measurement: Issues and Practice* 10 (1): 7–12.
- . 1999. Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice* 18 (1): 23–27.
- . 2001. *Student-involved classroom assessment*. 3rd ed. Upper Saddle River, NJ: Merrill/Prentice Hall.
- Stumpo, V. M. 1997. 3-tier grading sharpens student assessment. *Education Digest* 63 (4): 51–54.
- Thorndike, R. M. 1997. *Measurement and Evaluation*. 6th ed. Upper Saddle River, NJ: Merrill/Prentice Hall.
- Tombari, M., and G. Borich. 1999. *Authentic assessment in the classroom*. Upper Saddle River, NJ: Merrill/Prentice Hall.

Copyright of Clearing House is the property of Heldref Publications. The copyright in an individual article may be maintained by the author in certain cases. Content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.