

Differing Opinions on Testing Between Preservice and Inservice Teachers

KATHY E. GREEN

University of Denver

ABSTRACT Studies of teachers' use of tests suggest that classroom tests are widely used and that standardized test results are rarely used. What is the genesis of this lack of use? A previous comparison of pre- and inservice teachers' attitudes toward assessment suggested no differences. This study assessed the different opinions among sophomores ($n = 84$), seniors ($n = 152$), and inservice teachers ($n = 553$) about the use of classroom and standardized tests. Significant differences were found; preservice teachers had less favorable attitudes toward classroom testing than teachers did and more favorable attitudes toward standardized testing.

This study assessed differences among college students entering a teacher education program, students finishing a teacher education program, and inservice teachers concerning their opinions of some aspects of classroom and standardized testing. Although numerous studies of inservice teachers' attitudes toward testing have been conducted, little research is available regarding preservice teachers' views of testing and of the genesis of teachers' views of testing.

Interest in this topic stemmed from research findings suggesting that the results of standardized tests are not used by most teachers. If standardized testing is to continue, the failure to use results is wasteful. Other studies have identified some of the reasons for the lack of use. This study's purpose was to determine whether opinions about the usefulness of standardized and other tests were negative for students before they even entered the teaching profession. When were those attitudes developed? Are attitudes fixed by students' educational experiences prior to entry into a teacher education program? Are preservice teachers socialized by their educational programs into resistance to testing? Do negative attitudes appear upon entry into the profession because of socialization into the school culture? Or do they appear after several years of service as a teacher because of personal experiences in the classroom?

I found only one study that addressed differences in opinions of pre- and inservice teachers (Reeves & Kazelskis, 1985). That study examined a broad range of issues salient to first-year teachers; only one item addressed testing specifically. Reeves and Kazelskis found no significant differences between pre- and inservice teachers' opinions about testing, as measured by that item. In this study, I sought more information pertinent to the development of opinions about testing.

Test use in U.S. schools has been and continues to be extensive. It has been estimated that from 10 to 15% of class time is spent dealing with tests (Carlberg, 1981; Newman & Stallings, 1982). Gullickson (1982) found that 95% of the teachers he surveyed gave tests at least once every 2 weeks. The estimated percentage of students' course grades that are based on test scores is 40 to 50%, ranging from 0 to 100% (Gullickson, 1984; McKee & Manning-Curtis, 1982; Newman & Stallings). Classroom tests, thus, are used frequently and may, at times, be used almost exclusively in determining students' grades.

In contrast, a review of past practice suggests minimal teacher use of *standardized* test results in making instructional decisions (Fennessey, 1982; Green & Williams, 1989; Lazar-Morrison, Polin, Moy, & Burry, 1980; Rudell, 1985). Stetz and Beck (1979) conducted a national study of over 3,000 teachers' opinions about standardized tests. They noted that 41% of the teachers surveyed reported making little use of test results, a finding consistent with that of Goslin (1967) from several decades ago and that of Boyd, McKenna, Stake, and Yachinsky (1975). Test results were viewed as providing information that was supplemental to the wider variety of information that the teachers already possessed. Reasons offered for why standardized tests are given but results not always used by teachers include resistance to a perceived narrowing of the curriculum, resistance to management control, accountability avoidance (Darling-Hammond,

Address correspondence to Kathy E. Green, University of Denver, School of Education, Denver, CO 80208.

1985), and a limited understanding of score interpretation resulting from inadequate preservice training (Cramer & Slakter, 1968; Gullickson & Hopkins, 1987). Marso and Pigge (1988) found that teachers perceive a lower need for standardized testing skills than for classroom testing skills. They also found that teachers reported lower proficiencies in standardized test score use and interpretation than in classroom test score use and interpretation.

The results of those studies suggest that inservice teachers use classroom tests extensively but make little use of standardized test results. This suggests that inservice teachers, in general, hold positive attitudes toward classroom tests and less positive attitudes toward standardized tests. The literature does not lead to any predictions about preservice teachers' attitudes toward tests.

This study assessed differences between preservice and inservice teachers' opinions about testing and test use. The following research hypotheses were formulated to direct the study.

- H1. There are significant differences in opinions about the testing and test use between preservice and inservice teachers.
- H2. There are significant differences in opinions about testing between students beginning their preparation (sophomores) and students finishing their preparation (seniors).
- H3. There are significant differences among inservice teachers with differing years of experience.

Method

Samples

Three samples were drawn for this study. They were samples of (a) practicing teachers, (b) college sophomores beginning a teacher education program, and (c) college seniors completing a teacher education program (but prior to student teaching). For the first sample, survey forms were mailed in a rural western state to 700 teachers randomly selected from the State Department of Education list of all licensed educators. During the spring semester of 1986, teachers were sent a letter explaining the nature of the study, a survey form, and a stamped return envelope. With two follow-up mailings, a total of 555 questionnaires were received, or 81% of the deliverable envelopes. (Twelve questionnaires were undeliverable, 4 persons refused to respond, and 133 persons did not reply.) No compulsory statewide standardized testing program was in place in the state.

The second sample was a convenience sample of three sections of an educational foundations class typically taken by college sophomores who have just enrolled in a teacher preparation program ($n = 84$). The course examines educational thought and practice in the United States. The classes were taught in an 8-week block, meet-

ing for 50 min per day, 4 days per week. Survey forms were distributed in class and completed during class time.

The third sample was also a convenience sample of four sections of a tests and measurement class taken by college seniors ($n = 152$). The course is typically taken after coursework is almost complete, but prior to student teaching. The course provides instruction in basic statistics, classroom test construction and analysis, and standardized test use and interpretation. The course was also taught in an 8-week block, with the same schedule as the foundations course. Survey forms were distributed during the first week of class and completed during class time. Survey forms took from 10 to 30 min to complete. Responses were anonymous. Both sophomores and seniors were attending a public university in a small western town.

Table 1 presents descriptive information for the three samples.

Instruments

Three different forms with overlapping questions were used in this study. The survey form sent to the teachers contained questions regarding training in tests and measurement, subject and grades taught, tests given, and attitudes toward both standardized and classroom tests. The questionnaire was two pages in length, double-sided and contained 49 questions. The form given to the sophomores had 43 questions and was one page in length, double-sided. The form given to the seniors was three pages in length, single-sided. The latter two forms differed by the inclusion of an evaluation anxiety scale and items eliciting importance of contemporary measurement practices for the seniors. Although different formats may have affected responses to some extent, all the forms began with several demographic questions followed by the items relevant to this study. Any form differences would, then, likely be minimized for those initial items.

There were 18 items common to the three forms. Sixteen of the items were Likert items with a 1 to 6 (*strongly disagree* to *strongly agree*) response format. Likert-scale items were drawn from a previously developed measure of attitudes toward both standardized and classroom testing (Green & Stager, 1986). Internal consistency reliabilities of the measures ranged from .63 to .75. The re-

Table 1.—Description of Samples

Item	Sophomores ($n = 84$)	Seniors ($n = 152$)	Teachers ($n = 553$)
Percentage female	84	152	553
Mean age	73.0	75.9	63.6
Age range	18-33	20-45	—
Mean years in teaching	—	—	12

maining two items asked how many hours per week teachers spend in testing activities and how much of a student's grade should be based on test results. The study examined differences found among groups on those items. Item content is presented in Table 2, in which items are grouped by content (opinions about standardized tests, classroom tests, and about personal liking for tests).

Data were analyzed using multivariate analyses of variance, followed by univariate analyses of variance. If univariate results were significant, I used Tukey's HSD test to assess the significance of pairwise post hoc differences. Samples of both items and persons were limited; therefore, results may not be widely generalizable.

Results

Significant multivariate differences were found across opinion items (Wilks's lambda = .70, $p < .001$) when the three samples were compared (Table 2). Hypothesis 1 was supported. Differences were found between teachers and students for all items, with significance levels varying from .02 to .001 for individual items. Opinions were not consistently more positive across all items for teachers or for students. For instance, whereas teachers were most likely to feel that standardized tests address important educational outcomes, teachers were least likely to find that standardized tests serve a useful purpose. In general, though, students favored use of standardized tests for student or teacher evaluation more than teachers did. Al-

Table 2.—Means and Standard Deviations for Opinions About Testing by Group

Variable	Sophomores (<i>n</i> = 84)	Seniors (<i>n</i> = 152)	Teachers (<i>n</i> = 553)	<i>p</i>	1	2	3
Hours spent in testing/week	10.43 (6.72)	9.18 (6.43)	4.37 (4.05)	.001	*	*	—
Percentage grade based on test	49.63 (15.48)	46.94 (18.71)	41.31 (22.68)	.001	*	*	—
Standardized test items							
Standardized tests are the best way to evaluate a teacher's effectiveness.	2.79 (1.03)	2.83 (1.10)	2.12 (1.18)	.001	*	*	—
Teachers whose students score higher on standardized tests should receive higher salaries.	2.53 (1.07)	2.33 (1.17)	1.74 (1.01)	.001	*	*	—
Requiring <i>students</i> to pass competency tests would raise educational standards.	4.14 (1.13)	3.89 (1.09)	3.69 (1.26)	.001	*	*	—
Requiring <i>teachers</i> to pass competency tests would raise educational standards.	4.35 (.90)	4.09 (1.27)	3.30 (1.34)	.001	*	*	—
Standardized tests assess important educational outcomes.	3.47 (1.04)	3.54 (.87)	3.95 (.88)	.001	*	*	—
Standardized tests serve a useful purpose.	4.02 (.83)	3.97 (.81)	2.93 (.97)	.001	*	*	—
Standardized tests force teachers to "teach to the test."	3.05 (1.19)	2.74 (.98)	3.11 (1.22)	.02	—	*	—
Classroom test items							
Test construction takes too much teacher time.	4.57 (1.02)	4.36 (.85)	3.97 (.88)	.001	*	*	—
Test scores are a fair way to grade students.	3.42 (1.02)	3.32 (1.13)	4.04 (.84)	.001	*	*	—
Testing has a favorable impact on student motivation.	4.00 (.71)	3.88 (1.00)	4.16 (.88)	.01	—	*	—
Tests are of little value in identifying learning problems.	1.76 (.96)	1.43 (.84)	1.44 (1.05)	.01	*	—	*
It is relatively easy to construct tests in my subject area.	4.11 (1.25)	3.51 (1.34)	4.35 (.89)	.001	—	*	*
Tests measure only minor aspects of what students can learn.	2.92 (1.13)	3.01 (1.13)	3.24 (1.00)	.01	*	—	—
Personal reflections							
I do(did) well on tests.	4.05 (1.04)	4.00 (1.10)	4.46 (.94)	.001	*	*	—
I personally dislike taking tests.	3.13 (1.35)	3.12 (1.14)	3.46 (1.16)	.01	—	*	—
The tests I have taken were generally good assessments of my knowledge of an area.	3.65 (1.08)	3.41 (1.10)	4.09 (.82)	.001	*	*	—

Note. For opinion items, the scale ranged from *strongly disagree* (1) to *strongly agree* (6). Standard deviations are presented in parentheses. Asterisks(*) indicate significant ($p < .05$) differences between groups: 1 = teachers versus sophomores, 2 = teachers versus seniors, 3 = sophomores versus seniors.

though the students were less likely to say that they do well on tests and that tests previously taken were good assessments of their ability, the students were also less likely to say that they disliked taking tests. Students' opinions about classroom testing were less favorable than were teachers' opinions for all but one item. Differences were also found between teachers and students in estimates of time spent in testing and in the percentage of students' grades based on test scores.

Hypothesis 2 was not supported. Only two significant differences in means were found between the sophomores and the seniors. One difference was found for the item "It is relatively easy to construct tests in my subject area." Sophomores tended to agree with that statement more than the seniors did. Because the seniors were required to complete a task involving test construction, the impending course requirement may have influenced their opinions. The second difference was found for the item "Tests are of little value in identifying learning problems," with more positive opinions expressed by seniors than by sophomores.

Hypothesis 3 was tested by dividing teachers into three groups: 0 to 1 years, 2 to 5 years, and 5+ years of experience as a teacher. No significant multivariate or univariate differences were found, so Hypothesis 3 was not supported. However, there were few teachers with 0 to 1 years of experience in the sample. Because of the small number of teachers with 0 to 1 years of teaching (46 teachers; 8.7% of the data file), groups were reformed as follows: 0 to 3 years, 4 to 6 years, and 6+ years of experience. Still, no significant multivariate or univariate differences were found. (In addition, no differences were found between teachers with 0 to 3 years of experience and those with 6 or more years of experience.)

Discussion

This study was undertaken to examine whether differences in opinions about testing would be discerned between preservice and inservice teachers and whether those differences would suggest a progression. The differences found suggest that teacher education students are less favorable to classroom testing and more favorable to standardized testing than teachers are. Differences were *not* found between sophomores and seniors, however. Nor were opinions about testing found to depend upon years of experience in teaching. Those results do not reflect a developmental progression. The shift in opinion seems to occur when beginning a teaching position, suggesting effects that result from job requirements or socialization as a teacher more than from a developmental trend. Differences between students and teachers, then, seem likely to be caused by direct teacher experience with creating, administering, and using tests or by acculturation into life as a teacher in a school. That conclusion suggests that if

one wishes to affect teachers' opinions about testing, provision of inservice experiences may be a more profitable avenue than additional preservice education.

Test use. The teachers sampled in this study reported spending an average of about 11% of their time in testing, which is consistent with estimates reported in the literature (10 to 15%). The finding in this study that an average of 41% of the students' grades was based on test results is also consistent with estimates reported in the literature (40 to 50%). Estimates of the time needed for testing activities obtained from students sampled in this study were much higher (23% and 26% for seniors and sophomores, respectively) than the estimates obtained from the teachers' reports. Although students' estimates of the percentage of grade based on test scores were significantly higher than those of teachers, they were within the range reported in the literature. Students, then, who lack an experiential base, seem either to have exaggerated views regarding the time that teachers spend on testing-related activities or think that it will take them longer to construct tests.

Beginning teachers also lack an experiential base. One might ask whether beginning teachers spend more time in test-related activities than do teachers with more experience, because beginning teachers may not have files of tests to draw upon. Mean reported time spent in testing was higher for first- and second-year teachers (means of 5.4 and 5.7 hours per week) than for teachers with more experience (mean for third year = 2.3, 4th year = 2.8, 5th year = 3.8). Thus, students may be accurate in their perception of the time needed by novices for testing-related activities.

Standardized testing. The students' opinions ranged from neutral to positive regarding the use of standardized tests and were, on average, significantly more positive than the teachers' opinions. One explanation for the positive opinions may be that students have extremely limited personal experience with standardized tests (their own or their friends) and so have a limited basis upon which to judge test effectiveness. By college level, most students have taken a number of standardized tests but may not be aware of the results, may not have been directly affected by the results, or may have been affected by the results at a time when they were too young to understand or argue. Students may believe that the tests must be useful because "authorities and experts" sanction their administration. Students' opinions may, then, be shaped by the positive *public* value placed on tests, as well as by their educational programs. The tests and measurement course taken by many preservice teachers emphasizes how tests can be valuable if used properly. One can argue that most students view themselves as intending to use tests properly. In contrast, many teachers are required to give standardized tests, and they may also be required to take them.

Preservice-inservice differences might be even more extreme in states where the stakes attached to standardized test use are higher—where the teacher's job or salary depends upon test results. Teachers develop a broader base of experience with standardized testing, and they may be more aware of the limitations of the tests and of the controversy surrounding standardized testing. The measurement profession is unclear about the value of standardized testing; it is not surprising that teachers also have reservations.

Classroom testing. Differences were also found between teachers and teacher education students for most classroom test items, though differences were not as pronounced for these items. The result is in contrast to Reeves and Kazelskis's (1985) finding of no differences between similar groups. The result of somewhat less favorable opinions of preservice than inservice teachers toward classroom testing may have stemmed from the frequent test taking by students versus the frequent use of tests by teachers. By the time students are seniors in college, they will have taken a larger number of classroom tests than standardized tests and thus will have considerably more experience in evaluating their effectiveness. Students undoubtedly encounter classroom tests and test questions that they consider to be unfair assessments of their knowledge. Such experiences may temper their opinions toward classroom tests. In contrast, because most teachers rely to some extent on test results in assigning grades and in evaluating instruction, opinions may change to conform with this behavior. Teachers' opinions may also be influenced by an experiential understanding of testing gained through learning how informative test results can be.

Because it is unlikely that the widespread use of classroom and standardized tests will diminish, teachers will continue to be called upon to use tests to make decisions that are important in the lives of students. Teachers need to be competent in test construction and interpretation. However, if tests are to be used effectively as part of the instructional process, teachers must perceive the positive aspects of test use. If a teacher finds that task impossible, that teacher should discontinue traditional test use and seek alternative assessment techniques, within the boundaries allowed by the district. Teachers should communicate positive feelings about the tests they give to their students. Teachers will probably be more likely to do so if they have positive opinions of tests. Tests are often viewed as evaluative; they may more effectively be viewed as informative and prescriptive.

If teacher educators wish to affect prospective teachers' views, they may need to both clarify their own views about the place of testing in instruction and clearly present arguments about testing, pro and con, to their classes. Well-constructed classroom assessments, whether paper-and-pencil, portfolio, or performance measures, provide diagnostic and prescriptive information about

the students' progress and about the effectiveness of instruction. This information is valuable. Poorly constructed or standardized measures that do not address the curriculum provide little information of use in the classroom. The reasons for giving tests that do not provide information useful in instruction must be clearly explained. Such tests may be mandated to provide legitimate administrative, state, or national information.

But to what extent can teacher educators shape *prospective* teachers' views? The results of this study suggest that opinions held prior to and following preservice instruction may not survive the transition to the real world of the classroom. If this is the case, the preservice course—no matter how good it is—would be ineffective in influencing attitudes. (It may, however, be highly effective in influencing the quality of testing practices by providing basic skills in test construction and interpretation.) Inservice instruction may be a better vehicle to use to produce attitude change.

This study was cross-sectional in design. A longitudinal study that examined opinions over time (from preservice to inservice) is required to identify the extent to which opinions are shaped by school requirements. Additional information regarding school characteristics affecting preservice and inservice teachers' attitudes toward testing would be of interest, as would information about differences in testing skill levels between pre- and inservice teachers.

NOTES

An earlier version of this paper was presented at the 1990 annual meeting of the National Council on Measurement in Education, held April 1990 in Boston.

Appreciation is expressed to the *Journal of Educational Research* reviewers for their helpful suggestions.

REFERENCES

- Boyd, J., McKenna, B. H., Stake, R. E., & Yachinsky, J. (1975). *A study of testing practices in the Royal Oak (MI) public schools*. Royal Oak, MI: Royal Oak City School District. (ERIC Reproduction Service No. 117 161)
- Carlberg, C. (1981). South Dakota study report. Denver, CO: Midcontinent Regional Educational Laboratory.
- Cramer, S., & Slakter, M. (1968). A scale to assess attitudes toward aptitude testing. *Measurement and Evaluation in Guidance*, 1(2).
- Darling-Hammond, L., & Wise, A. E. (1985). Beyond standardization: State standards and school improvement. *Elementary School Journal*, 85, 315-336.
- Fennessey, D. (1982). Primary teachers' assessment practices: Some implications for teacher training. Paper presented at the annual conference of the South Pacific Association for Teacher Education, Frankston, Victoria, Australia.
- Goslin, D. A. (1967). *Teachers and testing*. New York: Russell Sage Foundation.
- Green, K. E., & Stager, S. F. (1986-87). Testing: Coursework, attitudes, and practices. *Educational Research Quarterly*, 11(2), 48-55.
- Green, K. E., & Stager, S. F. (1986). Measuring attitudes of teachers toward testing. *Measurement and Evaluation in Counseling and Development*, 19, 141-150.
- Green, K. E., & Williams, E. J. (1989, March). Standardized test use by

classroom teachers: Effects of training and grade level taught. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco.

Gullickson, A. R. (1982). The practice of testing in elementary and secondary schools. (ERIC Reproduction Service No. ED 229 391)

Gullickson, A. R. (1984). Teacher perspectives of their instructional use of tests. *Journal of Educational Research*, 77, 244-248.

Gullickson, A. R., & Hopkins, K. D. (1987). The context of educational measurement instruction for preservice teachers: Professor perspectives. *Educational Measurement: Issues and Practice*, 6, 12-16.

Karmos, A. H., & Karmos, J. S. (1984). Attitudes toward standardized achievement tests and their relation to achievement test performance. *Measurement and Evaluation in Counseling and Development*, 17, 56-66.

Lazar-Morison, C., Polin, L., Moy, R., & Burry, J. (1980). A review of the literature on test use. Los Angeles: Center for the Study of Evaluation, California State University. (ERIC Reproduction Service No. 204 411)

Marso, R. N., & Pigge, F. L. (1988). Ohio secondary teachers' testing

needs and proficiencies: Assessments by teachers, supervisors, and principals. *American Secondary Education*, 17, 2-9.

McKee, B. G., & Manning-Curtis, C. (1982, March). Teacher-constructed classroom tests: The stepchild of measurement research. Paper presented at the National Council on Measurement in Education annual conference, New York.

Newman, D. C., & Stallings, W. M. (1982). Teacher competency in classroom testing, measurement preparation, and classroom testing practices. Paper presented at the American Educational Research Association annual meeting. New York. (ERIC Reproduction Service No. ED 220 491)

Reeves, C. K., & Kazelskis, R. (1985). Concerns of preservice and inservice teachers. *Journal of Educational Research*, 78, 267-271.

Ruddell, R. B. (1985). Knowledge and attitudes toward testing: Field educators and legislators. *Reading Teacher*, 38, 538-543.

Stetz, F. P., & Beck, M. D. (1979). Comments from the classroom: Teachers' and students' opinions of achievement tests. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco.

SUBSCRIBE



THE JOURNAL OF EDUCATIONAL RESEARCH

ORDER FORM

☐ YES! I would like to order a one-year subscription to **The Journal of Educational Research**, published bimonthly. I understand payment can be made to Heldref Publications or charged to my VISA/MasterCard (circle one).

☐ \$33.00 individuals ☐ \$66.00 institutions

ACCOUNT# _____ EXPIRATION DATE _____

SIGNATURE _____

NAME/INSTITUTION _____

ADDRESS _____

CITY/STATE/ZIP _____

COUNTRY _____

ADD \$12.00 FOR POSTAGE OUTSIDE THE U.S. ALLOW 6 WEEKS FOR DELIVERY OF FIRST ISSUE.

SEND ORDER FORM AND PAYMENT TO:

HELDREF PUBLICATIONS, THE JOURNAL OF EDUCATIONAL RESEARCH

1319 EIGHTEENTH STREET, NW, WASHINGTON, DC 20036-1802

PHONE (202) 296-6267 FAX (202) 296-5149

SUBSCRIPTION ORDERS 1 (800) 365-9753

Since 1920, The Journal of Educational Research has contributed to the advancement of educational practice in elementary and secondary schools. Authors experiment with new procedures, evaluate traditional practices, replicate previous research for validation, and perform other work central to understanding and improving the education of today's students and teachers. The Journal of Educational Research is a valuable resource for teachers, counselors, supervisors, administrators, planners, and educational researchers.