DOCUMENT RESUME

ED 243 882 SP 024 372

AUTHOR Cornett, Lynn M.
TITLE A Comparison of Teacher Certification Test Scores and Performance Evaluations for Graduates in Teacher Education and in Arts and Sciences in Three Southern States. Final Report.
INSTITUTION Southern Regional Education Board, Atlanta, Ga.
SPONS AGENCY National Endowment for the Humanities (NFAH), Washington, D.C.
PUB DATE 84
GRANT OP-20102-83
NOTE 59p.
PUB TYPE Reports - Research Technical (143)

EDRS PRICE MF01 PC03 Plus Postage.
DESCRIPTORS Comparative Analysis; *Educational Certificates; Elementary School Teachers; Elementary Secondary Education; Employment Qualifications; Humanistic Education; *Liberal Arts; *Scores; Secondary School Teachers; Standardized Tests; State Norms; *Teacher Certification; Teacher Education Programs; Teacher Effectiveness; Teacher Evaluation; *Vocational Evaluation
IDENTIFIERS Georgia; Louisiana; North Carolina; *Provisional Certification

ABSTRACT

The regular certification of arts and science graduates who have not completed the necessary professional education courses has become state policy recently in several states and others are considering it. These four studies provide data on two research questions related to this policy: (1) What, if any, differences in scores on teacher certification tests occur for two groups of teachers: those who have completed teacher education or certification programs, and those who are temporarily certified (arts and sciences majors)? and (2) What, if any, differences exist in on-the-job performance of the aforementioned groups of teachers as measured by evaluation instruments currently in use? The methodology, results, and conclusions of four related studies are presented: (1) a study of scores on the Georgia Teacher Certification Tests for teachers with regular certification and provisional certification; (2) a study of scores on the National Teacher Examinations for teachers with regular and temporary certification in Louisiana; (3) a comparison of the performance of classroom teachers in a metropolitan school district who have graduated from arts and sciences or teacher education programs; and (4) a comparison of performance and test scores for a selected population of teachers with provisional and regular certification in North Carolina. (JD)

# A COMPARISON OF TEACHER CERTIFICATION TEST SCORES AND PERFORMANCE EVALUATIONS FOR GRADUATES IN TEACHER EDUCATION AND IN ARTS AND SCIENCES IN THREE SOUTHERN STATES

Lynn M. Cornett

A final report submitted by the Southern Regional Education Board
to the National Endowment for the Humanities
in partial fulfillment of Grant OP-20102-83.



Southern Regional Education Board
1340 Spring Street, N.W.
Atlanta, Georgia 30309
1984

This report summarizes a study partially funded by the National Endowment for the Humanities Grant No. OP-20102-83. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the National Endowment for the Humanities.

# A COMPARISON OF TEACHER CERTIFICATION TEST SCORES AND PERFORMANCE EVALUATIONS FOR GRADUATES IN TEACHER EDUCATION AND IN ARTS AND SCIENCES IN THREE SOUTHERN STATES

## INTRODUCTION AND PURPOSE

There is a growing awareness and determination across the nation that the curriculum of the public schools must be strengthened, and that the caliber of teachers must be improved. Insuring that prospective teachers master general education components and maintaining the quality of content to be taught are increasingly important objectives. Mandates in numerous states now require that beginning teachers demonstrate minimum competencies on written tests that focus on basic skills and/or content to be taught.

The emphasis on employing teachers who know their subject matter has refocused attention on the employment of liberal arts graduates with preparation in a major subject field, even when a pedagogy sequence has not been completed. This policy is a reaction to the shortage of qualified teachers in certain fields, such as science and mathematics, as well as the recognition that persons with such preparation may have much to offer in teaching positions.

The SREB Task Force on Higher Education and the Schools, in its report The Need for Quality, (SREB, 1981) recommended that states modify their certification requirements to allow provisional certification of all beginning secondary teachers, including arts and sciences graduates, "with safeguards to insure the quality of instruction." The 1980 report of the Commission on the Humanities (Commission, 1980) calls for state departments of education to base certification requirements on solid liberal education of applicants. The Commission expresses concern that the minimum competencies now required in teacher testing may in fact become norms. The Commission also questions the prevailing system that for the most part excludes those who are academically prepared in various humanities disciplines but lack a required number of education courses.

As states consider the possibility of opening certification to arts and sciences majors, several questions arise. Will such majors be able to handle the classroom? Are these graduates sufficiently aware of differences in student abilities and cultural origins? Does knowledge of subject matter imply being able to teach that information at the level of high school students?

There is a need for data on the actual performance of arts and sciences graduates in the classroom, and on how they score on teacher certification tests of subject matter knowledge and pedagogy. The increasing pressure on state certification officers for revision of certification rules calls for data and research on this subject.

Several policy questions are related. If it is found that the classroom performance of beginning teachers who have not had a pedagogy sequence is similar to those who have, then the effectiveness of pedagogy sequences is open to question. If, on the other hand, the performance of arts and sciences graduates does not measure up to the performance of teacher education graduates, then the policy of certifying them should be questioned. Data on performance after the first or second year may indicate whether on-the-job experience does or does not make a difference in the assessed performance of the teacher. Results will need to be interpreted in the context of methods used to assess performance.

An examination of the test scores of arts and sciences graduates and teacher education graduates should indicate whether or not a difference in subject matter, general education, and professional knowledge does exist among those entering a teaching career. Test results may reflect preparation programs or self-selection by students into particular college programs. Research has not substantiated any relationship between the level of ability to score on a test and performance in the classroom. This may indicate the complexity of the classroom. Current teacher certification tests have cutoff scores set at minimum levels, hence their present purpose is to screen out those who cannot pass a minimum standard, and they do not necessarily predict who might or

might not perform well in the classroom. In either case, policy questions are being asked about the attractiveness of teacher education programs and of the teaching profession to talented individuals, and data for making decisions will be helpful.

Data are now available from several states which permit comparisons of groups of arts and sciences graduates with teacher education graduates in terms of their knowledge base, as measured by a standardized test, and by their performance in the classroom, as measured by systematic evaluation instruments.

Two research questions are posed for each set of data. First, what differences exist in scores on certification tests which measure knowledge of subject matter, general education, professional education, or a combination thereof, for employed teachers who have graduated with arts and sciences degrees and those who have graduated from arts and sciences programs? A second question is: What differences exist in on-the-job performance of teachers, as measured by performance assessment instruments, for employed teachers who have graduated from arts and sciences programs compared with those who have graduated from teacher education programs?

Several areas of research are related to the questions posed. What is the present status of certifying arts and sciences graduates? What is the relationship between a liberal education for a teacher and pedagogical knowledge? How do arts and sciences majors and teacher education graduates compare on tests of academic knowledge as well as on-the-job performance? In addition, information is sought about the instruments used to evaluate classroom performance as well as tests used for certification purposes. What is the relationship between how well a teacher scores on a test or is rated in the classroom and student achievement and attitudes in school?

## RELATED LITERATURE AND ANALYSIS

### Certification of Arts and Sciences Graduates

Suggestions or recommendations have been advanced in several states to provisionally certify arts and sciences graduates on the same basis as teacher education graduates. In the SREB region, only two states have moved to permit certification of arts and sciences graduates for secondary school positions. In Virginia, arts and sciences graduates may be provisionally certified and then must complete nine hours of approved courses or an alternate program approved by the state superintendent of education. In Florida, legislation has authorized the employment of persons with subject matter preparation but no professional education courses, to be implemented by districts under Board of Education guidelines. Nationally, California and New Hampshire have provisions for certifying arts and sciences graduates, and several other states are discussing proposals. Other states may give provisional certification to arts and sciences graduates, but in some of the states that have state salary schedules such teachers do not earn the same salary as those who have graduated from an approved program or have completed the necessary education hours. Teachers in these states have to earn the necessary education hours to attain regular certification.

### Teacher Education versus Liberal Arts Education for Teachers

The question of whether or not the education of teachers should include courses on pedagogy is not a new dimension in examining the quality of teachers. The influence different types of training programs have on the effectiveness of a teacher were summarized in the 1969 edition of the Encyclopedia of Educational Research, which discussed several studies that focused on the same question during the early Sixties. Results were mixed but the authors concluded that the "results are not very encouraging"; several studies indicated that having gone through teacher training programs did make

a difference and others reported that training to be a teacher made no difference except that teachers who were trained had greater objectivity.

Present arguments appear to center around the inability of schools of education to attract large numbers of bright students and on whether or not students studying to be teachers should concentrate their work in academic areas as opposed to courses in pedagogy. According to a recent summary of the question in Educational Testing Services' Teacher Competence, (1982) the National Accrediting Association for Colleges of Education (NCATE) notes that there has been an erosion in the amount of time devoted to pedagogy coursework, with an increase in arts and sciences courses. The intellectual content of the education programs has been attacked by J. Myron Atkin, dean of the College of Education of Stanford University. Others argue that teachers must learn by doing and that teacher education cannot be fully taught in the college classroom.

## Performance Evaluation

Studies examining the on-the-job performance of teachers have produced mixed results, but generally have found that the teacher who has regular certification or has completed a teacher education program performed better in the classroom according to ratings from superiors.

When effectiveness of teachers was measured by a rating scale administered by school principals, provisionally certified elementary teachers did less well than regularly certified teachers in planning and preparation, subject area knowledge, student evaluation, and pupil-teacher relations. No significant differences were found in the following areas: instruction (ability to motivate students, use instructional materials), parent-teacher relations, and human relations with colleagues in the school and community (LuPone, 1961).

Beery's (1962) research in elementary and secondary schools in Florida found that first-year teachers who had completed sequences in education courses were rated

significantly higher than those who had not. The ratings were made by education personnel as well as laypersons from outside the school. When principals' ratings only were used to compare teachers in the same school, the mean ratings favored the fully certified teachers, but not at a statistically significant level.

Hall (1964) completed a study in Florida of first-year elementary teachers and found that student achievement gains were significantly related to hours of education courses that had been completed by teachers. Student achievement scores favored the fully certified teachers, especially in word meaning, paragraph meaning, and language.

In a study conducted in Georgia in 1967, teachers who had regular certification were compared to those who held temporary certification because they had not completed the necessary number of education hours. Using an instrument similar to the one used by Beery, the regular teachers were found more systematic and responsible, more skilled in use of teaching media, and generally more competent (Bledsoe et al., 1967).

A 1971 study by Popham compared experienced teachers with persons from outside the schools, such as an electrician, in their ability to teach high school students a unit lasting from 4 to 9 hours. The achievement of the students was measured through post-testing after the material had been presented. The students of the experienced teachers scored higher, but differences were not significant. A study by Bausell and Moody (1972) reported that students taught by inexperienced student teachers learned as much as those taught by experienced teachers.

Copley, in his 1975 study of ratings given to teachers by principals, found that in areas such as communication skills and consideration of pupils, teacher education graduates were rated higher than those who had no student teaching or those who possessed an arts and sciences degree, but no difference was found in planning and organization, knowledge of subject matter, or personal characteristics.

Test Results for Teacher Education and Other Majors

Students who intend to major in education have historically ranked near the bottom third for all majors on achievement and aptitude testing, and the situation is getting worse. Declines in English and mathematics scores between 1970 and 1976 were shown to be greater for prospective education majors than for all other majors (Weaver, 1979). Scholastic Aptitude Test (SAT) scores for 1982 indicate that the average verbal SAT score ranked 32 points below and mathematics scores averaged 48 points below the national averages. Education ranked 26th out of 29 majors; only home economics, ethnic studies, and vocational studies ranked lower.

Recent information from the National Center for Education Statistics (NCES, 1982) indicates that for high school graduates of the class of 1980, those intending to major in education had lower achievement scores and lower self-reported grade averages, had taken fewer mathematics and science courses, and a smaller proportion of them had been in academic programs in high school, compared to students declaring all other majors.

In a study of 1979 graduates of public and private colleges in Virginia, it was found that mean SAT scores of graduates certified to teach were lower than those not certified to teach. However, it was noted that the difference in scores by institution was greater than whether or not graduates had the qualifications to be certified. The researchers do indicate that conclusions drawn from the evidence and general findings should be viewed with the knowledge of the aggregation of data and nature of the SAT (State Council of Higher Education for Virginia, 1981).

Results of the Graduate Record Examination (GRE) scores show similar trends. Verbal scores of intended education majors dropped 22 points from 1974 to 1981; mathematics scores dropped 7 points. For all candidates the drop on verbal was 20 points, but mathematics scores increased by 13 points from 1974 to 1981.

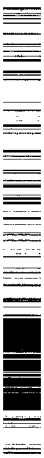
intended education majors also ranked near the bottom in comparison with all other intended majors (Hardy, 1982).

In addition, it is pointed out that schools have been just as willing to hire low scoring teachers as high scoring teachers (Weaver, 1979; Vance and Schlechty, 1982).

## Teacher Testing and Its Relation to Performance Evaluation

Several studies have examined the relationship between the National Teacher Examinations (NTE) and performance in the classroom. For the most part, studies have been conducted using observation instruments that principals or supervisors employed to assess the performance of student teachers or teachers. In general, low correlations have been observed (Quirk, Witten, and Weinberg, 1973). A study of the relationship between the Common subtest scores of the NTE and teaching styles found correlations with a median of .25 (Medley and Hill, 1970). Another study which examined Weighted Common scores, selected Area Examinations, and supervising teachers' ratings reported signif ant correlations, both positive and negative. The authors questioned the rating instruments that were used (Andrews, Blackmon, Davidson and Mackey, 1982). A study of performance of student teachers' assessments by university supervisors in relation to the NTE elementary education test reported a significant correlation of .43 (Piper and Sullivan, 1981). Correlations of the NTE area tests in mental retardation, early childhood, and physical education as well as the professional education subtest of the Common test and the Common score were correlated with on-the-job performance as measured by the Georgia Teacher Assessment Instrument (TPAI). Correlations ranged from -.12 to .52 (SREB, 1982). In general, the findings of the studies support the contention of Educational Testing Service that knowledge is only one part of the complexity of classroom teaching (ETS, 1978).

A recent study examined the relationship of the scores on the Georgia Teacher Certifications Test with the results of the Teacher Performance Assessment Instrument



data for beginning teachers and found that for the first assessments in the fall, correlations tended to be positive, but very small, ranging from .09 to .35 for academic subjects. The subjects of art, home economics, and music all showed very small negative correlations ranging from -.03 to -.15 (SREB 1982).

Teacher Testing, Performance Evaluation, and Student Achievement

There has been little research on the relationship of performance assessment or testing of teachers with achievement and attitudes of students. Mixed results were obtained in a study which examined student achievement and attitudes as they related to low inference teaching measures (low inference instruments measure the presence or absence of teacher behaviors in a classroom) (Coker, Medley, Soar, 1980). A study examining the relationship of the Georgia Teacher Performance Assessment Instrument with student achievement found a significant relationship between the TPAI ratings and teacher-made tests. When student achievement was measured by standardized tests, the results were mixed (Capie, 1981).

Some studies report relationships between the National Teacher Examinations and student achievement. However, many of the studies are beset with methodological problems. An early study by Lins (1946) showed moderate correlations between student achievement and the NTE Common Examinations, however, small samples were used. A study of the relationship between the National Teacher Examinations Common test and student achievement in mathematics and vocabulary found significant relationships (Sheehan and Marcus, 1978). However, they cite another study showing an inverse relationship between teachers' NTE scores and pupil achievement.

Teacher Performance Instruments

Questions arise as to the validity and reliability of teacher observation instruments used by states and districts to evaluate teachers on the job.

Most of the instrumentation appears to have sufficient reliability and validity to support use. However, it might be noted that most reliability data concerns interrater agreement, which provides information on the use of the instrument, the clarity of the language, and training of the observers. According to Medley and Mitzell (1963) this coefficient of observer agreement tells us something about the objectivity of the observations, but nothing about how closely the obtained score approximates a true score (the day-to-day behavior of the teachers as opposed to what is observed on one occasion). Other reliability of the instrument can be low, even if two observers agree exactly, because of errors that arise from changes in behavior of the teacher from one occasion to another. Dickson and Wiersma (1980) note that teacher observation depends highly on the instrument in use.

Instruments currently in use are generally developed from a "consensus" model, from a research perspective, or from a combination of both. In a consensus model, competencies that practitioners "believe" to be associated with effective practice are assembled. However, what is thought to be characteristically effective practice may not be the case when the behaviors are correlated to student achievements and attitudes, as reported in the works of several authors (Rosenshine, 1976; Medley, 1977; Coker, Medley & Soar, 1980). A second method of developing evaluation systems is to use the current research base, which links particular teacher behaviors with student outcomes. Most of the research in this area has been conducted in elementary schools with low socioeconomic students, therefore, teacher behaviors linked with student achievement may not be the same as behaviors linked to achievement of secondary or high achieving students. A recent study has questioned the generic character of the observation instruments in use (Southern Regional Education Board, 1982).

Research on assessment instrumentation is sparse, especially that which examines specific teacher competencies associated with student achievement and attitudes across a variety of ages and subject areas. According to a recent critique of teacher evaluation, methods used to evaluate teachers are inadequate. Soar, Medley and Coker (1983) criticize rating scales to evaluate teachers because of their inaccuracy in measuring the performance of teachers, their lack of validity, and because they are susceptible to the "halo effect." They report a number of studies to support their contention that ratings of teachers' performance by trained observers have "no validity as predictors of teacher effectiveness" (p. 244) when the criteria is student learning.

Questions also arise concerning who should be doing the observations in the classroom--principals, peers from within the school, or peers from outside the school. Studies have shown that there are not substantive overall differences in ratings related to the type of observer (Hawkins & Stoops, 1966; Hamm, Liu, Brinlee, 1983). However, it has been noted that differences occurred in ratings that teachers, district administrators, and peer teachers gave on classroom management and communication on an instrument used statewide. District administrators scored teachers significantly lower than teachers or principals on classroom management, but on communication, principals scored teachers higher than district adminstrators and peer teachers (Hamm, et al., 1983).

# METHODOLOGY, RESULTS, AND CONCLUSIONS OF FOUR RELATED STUDIES

## A Study of Scores on the Georgia Teacher Certification Tests for Teachers with Regular Certification and Teachers with Provisional Certification

This is an ex post facto study examining the differences between the scores on the Georgia Teacher Certification Tests (TCT). The study assesses teaching field knowledge for two groups of teachers: (1) those who have completed a regular teacher education program leading to a bachelor's or master's degree, and (2) those who have completed an arts and sciences program leading to a bachelor's or master's degree and do not possess enough hours in professional education courses to be regularly certified. The variables of the study were:

1. The scaled scores on the Georgia Teacher Certification Test.
2. Area in which the teacher took the Teacher Certification Test.
3. Type and level of program completed by the teacher: approved teacher education program at the bachelor's or master's level or an arts and sciences degree at the bachelor's or master's level, as ascertained from type and level of certification.

### Populations

The population for the study was composed of teachers who were employed in public school systems in the state of Georgia for the 1982-83 school year, had taken the Georgia Teacher Certification Test during the three testings in 1981-82 or during the three testings in 1982-83, graduated from approved teacher education programs and possessed professional certificates (NT-4 T-4, NT-5, T-5), or graduated with arts and sciences degrees and possessed provisional certificates (B-4, B-5) as their only certification. (See Appendix B for information on certification.) All teachers had three or fewer years of experience teaching in Georgia.

## Data Collection

The computer data base of the State of Georgia Department of Education was the data base for the project. All data were treated with utmost security. Group data were used with no identification of individuals at any time.

## Instrumentation

Georgia Teacher Ce. .ification Tests. The Georgia tests are criterion-referenced tests that assess an individual's knowledge of content in his or her teaching field. The tests were developed as part of the performance-based certification policy of the State of Georgia. Cutoff scores and minimal performance standards were set in 1977-1978 by the Department of Education; cutoff scores are two and one-half standard errors of measurement below the determined minimal level. The tests were designed to reflect the curriculum in Georgia public schools. A large number of teachers reviewed objectives of the tests in order to maximize the degree of content validity. The reliability data are expressed as Kuder-Richardson 20 coefficients. The coefficients on the tests range from 0.85 to 0.94. The final scaled score is an adjustment of raw scores so that the same cutoff score of 70 on all of the tests corresponds to different percentage-correct values, depending on the test taken (Georgia State Board of Education, n.d.). A study, which examined passing rates of graduates of the University System of Georgia the first time that the tests were taken, found passing rates based on the cutoff score of 70 varied for different subject areas. Passing rates ranged from 55 percent on home economics to 97 percent on music education. These differences may indicate a difference in difficulty of tests or differences in populations taking the tests. Since passing rates are for first-time test-takers it is not known if differences continue to occur after retakes on the tests (SREB, 1982).

Since data for the study were obtained from records of currently employed teachers who hold a certificate, scores are all at or above the passing scaled score of 70. The scores recorded represent either the first attempt at passing the test, if an acceptable score was achieved, or may represent the last score, if a score of 70 was achieved in a series of multiple attempts.

Analyses of the Data

To determine if differences in knowledge level of subject matter (as measured by the Georgia Teacher Certification Test) exist, a statistical summary was made for four groups: graduates of approved teacher education programs at the (1) bachelor's and (2) master's levels, and arts and sciences graduates at the (3) bachelor's and (4) master's levels. In addition, test scores were broken down into 5 ranges, and numbers and percentages of teachers falling into each range were calculated for the programs (teacher education, arts and sciences) as well as the levels (bachelor's, master's). The data were also examined to determine differences by program as well as level, with interactions included. The data were analyzed to determine if differences occurred by testing year. Data for the groups were analyzed separately by test areas having sufficient numbers of teachers. A grouping was made of all teachers, which included: communicative arts, social studies, mathematics, science, French, Latin, Spanish, German, business, distributive education, agriculture, home economics, and industrial arts. Another grouping was made for humanities which included: communicative arts, social studies, and all languages. All calculations were completed using the Statistical Package for the Social Sciences (SPSS) (Nie, Hull, Jenkins, Steinbrenner & Bent, 1975).

## Results

The analysis of the data by testing year (1981-82, 1982-83) reveals that the results are comparable and that few differences exist in the data from year to year (see Table 1). This analysis was completed to test the use of only two years of testing data for the study.

TABLE 1

Mean Scores on the
Teacher Certification Test by Year

| | N | 1981-82 | N | 1982-83 |
|---|---|---|---|---|
| All* | 480 | 79.2 | 223 | 79.1 |
| Humanities** | 259 | 80.0 | 116 | 81.4 |
| Communicative Arts | 99 | 78.7 | 37 | 78.2 |
| Social Studies | 143 | 80.7 | 71 | 81.7 |
| Mathematics | 55 | 79.6 | 40 | 78.6 |
| Science | 54 | 79.0 | 30 | 77.3 |

* Includes the total population with the following subject areas represented: communicative arts, social studies, mathematics, science, French, Latin, Spanish, German, business, distributive education, agriculture, home economics, industrial arts.

** Humanities includes communicative arts, social studies, and all languages (Spanish, French, German, Latin).

NOTE: Subject breakdowns are shown only for areas with sufficient numbers for analysis.

The analyses of the data reveal that for employed teachers with fewer than 3 years of experience in Georgia schools the knowledge level as measured by the Georgia Teacher Certification Tests is more dependent on level of degree than on program (see Tables 2 and 3). In general, the bachelor's level teachers scored lower than the master's level teachers by around 3.8 points. On the other hand, arts and sciences graduates scored slightly higher than teacher education graduates overall, but the differences are not consistent across fields. Differences by levels were more apparent in

TABLE 2

Georgia Teacher Certification Test
Scores for a Selected Population by Program/Level

| | N | Grand Mean | Discrepancies from the Grand Mean | | | |
|---|---|---|---|---|---|---|
| | | | Arts & Sciences | Teacher Education | Bachelor's | Master's |
| All* | 703 | 79.2 | .4 | -.2 | -.4 | 3.4 |
| Humanities** | 375 | 80.4 | 1.4 | -.9 | -.5 | 3.1 |
| Communicative Arts | 136 | 78.6 | 1.4 | -.8 | -.6 | 4.5 |
| Social Studies | 214 | 81.0 | .7 | -.5 | -.3 | 1.9 |
| Mathematics | 95 | 79.1 | -.4 | .4 | -.4 | 5.5 |
| Science | 84 | 78.4 | -.5 | .4 | 0.1 | -2.0 |

* Includes the total population with the following subject areas represented: communicative arts, social studies, mathematics, science, French, Latin, Spanish, German, business, distributive education, agriculture, home economics, industrial arts

** Humanities includes communicative arts, social studies, and all languages (Spanish, French, German, Latin).

NOTE: Subject breakdowns are shown only for areas with sufficient numbers for analysis.

some particular fields, with greatest differences for communicative arts and mathematics. Science teachers at the bachelor's level, however, had higher scores than those at the master's. Differences by program were most marked for humanities and communicative arts, with graduates from an arts and sciences program scoring slightly over two points higher than teacher education graduates.

Differences within level indicate that at the bachelor's level, the arts and sciences groups scored slightly higher than the teacher education groups, with the exception of the mathematics and science teachers. (See Appendix A for complete data.) However, at the master's level the arts and sciences group scored lower than the teacher education

TABLE 3

Scores on Georgia Teacher Certification Tests by Levels for Graduates of Arts and Sciences and Teacher Education Programs

| | All* | | Humanities** | | Communicative | | Social Studies | | Mathematics | | Science | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Arts & Sciences | Teacher Education | Arts & Sciences | Teacher Education | Arts & Sciences | Teacher Education | Arts & Sciences | Teacher Education | Arts & Sciences | Teacher Education | Arts & Sciences | Teacher Education |
| Total=703 | | | | | | | | | | | | |
| Number | 267 | 436 | 147 | 228 | 48 | 88 | 90 | 124 | 40 | 55 | 36 | 48 |
| Bachelor's | | | | | | | | | | | | |
| Number | 230 | 399 | 119 | 203 | 42 | 79 | 70 | 112 | 36 | 52 | 34 | 44 |
| Mean Score | 79.3 | 78.6 | 81.5 | 78.9 | 79.1 | 77.4 | 82.0 | 79.9 | 78.3 | 78.9 | 78.2 | 78.8 |
| Master's | | | | | | | | | | | | |
| Number | 37 | 37 | 28 | 25 | 6 | 9 | 20 | 12 | 4 | 3 | 2 | 4 |
| Mean Score | 81.3 | 83.8 | 82.4 | 84.8 | 86.3 | 80.9 | 81.1 | 85.9 | 81.3 | 88.9 | 71.7 | 78.8 |

* Includes the total population with the following subject areas represented: communicative arts, social studies, mathematics, science, French, Latin, Spanish, German, business, distributive education, agriculture, home economics, industrial arts.

** Humanities includes communicative arts, social studies, and all languages (Spanish, French, German, Latin).

group. This holds true for all groups except communicative arts. In science and mathematics, teacher education graduates outscored the arts and sciences teachers at both levels. For communicative arts, the arts and sciences group outscored the teacher education group at both levels. The small numbers at the master's level must be considered when examining data, particularly in science, mathematics, and communicative arts categories.

Conclusion

The results for all teachers in this selected population, as well as breakdowns by subject matter, indicate that subject matter and, to a slight degree, professional knowledge, as measured by the Georgia TCT, are more dependent on level than on program. A much higher percentage of the teachers who have completed a bachelor's degree fall into the bottom of the score ranges than do the teachers with a master's degree. Similarly, almost twice as many master's level teachers fall into the upper ranges when compared to the bachelor's level teachers.

To examine the distribution of scores by program and level, total score ranges (from 70% to 99.9%) were divided into fifths (see Table 4). By levels for all teachers, approximately 45 percent of the master's teachers fell into the bottom two-fifths of the total range, with 70 percent of the bachelor's level teachers in the bottom two-fifths. Ten percent of the bachelor's level teachers fell in the upper two-fifths; approximately 20 percent of the master's graduates were in that range.

By program, the distributions were more similar; 66 percent of the arts and sciences graduates and 68 percent of the teacher education graduates were in the lower two-fifths. Approximately 12 percent of the arts and sciences graduates fell into the upper two-fifths of the range; 9 percent of the teacher education graduates were in those categories.

TABLE 4

Percent of Teachers Scoring
at Various Levels of Georgia Teacher Certification Test

| | Test Scores | | | | |
|---|---|---|---|---|---|
| | 94+% | 88% - 93.9% | 82% - 87.9% | 76% - 81.9% | 75.9% or less |
| All | | | | | |
| Arts & Sciences | 2.2(6)* | 9.7(26) | 22.1(59) | 34.8(93) | 31.1(83) |
| Teacher Education | 1.8(8) | 7.3(32) | 22.7(99) | 30.0(131) | 38.1(166) |
| Bachelor's | 1.6(10) | 7.3(46) | 21.1(133) | 33.1(208) | 36.9(232) |
| Master's | 5.4(4) | 16.2(12) | 33.8(25) | 21.6(16) | 23.0(17) |
| Humanities | | | | | |
| Arts & Sciences | 1.4(2) | 13.6(20) | 27.9(41) | 34.0(50) | 23.1(34) |
| Teacher Education | 3.1(7) | 7.0(16) | 27.2(62) | 27.2(62) | 35.5(81) |
| Bachelor's | 1.9(6) | 8.4(27) | 25.4(82) | 31.1(100) | 33.2(107) |
| Master's | 5.7(3) | 17.0(9) | 39.6(21) | 22.6(12) | 15.1(8) |
| Communicative Arts | | | | | |
| Arts & Sciences | --- | 4.2(2) | 30.3(15) | 37.5(18) | 27.1(13) |
| Teacher Education | --- | 1.1(1) | 21.6(19) | 36.4(32) | 40.9(36) |
| Bachelor's | --- | 1.7(2) | 21.5(26) | 38.0(46) | 38.8(47) |
| Master's | --- | 6.7(1) | 53.3(8) | 26.7(4) | 13.3(2) |
| Social Studies | | | | | |
| Arts & Sciences | 2.2(2) | 17.6(16) | 26.4(24) | 33.0(30) | 20.9(18) |
| Teacher Education | 4.8(6) | 8.9(11) | 32.3(40) | 22.6(28) | 30.5(39) |
| Bachelor's | 3.3(6) | 12.0(22) | 28.6(52) | 28.0(51) | 28.0(51) |
| Master's | 6.3(2) | 15.6(5) | 37.5(12) | 21.9(7) | 18.8(6) |
| Mathematics | | | | | |
| Arts & Sciences | 5.0(2) | 5.0(2) | 17.5(7) | 32.5(13) | 40.0(16) |
| Teacher Education | 0 | 14.5(8) | 20.0(11) | 30.9(17) | 34.5(19) |
| Bachelor's | 1.1(1) | 10.2(9) | 17.0(15) | 33.0(29) | 38.6(34) |
| Master's | 14.3(1) | 14.3(1) | 42.9(3) | 14.3(1) | 14.3(1) |
| Science | | | | | |
| Arts & Sciences | 2.8(1) | 5.6(2) | 8.3(3) | 38.9(14) | 44.4(16) |
| Teacher Education | 0 | 12.5(6) | 16.7(8) | 31.3(15) | 39.6(19) |
| Bachelor's | 1.3(1) | 9.0(7) | 14.1(11) | 37.2(29) | 38.5(30) |
| Master's | 0 | 16.7(1) | 0 | 0 | 83.3(5) |

*Numbers in parentheses indicate numbers of teachers

Due to rounding totals may not equal 100%.

For all teachers at the bachelor's level, arts and sciences majors score slightly above the teacher education graduates, but do not score as well at the master's level. If scores within programs are compared by level, the differences between the bachelor's and the master's level are much greater for the teacher education group. This may be an indication of selection processes that are taking place in the teaching profession, or the testing process itself. Since the TCT has been designed to test minimum competencies and its content is no more complex than the curriculum of the Georgia public schools, it probably does not test higher complexities in the subject area. However, teacher education graduates possessing a master's degree outscore all other groups for this population.

These conclusions are derived from analyses of data for teachers who have been hired for teaching positions, and who have passed the TCT with at least a scaled score of 70. It therefore does not reflect upon a total number of persons prepared in teacher education programs or arts and sciences programs. The data does support the generalization of the two years of test data for employed teachers to a larger group which has been tested over a period of time, and has similar characteristics to the selected population in terms of programs and experience.

## A Study of Scores on the National Teacher Examinations for Teachers with Regular and Temporary Certification in Louisiana

This is a study to examine the differences between the scores on the National Teacher Examinations (NTE)--both the Common Examinations, and selected Area Examinations for two groups of teachers: 1) those who have completed a regular teaching education program leading to a bachelor's or master's degree and have regular certification, and 2) those who have completed an arts and sciences program leading to a bachelor's or master's, without the necessary hours to be a fully certified teacher in Louisiana. The variables of the study were:

1. The scores on the Weighted Common Examinations of the National Teacher Examinations (NTE).
2. The scores on the Area Examinations of the National Teacher Examinations.
3. The number of hours of professional education completed.
4. The type of program completed.
5. The area of certification.

Population

The population was composed of all teachers who received certificates to be employed in the state of Louisiana from July 1982 to July 1983; all had met the minimum NTE scores necessary to be certified.

In Louisiana, teachers who have not graduated from an approved teacher education program or do not hold the necessary hours for certification may be given temporary one-year certificates upon request of the employing district. These certificates expire at the end of one year, but may be renewed by taking 6 hours of coursework leading to certification.

The total population of teachers who were issued temporary certificates during 1982-83, and held no other regular certification in the state, was used in the study (N = 89). In this group, the number of education hours taken ranged from 0 to 36, with an average of 9.5 hours. The average number of education hours for those taking the elementary education test was 13. It was determined that these hours had been obtained before taking the NTE tests. Eighty-three (83) of the teachers possessed bachelor's degrees; 6 held the master's degree. Of the group, 21 had taken the NTE Education in the Elementary School examination; 10 had taken the social studies test; and 9 had taken the English Language and Literature Examination. Degrees in psychology (13) and English (9) were the most Common. Twenty-six of these teachers held temporary certification in elementary education, with 22 in special education.

A random sample of 105 teachers with regular certification was drawn from the initial population. Twelve of this sample held master's degrees, the remaining 93 held bachelor's. The most common certification area was elementary education (42), followed by physical education (13), English (9), and special education (9).

## Data Collection

The computer data base of the state of Louisiana, as well as certification records, were used for data aggregation. All data were treated as group data, with no identification of individuals at any time.

## Instrumentation

National Teacher Examinations. The National Teacher Examinations (NTE) are composed of Area Examinations and Weighted Common Examinations (WCET). The Area Examinations test the content of a special field of major in undergraduate education. Scores range from 250 to 990. Area Examination scores cannot be compared across

areas, and even though scores appear to be similar, a particular score does not represent the same level of proficiency from test to test. The tests measure principles and concepts from teacher education programs (ETS, 1981). Scaled scores approximate interval data. Reliabilities are reported to range from .91 to .95 (Kuder-Richardson 20) with a standard error of measurement from 20 to 29 for most tests (ETS, 1978). The Education in the Elementary School Examination focuses on the nature of the child and the process of teaching in the elementary school. The nature of the child includes child development and individualization. The process of teaching includes classroom organization, diagnoses of student needs, and evaluation.

The Weighted Common Examinations (WCET) consist of a test in professional education and one in general education. The latter has 3 subparts--written English expression; social studies, literature, and the fine arts; and science and mathematics. The WCET total score is a combination of the above tests with the following weighting: professional education--4.0, social studies--2.5, written English expression--1.0, and science and mathematics--2.5. The coefficient of reliability (Kuder-Richardson 20) for the WCET is reported to be .96, with a standard error of measurement of 21. A combination of the WCET and the Area Examination produces a composite score with ranges from 500 to 1980 (ETS, 1978). Tests have been validated against content of teacher education programs by the state department of education.

Analyses

Statistical summaries of data were obtained on the WCET scores for all teachers in the two groups, and for teachers with no education hours. Analyses of the WCET, Area Examination scores, and composite scores were completed for the teachers who had taken the Education in the Elementary School Area Examination. Because of small numbers, no other subject area groups were analyzed.

Results

WCET mean test scores for the two groups of teachers varied--the temporarily certified group as a whole scored higher (619) than the regularly certified teachers (602); the temporarily certified teachers who had taken no professional education courses also scored higher (611) (see Table 5). The same pattern held true for the teachers taking the elementary education test (630 versus 590). For the mean scores on the elementary education area test the reverse was true; temporarily certified teachers scored 23 points lower than the regularly certified, however, the temporarily certified teachers received composite scores 17 points above the regularly certified teachers.

TABLE 5

Scores on the NTE Weighted Common Examinations (WCET) and the Education in the Elementary School Area Test for Louisiana Teachers holding Regular and Temporary Certificates

| | Temporary Certificates with 0 Education Hours | All Temporary | Regular |
|---|---|---|---|
| All Teachers | | | |
| Number | 35 | 89 | 105 |
| WCET Scores | | | |
| Mean | 611 | 619 | 602 |
| Standard Deviation | 54.8 | 56.2 | 58.0 |
| Teachers taking Elementary Education Test (NTE) | | | |
| Number | | 21 | 42 |
| Area Test Scores | | | |
| Mean | | 628 | 651 |
| Standard deviation | | 46.8 | 42.2 |
| WCET Scores | | | |
| Mean | | 630 | 590 |
| Standard deviation | | 52.8 | 51.8 |
| Composite Scores | | | |
| Mean | | 1258 | 1241 |
| Standard deviation | | 87.8 | 84.3 |

## Conclusion

The teachers in this population who had received little or no training in professional education appeared to be able to score better on the WCET than those who had completed a regular sequence of education courses, even though 40 percent of the WCET is weighted for professional education content. Temporarily certified teachers with no education hours outscored the teacher education group. General education or specialized content courses may make up for the lack of education courses, or possibly, there were undetermined differences between the groups taking the test. It is interesting, however, to note that for the more specialized information that is measured on the elementary education area examination, the teachers who had not completed a teacher education program (an average of 13 hours of education courses) did not score as high as those who had.

## A Comparison of the Performance of Classroom Teachers in a Metropolitan School District who have Graduated from Arts and Sciences or Teacher Education Programs

This is an examination of the performance of classroom teachers, as measured by a locally designed teacher evaluation system, for two groups: 1) those who have regular certification in the state of Georgia (having completed a teacher education program or the necessary hours for certification), and 2) those who hold provisional certification because of a lack of the necessary hours in professional education courses. The variables for the study were:

1. Scores on the district teacher evaluation instrument.
2. Type of certification held by the teacher: regular (the teacher had completed a teacher education program or the necessary hours of professional education) or provisional (arts and sciences graduates who had not taken the necessary number of education hours to be regularly certified).
3. Teaching assignment.
4. Years of experience as a classroom teacher.

Population

All teachers who were graduates of arts and sciences programs (provisional certification) in the district during the 1982-83 school year (no regular certificates in any area--N = 21) were included in the study. Eighteen of these were secondary teachers, three were teachers at the elementary level. Of these, 11 were first-year teachers; the overall average was 2.3 years experience. A random sample of all regularly certified teachers was drawn (N = 27) for comparison. Of these, 13 were secondary and 14 were elementary teachers. In this group, only one was a first-year teacher. The group had an average experience level of 7.3 years. Because of the differences in teaching levels

and years of experience, a second sample was drawn. It was matched as closely as possible by subject area and by years of experience. Since no first-year teachers with regular certification are assessed in the district because of the statewide evaluation in Georgia of beginning teachers, it was necessary to move the experience level up one year for the sample. That is, 11 teachers at the second year of experience were used, and so on. The experience level for the second group was 5.2 years; subject area and level taught were matched.

Data Collection

All data was obtained from school district records, with no identification of any individual at any time. All data were grouped for reporting purposes.

Instrumentation

The evaluation instrument, Teacher Performance Observation Record, used in the district was developed from a statewide evaluation instrument for assessing beginning teachers in Georgia. (See Appendix C for a copy of the instrument.) Evaluations are completed by the principal in each teacher's school. All principals in the district have undergone training in use of the statewide instrument; this county assumes that the training carries over to the district instrument. The district instrument contains 10 categories or competencies with 33 indicators under the competencies. Scores range from 1 to 5 on each indicator; a level of 4 or 5 indicates a satisfactory level, depending on the indicator. The following are the competencies:

Plans Instruction

Uses Techniques, Methods, and Media Related to the Objectives

Communicates with Learners

Demonstrates a Variety of Teaching Methods

Reinforces and Encourages Learner Involvement in Instruction

Demonstrates an Understanding of the Subject

Organizes Time, Space, Materials, and Equipment for Instruction

Demonstrates Enthusiasm for Teaching, Learning, and the Subject

Helps Learners Develop Positive Self-Concepts

Manages Classroom Interactions

## Analyses

Statistical summaries for both groups (regularly certified, or provisionally certified) using the population data for the provisional group and the two sample groups with regular certification were completed. The small numbers in all groups must be taken into consideration when examining the results of the study.

## Results

The ranges of scores for the provisionally and regularly certified teachers were distributed as follows:

TABLE 6

Score Ranges on the Teacher Performance Observation Record for Provisionally and Regularly Certified Teachers

| Total Scores | Numbers of Teachers | | |
|---|---|---|---|
| | Provisional Certification | Regular Certification | |
| | | First Sample | Second Sample |
| 158-165 | 6 | 22 | 14 |
| 151-157 | 7 | 2 | 4 |
| 144-150 | 4 | 2 | 2 |
| 137-143 | 1 | 1 | 1 |
| Below 137 | 3 | 0 | 0 |

The mean score for the provisional teachers was 150, with a standard deviation of 17. The scores ranged from 90 to 165 of the possible 165. Classroom experience averaged 2.3 years. For those teachers holding regular certificates in the first sample, the mean

score was 160, with a standard deviation of 6. The scores ranged from 140 to 165. The teachers averaged 7.3 years of experience. The mean score for the second sample was 158, with a standard deviation of 8. The scores ranged from 140 to 165. The teachers averaged 5.2 years of experience.

Conclusion

The teachers with regular certificates had higher performance ratings than the provisionally certified ones; however, the former were considerably more experienced than the latter group. The scores of provisional teachers showed a broader distribution than those for the regularly certified teachers. Most of the regularly certified teachers received perfect or nearly perfect ratings on most competencies. Since a score of 4 or 5 is needed for a satisfactory rating on competency, the total scores indicate that a very small percentage of the teachers are receiving less than satisfactory ratings, especially the highly experienced teachers. On the other hand, a larger proportion of the provisionally certified teachers received lower ratings.

Whether or not the results indicate an essential difference in the groups or a difference due to experience levels cannot be concluded from the data. It does appear that the instrument may be used by principals to assess strengths and weaknesses to a greater degree for those teachers who are either less qualified according to certificates held or less experienced. The principals may tend to rate an experienced teacher high on all categories because of other factors.

All conclusions should be made with caution due to the small numbers included in this study.

## A Comparison of Performance and Test Scores for a Selected Population of Teachers with Provisional and Regular Certification in North Carolina

This study examined scores on teacher certification tests and on-the-job evaluations to determine if differences exist for two groups of teachers: 1) those with provisional certification and 2) those with regular certification. Teachers with provisional certification have completed arts and sciences programs at the baccalaureate level or have 90 hours leading to a degree. These teachers do not have the necessary hours of professional education for a regular certificate. Teachers with regular certification have a bachelor's degree in teacher education or a program which includes the necessary hours of professional education courses. The variables of the study were:

1. The scores on the National Teacher Examinations: Weighted Common Examinations (WCET), area test score, composite score (combination of the WCET and area test score).
2. On-the-job performance as measured by two types of evaluation instruments: a statewide instrument recently developed for the purpose of evaluating classroom teachers (North Carolina Teacher Performance Appraisal Instrument) and district-designed instruments used for the evaluation of teachers within a district.
3. The type of certification held by each teacher.
4. Experience level of each teacher at the time of the evaluation.
5. Degree level.

### Population and Samples

Evaluation of Teacher Performance. All teachers who were employed with provisional certification during the school years 1978-79 to 1982-83 were included in the selected population of provisionally certified teachers or the arts and sciences group (n = 191). Of this population, 31 of the teachers possessed less than a bachelor's degree. A random

sample of 348 teachers was selected from a population of 21,000 teachers who held regular certification at the bachelor's level and had been certified in North Carolina from 1978-79 through the 1982-83 school year.

For the population of provisionally certified and the sample of regularly certified teachers, on-the-job evaluations were requested by the North Carolina Department of Public Instruction from the districts. For the regularly certified teachers, the last district of employment was contacted for evaluations. For the provisionally certified teachers, the district where the teacher was employed was contacted. Of the original 539 names sent to districts, 292 evaluations were received. The return rate for the provisional teachers was 59 percent; for the regularly certified group it was 51 percent. Reasons for non-return of the evaluations included reluctance by districts to release the data and the non-availability of evaluations requested. For all teachers in the provisional group no upgrading of certification to regular status had been made, although the number of professional education courses or staff development hours taken since receiving the provisional certification was not known.

National Teacher Examinations. In addition, because the NTE data for 5 years were available for the provisionally certified teachers as well as regularly certified teachers, these data were used for the analyses of the scores on the National Teacher Examinations. Scores were not available for all members of the populations because some teachers had received certification before the date the NTE was required, or because Graduate Board Examination Scores were used in lieu of NTE scores.

## Data Collection

All data were obtained from the North Carolina Department of Public Instruction. No individuals were identified and all information was confidential.

Instrumentation

Performance Evaluation. Two general types of evaluations were used in this study to examine the performance of the two groups of teachers: the North Carolina Teacher Performance Appraisal Instrument and district-designed evaluation instruments used in local school systems. The statewide instrument was developed to carry out legislation of the North Carolina General Assembly which mandated performance standards and criteria to be used in evaluating all public school professional employees. The Teacher Performance Appraisal Instrument provides for evaluations to be made in three general areas: broad program functions, particular technical functions, and indirect facilitating functions. The instrument contains 33 basic elements of teaching which are rated by the evaluator. Evaluators are school principals, trained by the State Department of Public Instruction. No reliability information was obtained from the Department. The rating scale on the instrument provides 6 choices from "unsatisfactory performance" to "superior performance," with a "not applicable" category. (See Appendix D for a copy of the instrument.) However, because of local variations allowed, school districts used various combinations of the 6-point scale. For analytical purposes, the evaluations were grouped into four types--A, B, C, and D (see Appendix E). All used "meets performance standards or expectations" with categories above and below to indicate less than satisfactory or above satisfactory performances. To equate the instruments, the ratings on the statewide instrument were converted on each basic element to "meets standards" (=3), "below standards," (= 2), or "above standard expectations" (= 4). "Not applicable" ratings were scored as 0 and omitted from a final mean score for each teacher. A total mean score was obtained as well as mean scores for each of the three general areas. Mean scores ranged from 2.0 to 4.0. (See Appendix E for types and means of scoring systems.)

The district evaluations, unlike the statewide instrument, varied considerably. Some contained rating scales, others were written descriptions of the teacher's performance in and out of the classroom. Two independent raters subjectively rated the evaluation given the teacher by the principal on a 5-point scale: 1 = decidedly below average, 2 = below average performance, 3 = meets standards, 4 = above average performance, and 5 = decidedly above average. An 82 percent agreement was reached between the two raters. The 3-point scale similar to that applied for the statewide instrument was then used. All ratings of 1 became 2, and all ratings of 5 became 4 to convert to a 3-point scale.

National Teacher Examinations. Scores on the Weighted Commons Examinations (WCET), the Area Examinations, and the composite score were reported for the groups. (See the Louisiana study for a discussion of the instruments.)

## Analyses

All data were analyzed using the SPSS statistical package (Nie, et al., 1975). Statistical summaries and distribution of these scores by experience level as well as score ranges were calculated for both groups, and by type of instrument (statewide or district level). Statistical summaries of the test data were computed for the Weighted Common Examinations (WCET).

## Results

Evaluation of Teacher Performance. Of the total number of evaluations used in the study, half were the statewide instrument and the other half were the various district evaluations. An examination of the mean scores on the statewide instrument indicates some differences for this group of teachers, in terms of the type of scale which was adopted by the local district during this first year of implementation. The

largest difference occurs between the scale which uses 4 choices (Evaluation Type A), two of which are below satisfactory and one above, as compared to a scale of 5 choices (Evaluation Type C) of which two are below and two above satisfactory. However, the means by which the final score was tallied--using only scores of 2 for below average, 3 for satisfactory, and 4 for above satisfactory--could be an explanation for the results (see Appendix E).

The mean scores for all evaluations, for all teachers, did not differ for the two groups: those who had regular certification and those who possessed provisional certification after having completed an arts and sciences program or, in a limited number of cases, having no degree. Differences were not noted for breakdowns by years of experience at the time the teachers were evaluated. The same pattern was found regardless of type of evaluation instrument used (see Table 7). In examining scores for competency areas on the Teacher Performance Appraisal Instrument (such as planning, the technical functions in the classroom, or the indirect facilitating functions) little difference was noted.

For both district and statewide evaluations, there was no real difference in the distribution of scores between regular and provisionally certified teachers, nor between experienced and less experienced teachers. Very few teachers in either group score below the satisfactory levels. In fact, 31 percent and 43 percent of all teachers were rated at the highest levels under the district and statewide evaluations, respectively (see Table 8). An examination of the distribution of scores for all teachers reveals that less than 3 percent scored in the lowest range on the statewide instrument. Approximately 12 percent scored in the lowest range on the district evaluations. The numbers of first-year teachers evaluated on the statewide instrument is small and results should be interpreted with care.

TABLE 7

Mean Scores for On-the-Job Evaluations of Regularly and Provisionally Certified Teachers in North Carolina

| Years of Experience | Regular Certification | | | Provisional Certification | | |
|---|---|---|---|---|---|---|
| | Number | Mean Score | Standard Deviation | Number | Mean Score | Standard Deviation |
| **All Evaluations (N = 292)** | | | | | | |
| 1 year or less | 67 | 3.3 | .6 | 37 | 3.4 | .7 |
| 5 years or less | 139 | 3.3 | .5 | 73 | 3.4 | .6 |
| More than 5 years | 40 | 3.3 | .5 | 40 | 3.3 | .6 |
| Total | | 3.3 | .5 | | 3.3 | .6 |
| **District Evaluations (N = 146)** | | | | | | |
| 1 year or less | 52 | 3.4 | .7 | 31 | 3.5 | .7 |
| 5 years or less | 67 | 3.4 | .7 | 42 | 3.4 | .7 |
| More than 5 years | 14 | 3.4 | .7 | 23 | 3.3 | .8 |
| Total | | 3.4 | .7 | | 3.4 | .7 |
| **State Evaluations (N = 146)** | | | | | | |
| 1 year or less | 15 | 3.3 | .4 | 6 | 3.4 | .3 |
| subscore X* | | 3.2 | .4 | | 3.1 | .1 |
| subscore Y | | 3.3 | .4 | | 3.4 | .5 |
| subscore Z | | 3.3 | .4 | | 3.4 | .1 |
| 5 years or less | 72 | 3.3 | .4 | 31 | 3.3 | .4 |
| subscore X | | 3.2 | .4 | | 3.3 | .4 |
| subscore Y | | 3.3 | .4 | | 3.4 | .4 |
| subscore Z | | 3.3 | .4 | | 3.3 | .3 |
| More than 5 years | 26 | 3.3 | .4 | 17 | 3.3 | .3 |
| subscore X | | 3.3 | .4 | | 3.2 | .3 |
| subscore Y | | 3.4 | .4 | | 3.3 | .3 |
| subscore Z | | 3.3 | .4 | | 3.3 | .3 |
| Total | | 3.3 | .4 | | 3.3 | .4 |

* Subscore X: broad program functions which include operating and updating the instructional programs

Subscore Y: technical functions which refer to the means by which the teacher carries out daily instruction

Subscore Z: indirect facilitating functions which do not involve teachers or students

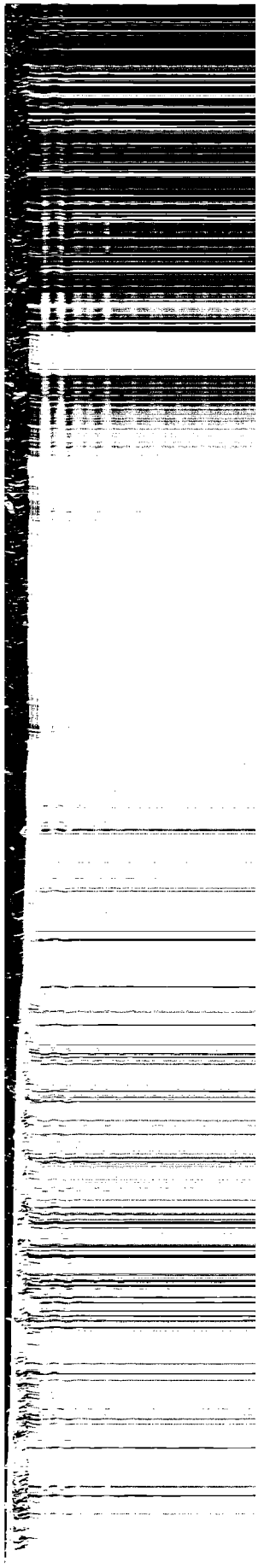


TABLE 8

Cross-tabulations of Mean Scores for On-the-Job Evaluations of North Carolina Teachers by Type of Certification

District Evaluations

| | First-Year Teachers (N = 83) | | | All Teachers (N = 146) | | |
|---|---|---|---|---|---|---|
| | Regular | Provisional | All | Regular | Provisional | All |
| Score=2 | | | | | | |
| Number | (5) | (4) | (9) | (8) | (10) | (18) |
| Row Percent | 55.6 | 44.4 | 10.8% | 44.4 | 55.6 | 12.3% |
| Column Percent | 9.6 | 12.9 | | 9.9 | 15.4 | |
| Total Percent | 6.0 | 4.8 | | 5.5 | 6.8 | |
| Score=3 | | | | | | |
| Number | (23) | (9) | (32) | (33) | (21) | (54) |
| Row Percent | 71.9 | 28.1 | 38.6% | 61.6 | 38.9 | 37.0% |
| Column Percent | 44.2 | 29.0 | | 40.7 | 32.3 | |
| Total Percent | 27.7 | 10.8 | | 22.6 | 14.4 | |
| Score=4 | | | | | | |
| Number | (24) | (18) | (42) | (40) | (34) | (74) |
| Row Percent | 57.1 | 42.9 | 50.6% | 54.1 | 45.9 | 50.7% |
| Column Percent | 46.2 | 58.1 | | 49.4 | 52.3 | |
| Total Percent | 28.9 | 21.7 | | 27.4 | 23.3 | |
| | (52) | (31) | (83) | (81) | (65) | (146) |
| | 62.7% | 37.3% | 100.0% | 55.5% | 44.5% | 100.0% |

Statewide Evaluation Instrument

| | First-Year Teachers (N = 21) | | | All Teachers (N = 21) | | |
|---|---|---|---|---|---|---|
| | Regular | Provisional | All | Regular | Provisional | All |
| Scores = 2.0-2.6 | | | | | | |
| Number | — | — | | (3) | (1) | (4) |
| Row Percent | — | — | | 60.0 | 40.0 | 2.7% |
| Column Percent | — | — | | 3.1 | 2.1 | |
| Total Percent | — | — | | 2.0 | 0.7 | |
| Score = 2.7-3.3 | | | | | | |
| Number | (10) | (2) | (12) | (57) | (23) | (80) |
| Row Percent | 83.8 | 16.7 | | 71.3 | 28.8 | 54.8% |
| Column Percent | 67.7 | 33.3 | 57.1% | 58.2 | 47.9 | |
| Total Percent | 47.6 | 9.5 | | 38.8 | 15.6 | |
| Score = 3.4-4.0 | | | | | | |
| Number | (5) | (4) | (9) | (38) | (24) | (62) |
| Row Percent | 55.6 | 44.4 | 42.9% | 61.3 | 38.7 | 42.5% |
| Column Percent | 33.3 | 66.7 | | 38.8 | 50.0 | |
| Total Percent | 23.8 | 19.0 | | 25.9 | 16.3 | |
| | (15) | (6) | (21) | (98) | (48) | (146) |
| | 71.4% | 28.6% | 100.0% | 67.1% | 32.9% | 100.0% |

ue to rounding totals may not add to 100.0 percent.

National Teacher Examinations. The Weighted Common Examinations (WCET) mean score for the provisionally certified (arts and sciences graduates) was slightly higher than for the regularly certified teachers or teacher education graduates (see Table 9). Breakdowns of the WCET scores by subject area were not available because of small numbers. Area Examination scores and Composite scores were not reported for the same reason.

TABLE 9

Results on the National Teacher Examinations (WCET) for a Selected Group of Regularly and Provisionally Certified Teachers in North Carolina

| | Regular Certificates | Provisional Certificates |
|---|---|---|
| Number | 15,644 | 118 |
| Mean Score | 580 | 582 |
| Standard Deviation | 78 | 96 |

## Conclusion

Evidence concerning the on-the-job performance of teachers as evaluated by school principals shows that there was very little difference between the teachers who had completed a teacher education program or the necessary hours for regular certification, and those with provisional certification. Additional training and numbers of professional education courses that had been completed by the provisionally certified teachers are not known. In terms of performance, as measured for beginning teachers with one year of experience compared to those with more experience, there were no differences. The instruments and methods of evaluating teachers appeared to place the majority of teachers in a very favorable light. The use of the statewide instrument for this group of teachers indicates that very few (less than 3 percent) of the teachers evaluated fell into a less than satisfactory category, based on a total score.

These results call into question whether or not these instruments are discriminating enough to reveal real differences that might occur between teachers, regardless of how they were prepared. Similarity in the means of a total score may not be revealing weaknesses in particular areas, although for the statewide instrument the breakdown by functions such as planning, classroom activities, and activities outside the classroom revealed no differences for the two groups. The purpose of the instruments is to identify weakness areas and to serve for decision-making within the districts, not for comparisons of groups of teachers.

On the other hand, if the instruments are valid for these purposes--and they are being used as the bases for employment decisions within districts--the ability of those persons who have not finished a traditional teacher education program or have not accumulated enough hours to become regularly certified to attain similar ratings on these types of instruments indicate that they appear to be functioning as well in the classroom as those who have completed the necessary requirements. At least weaknesses and strengths are identified to an equivalent degree through the use of the instruments.

The scores of the Weighted Common Examinations, which to a large extent deal with professional education, reveal that the arts and sciences graduates outscore the teacher education group to a slight degree; the scores should be interpreted as roughly equivalent. A breakdown of scores within the WCET needs to be examined to identify areas of weakness or strength for each group. Conclusions related to area test scores were not available because of small numbers of teachers in each subject area.

All interpretations of the data, both for teacher evaluations and for teacher test scores should be interpreted with the populations and samples in mind. Generalizations to any other groups should be made with caution.

## DISCUSSION OF THE FOUR RELATED STUDIES

The regular certification of arts and sciences graduates who have not completed the necessary professional education courses has become state policy recently in several states; other states are discussing the issue. These four studies provide data on two research questions related to this policy:

1. What, if any, differences in scores on teacher certification tests occur for two groups of teachers: those who have completed teacher education or certification programs and those who are temporarily certified (arts and sciences majors)?
2. What, if any, differences occur in on-the-job performance of such groups of teachers, as measured by evaluation instruments currently in use?

### Teacher Certification Tests

The results of these studies reflect data for employed teachers from selected populations and do not necessarily indicate how all graduates might perform on teacher certification tests. All teachers in the studies had received scores at or above the minimums needed for certification in each state.

Research has not established a relationship between scores on teacher certification tests and student achievement or attitudes in the classroom. Neither has it established what minimum level of knowledge is needed by an effective (as defined in terms of student achievement) teacher in the classroom. However, states across the nation, especially in the South, have mandated that teachers must pass minimum competency tests in general education, professional education, subject matter knowledge, or combinations thereof, to insure that teachers being granted certification possess a minimum level of knowledge to be able to function in the classroom.

In these studies, graduates of arts and sciences programs who had provisional or temporary certification generally outscored teacher education graduates in tests of general education and professional education (NTE Weighted Common Examinations),

despite the fact that 40 percent of the WCET score was weighted for professional education content. The arts and sciences graduates had completed varying amounts of professional education courses; however, the scores of the temporarily certified teachers in Louisiana who had not taken any professional education courses prior to taking the WCET attained higher scores than the teacher education graduates.

Test results on subject area examinations were compiled from the NTE Area Examinations and the data from the Georgia Teacher Certification Tests, which are heavily weighted toward subject matter content, although some professional education is included. The Georgia data revealed a greater difference by level of degree (bachelor's or master's) than by separation between teachers possessing regular or provisional certification. At the bachelor's level the arts and sciences graduates in general scored slightly higher as a group, but at the master's level the teacher education graduates outscored the arts and sciences persons. Whether this is a function of the selection processes employed or the fact that the test is designed to test minimum levels of subject matter cannot be ascertained from this study.

Because of the small numbers associated with the respective NTE scores, data could not be analyzed by level, and no generalizations other than those made for elementary teachers were possible. The teacher education graduates in elementary education outscored those who were provisionally or temporarily certified on the Area Examinations in Louisiana. It appears on the basis of these limited data that at the elementary level, in the specialized knowledge tested for in the NTE Area Examinations, the teacher education graduates are more knowledgeable than the arts and sciences graduates. Because of the small number of teachers being provisionally certified in Georgia for elementary education, no findings were possible. The differences for the elementary and secondary fields may indicate similarities in terms of content focus at the secondary level for both teacher education and for arts and sciences graduates. The fact that programs to prepare elementary teachers are generally weighted toward the teaching of children--tested for in the area test--could explain the differences.

Evaluation of On-the-Job Performance of Teachers

The data from these studies indicate few differences in the on-the-job performance of teachers as measured by several types of currently used evaluation instruments in which the school principal is the rater of a teacher's on-the-job performance.

The data from the Georgia metropolitan district do indicate that teachers who are regularly certified receive a better rating than the provisionally certified arts and sciences graduates when rated on the district instrument. Because of problems in equating years of experience and because of small numbers in the study, it is not known to what extent the differences are a function of the course background of the teacher or of the experience level.

For the much larger North Carolina population, no differences were found between the provisionally certified teachers and those with regular certification, nor were differences found when the two groups were compared by experience level. The first-year teachers in both groups received similar ratings, which did not differ from the ratings of experienced teachers.

Data from both groups revealed that for experienced teachers, most evaluations are at the satisfactory or above satisfactory level. This may be, in part, because those teachers who were not able to perform were no longer on the job or that principals tend to give similar ratings to highly experienced teachers. However, the North Carolina data show that even among first-year teachers, very few received a mean score which would be considered less than satisfactory.

All the performance results should be examined with the knowledge that evaluation instruments have not been validated against student learning, and that the ratings are confined to the principals' perceptions of the ability of the teacher to perform a particular activity. The fact that few differences occurred in the North Carolina data in terms of type of certification or level of experience calls into question whether or not the

comparison of performance in terms of a total mean score really reveals differences in the performance of the teacher. On the other hand, the evaluations are designed for employment decisions and to identify weaknesses of the teachers. In this respect, the instruments do not seem to distinguish between the two groups.

Because of the nature of the instrumentation and the state of previous research, further studies are needed on other populations using the instruments examined in this study as well as other evaluation instruments. In particular, the instruments need to be validated against student achievement to determine whether or not the competencies being rated are those demonstrated by teachers who are effective in terms of student achievement. The ability of the school principal to effectively "rate" teacher competencies needs further research.

Certification of Arts and Sciences Majors

For the populations studied, the data support the contention that, as measured by certification tests, arts and sciences majors are as well--or possibly better--qualified than teacher education graduates in general education, general professional knowledge, and subject area knowledge at the secondary level. At the elementary level, more research needs to be conducted to determine if the preliminary indication of these studies that teachers prepared in teacher education programs have a greater knowledge of what is measured on the NTE Elementary Education Area Examination than the arts and sciences majors is similar for other populations.

The studies reveal mixed results in terms of performance in the classroom. The North Carolina data, using different types of evaluation instruments, revealed no real differences, while the results from the Georgia metropolitan district, based on small numbers of teachers, favored the teacher education graduates. The instrumentation and the validity of the instruments for measuring the effectiveness of teachers is certainly

brought into question. The interpretation of the results of these studies must be made with that in mind.

As states begin or consider to begin certifying arts and sciences graduates, it will be important to gather data on teachers currently in the classroom as well as those who enter the classroom under these new circumstances. Regionwide pooling of data for research would be helpful. If certification is to function for quality control, then states need to monitor practices and consider whether old or new policies are in the best interests of students.

States also need to carry out further research to determine the ability of evaluation instruments to distinguish between the good and the poor teacher in the classroom. Can one generic instrument be used across all subject and experience levels? Studies on inclusion of data on student achievement and attitudes for validation of teacher performance instruments should be given high priority.

## REFERENCES


Andrews, J., Blackmon, C. R., Davidson, C. W., & Mackey, J. A. (1982). Predicting seniors' National Teacher Examination scores while sophomores. The Boardman, Baton Rouge, LA: Louisiana School Boards Association.

Bausell, R. B., & Moody, W. B. (1972, January). Are teacher preparatory institutions necessary? Phi Delta Kappan, p. 298.

Beery, J. R. (1962). Does professional preparation make a difference? Journal of Teacher Education, 13, pp. 386-395.

Benderson, A. (1982). Focus: Teacher Competence Princeton, NJ: Educational Testing Service.

Bledsoe, J., Cox, J. V., & Burnham, R. (1967). Comparison between selected characteristics and performance of provisionally and professionally certified beginning teachers in Georgia. Athens, GA: University of Georgia (ERIC Reproduction Service No. ED 015 553).

Capie, W. (1981). Teacher performance assessment instruments, administration technical manual. Athens, GA: University of Georgia.

Coker, H., Medley, D. M., & Soar, R. S. (1980, October). How valid are expert opinions about effective teachers? Phi Delta Kappan, pp. 131-134.

Copley, P. O. (1975). A study of the effect of professional education courses in beginning teachers. Springfield, MO: Southwest Missouri State University. (ERIC Document Reproduction Service No. ED 098 147).

Commission on Humanities (1980). The humanities in American life, a report of the Commission on the Humanities. Berkeley, CA: University of California Press.

Dickson, G. E., & Wiersma, W. (1980, May). Research and evaluation in teacher education. Toledo, OH: The Center for Educational Research and Services, University of Toledo.

Ebel, R. L., Noll, V. H., & Bauer, R. M. (Eds.). (1969). The encyclopedia of educational research (4th ed.). London: The Macmillan Company.

Educational Testing Service. (1981). Interpreting National Teacher Examinations scores. Princeton, NJ: ETS.

Educational Testing Service. (1982, November 3 and 10). Correspondence with Roy A. Hardy.

Educational Testing Service. (1978, April 14). The National Teacher Examinations, notes on their reliability and standard error of measurement.

Georgia State Board of Education. (no date). Georgia teacher testing program. . . a review. Atlanta, GA: Report to Georgia State Board of Education.

Hall, H. O. (1964). Professional preparation and teacher effectiveness. Journal of Teacher Education, 15, pp. 72-76.

Hamm, D. W., Liu, J. M., & Brinlee, P. (1983, April). The observer as instrument: Implications for teacher evaluation observation systems and observer training programs. Paper presented at annual meeting of American Educational Research Association, Montreal, Canada.

Hawkins, E. E. & Stoops, E. (1966). Objective and subjective identification of outstanding elementary teachers. The Journal of Educational Research, 59 (8), pp. 344-346.

Lins, L. (1946). The prediction of teaching efficiency. Journal of Experimental Education, 15, pp. 2-60.

LuPone, O. J. (1961). A comparison of provisionally certified and permanently certified elementary school teachers in selected school districts in New York state. The Journal of Educational Research, 55 (2), pp. 53-63.

Medley, D. M. (1977). Teacher Competence and Teacher Effectiveness. Washington, D. C.: American Association of Colleges for Teacher Education.

Medley, D. M., & Hill, R. A. (1970). Cognitive factors in teaching style. Paper presented at the annual meeting of the American Educational Research Association, Minneapolis, MN.

Medley, D. M., & Mitzel, H. E. (1963). Measuring classroom behavior by systematic observation. In N. L. Gage (Ed.), Handbook of research on teaching (pp. 247-328). Chicago, ILL: Rand McNally & Company.

National Center for Education Statistics. (1982. November). Bulletin. Washington, D. C.: U. S. Department of Education.

Nie, H. M., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. M. (1975). SPSS: Statistical package for the social sciences (2nd ed.). New York: McGraw-Hill

Piper, M. K., & Sullivan, P. S. (1981). The National Teacher Examination: Can it predict classroom performance? Phi Delta Kappan, p. 401.

Popham, J. W. (1971, June). Teaching skill under scrutiny. Phi Delta Kappan, pp. 599-601.

Quirk, T. J., Witten, B. J., & Weinberg, S. F. (1973). Review of studies of the concurrent and predictive validity of the National Teacher Examinations. Review of Educational Research, 43 (1).

Rosenshine, B. (1976). Classroom instruction. In National Society for Study of Education Yearbook, Psychology of Teaching Methods. Chicago, IL: University of Chicago Press.

Sheehan, D. S., & Marcus, M. (1978). Teacher performance on the National Teacher Examinations and student mathematics and vocabulary achievement. The Journal of Educational Research, 71, pp. 134-136.

Soar, R. S., Medley, D. M., & Coker, H. (1983, December). Teacher evaluation: A critique of currently used methods. Phi Delta Kappan, pp. 239-246.

Southern Regional Education Board. (1981). The need for quality: A report to the Southern Regional Education Board by its Task Force on Higher Education and the Schools. Atlanta, GA: SREB.

Southern Regional Education Board. (1982). Teacher testing and assessment: An examination of the National Teacher Examinations (NTE), The Georgia Teacher Certification Test, (TCT) The Georgia Teacher Performance Assessment Instrument (TPAI) for a selected population. Atlanta, GA: SREB.

State Council of Higher Education for Virginia. (1981). Final report on the study of teacher preparation programs in Virginia pursuant to house joint resolution No.100 (1980 Session). Richmond, VA.

Vance, V. S., & Schlechty, P. C. (1982, September). The distribution of academic ability in the teaching force: Policy implications. Phi Delta Kappan, pp. 22-30.

Weaver, W. T. (1979, September). In search of quality: The need for talent in teaching. Phil Delt Kappan, pp. 29-32, 46.

APPENDIX A

GEORGIA TEACHER CERTIFICATION TEST RESULTS
FOR A SELECTED POPULATION

| | ALL* | | HUMANITIES** | | COMMUNICATIVE ARTS | | SOCIAL STUDIES | | MATHEMATICS | | SCIENCE | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Arts & Sciences | Teacher Ed. | Arts & Sciences | Teacher Ed. | Arts & Sciences | Teacher Ed. | Arts & Sciences | Teacher Ed. | Arts & Sciences | Teacher Ed. | Arts & Sciences | Teacher Ed. |
| Bachelor's 1981-82 | | | | | | | | | | | | |
| Number | 109 | 321 | 59 | 165 | 28 | 58 | 29 | 96 | 12 | 37 | 15 | 35 |
| Mean | 79.6 | 78.6 | 80.3 | 79.0 | 79.0 | 77.5 | 81.7 | 79.9 | 79.7 | 79.0 | 79.2 | 79.7 |
| Standard Deviation | 6.2 | 6.0 | 6.0 | 6.2 | 4.8 | 4.7 | 6.7 | 6 7 | 5.9 | 6.1 | 7.0 | 6.3 |
| Bachelor's 1982-83 | | | | | | | | | | | | |
| Number | 121 | 78 | 60 | 38 | 14 | 21 | 41 | 16 | 24 | 15 | 19 | 9 |
| Mean | 79.0 | 78.1 | 82.9 | 78.4 | 79.2 | 77.1 | 82.0 | 80.1 | 77.7 | 78.9 | 77.5 | 75.5 |
| Standard Deviation | 8.4 | 5.5 | 6.3 | 5.8 | 5.0 | 4.6 | 6.5 | 7.0 | 6.2 | 5.6 | 5.9 | 3.0 |
| Master's 1981-82 | | | | | | | | | | | | |
| Number | 19 | 31 | 13 | 22 | 4 | 9 | 8 | 10 | 3 | 3 | 2 | 2 |
| Mean | 81.1 | 83.4 | 84.0 | 84.1 | 87.9 | 80.9 | 81.2 | 84.8 | 76.8 | 88.9 | 71.6 | 73.9 |
| Standard Deviation | 7.6 | 7.3 | 7.1 | 7.2 | 3.6 | 5.3 | 7.3 | 7.3 | 8.0 | 6.2 | 4.0 | .50 |
| Master's 1982-83 | | | | | | | | | | | | |
| Number | 18 | 6 | 15 | 3 | 2 | 0 | 12 | 2 | 1 | 0 | 0 | 2 |
| Mean | 81.5 | 86.1 | 81.0 | 89.6 | 83.1 | --- | 81.0 | 91.0 | 94.9 | --- | --- | 83.7 |
| Standard Deviation | 6.8 | 7.8 | 6.1 | 5.4 | 4.8 | --- | 6.6 | 6 9 | --- | --- | --- | 12.87 |
| Total=703 | | | | | | | | | | | | |
| Number | 267 | 436 | 147 | 228 | 48 | 88 | 90 | 124 | 40 | 55 | 36 | 48 |
| Bachelor's | | | | | | | | | | | | |
| Number | 230 | 399 | 119 | 203 | 42 | 79 | 70 | 112 | 36 | 52 | 34 | 44 |
| Mean | 79.3 | 78.6 | 81.6 | 78.9 | 79.1 | 77.4 | 82.0 | 79.9 | 78.3 | 78.9 | 78.2 | 78.8 |
| Master's | | | | | | | | | | | | |
| Number | 37 | 37 | 28 | 25 | 6 | 9 | 20 | 12 | 4 | 3 | 2 | 4 |
| Mean | 81.3 | 83.8 | 82.4 | 84.8 | 86.3 | 80.9 | 81.1 | 85.9 | 81.3 | 88.9 | 71.7 | 78.8 |

* Includes the total population with the following subject areas represented: communicative arts, social studies, mathematics, science, French, Latin, Spanish, German, business, distributive education, agriculture, home economics, industrial arts.

** Humanities includes communicative arts, social studies, and all languages (Spanish, French, German, Latin).

NOTE: Subject breakdowns are shown only for areas with sufficient numbers for analysis.

## APPENDIX B

## TYPES OF CERTIFICATES CURRENTLY ISSUED TO TEACHERS IN GEORGIA

| Educational Minimum Degree Requirements | Preparation Completed | Designation | Validity | Provisos |
|---|---|---|---|---|
| Baccalaureate | Before 5/1/80 | T4 | 5 yrs renewable<br>1 yr extendable for 4 years | All conditions* have been met<br>While all conditions* are being met |
| | On or after 5/1/80 | NT4 | 5 yrs, convertible to PBT | All conditions* have been met, and convertible to PBT after 2 successful on-the-job assessments |
| | | | 1 yr extendable for 2 years | While all conditions* are being met |
| | | PBT4 | 5 yrs renewable | After meeting on-the-job assessments |
| | | NB4 | 1 yr not renewable | All conditions* have been met except passage of CRT, convertible to NT4 on passage of CRT |
| | | B4 | 1 yr renewable for 2 yrs if professional courses & conditions* satisfied | Field requirements are met, but lacks professional preparation and employer requests |
| Master's | Before 5/1/80 | BPA4 | 1 yr renewable | |
| | | T5 | 5 yrs renewable | see T4 |
| | | NT5 | 5 yrs convertible to PBT | see NT4 |
| | On or after 5/1/80 | PBT5 | 5 yrs renewable | see PBT4 |
| | | NB5 | 1 yr not renewable | see NB4 |
| | | B5 | 1 yr renewable 2 yrs | see B4 |
| | | PA5 | 1 yr renewable 4 times | see PA4 |

Source: Galambos, E. C., (1981). Certificates in Georgia with comparisons for other states. Georgia Professional Standards Commission, Atlanta, GA.

*Conditions: Preparation finished after 7/1/76, and completed: exceptional children course, reading course, and recency of study requirements. If preparation finished after 5/1/80, these requirements govern plus passage of Teacher Certification Tests.

N = Nonrenewable
PB = Performance-Based
NB = Nonrenewable Provisional
Provisional

APPENDIX A

GEORGIA METROPOLITAN DISTRICT

# TEACHER PERFORMANCE OBSERVATION RECORD

| COMPETENCIES AND INDICATORS | Improvement Needed | Satisfactory |
|---|---|---|
| **PLANS INSTRUCTION** | | |
| 1.01 Specifies or selects learner objectives for lessons | 1 2 3 | 4 5 |
| 1.02 Specifies or selects teaching methods for lessons | 1 2 3 | 4 5 |
| 1.03 Specifies or selects contents, materials, and equipment for lessons | 1 2 3 | 4 5 |
| 1.04 Specifies or selects materials and procedures for assessing learner progress on the objectives | 1 2 3 | 4 5 |
| 1.05 Plans instruction at a variety of levels | 1 2 3 | 4 5 |
| 1.06 Identifies and refers learners who require the assistance of specialists | 1 2 3 4 | 5 |
| **USES TECHNIQUES, METHODS, AND MEDIA RELATED TO THE OBJECTIVES** | | |
| 2.07 Uses teaching methods appropriate for objectives, learners, and environment | 1 2 3 4 | 5 |
| 2.08 Uses instructional equipment and other instructional aids | 1 2 3 | 4 5 |
| 2.09 Uses instructional materials that provide learners with appropriate practice on objectives | 1 2 3 4 | 5 |
| **COMMUNICATES WITH LEARNERS** | | |
| 3.10 Gives directions and explanations related to lesson content | 1 2 3 4 | 5 |
| 3.11 Provides feedback to learners throughout the lesson | 1 2 3 4 | 5 |
| 3.12 Uses comprehensible written and oral expression with learners | 1 2 3 4 | 5 |
| **DEMONSTRATES A VARIETY OF TEACHING METHODS** | | |
| 4.13 Implements learning activities in a logical sequence | 1 2 3 | 4 5 |
| 4.14 Uses a variety of teaching methods | 1 2 3 | 4 5 |
| 4.15 Works with individuals, small groups, and large groups effectively | 1 2 3 | 4 5 |
| **REINFORCES AND ENCOURAGES LEARNER INVOLVEMENT IN INSTRUCTION** | | |
| 5.16 Uses procedures to maintain learners involved in lessons | 1 2 3 | 4 5 |
| 5.17 Maintains learner involvement in lessons | 1 2 3 | 4 5 |
| 5.18 Reinforces and encourages the efforts of learners to maintain involvement | 1 2 3 4 | 5 |
| **DEMONSTRATES AN UNDERSTANDING OF THE SUBJECT** | | |
| 6.19 Helps learners recognize the purpose/importance of topics or activities | 1 2 3 | 4 5 |
| 6.20 Demonstrates knowledge in the subject area | 1 2 3 4 | 5 |
| **ORGANIZES TIME, SPACE, MATERIALS, AND EQUIPMENT FOR INSTRUCTION** | | |
| 7.21 Attends to non-instructional tasks | 1 2 3 | 4 5 |
| 7.22 Uses instructional time efficiently | 1 2 3 4 | 5 |
| 7.23 Provides a learning environment that is attractive and orderly | 1 2 3 | 4 5 |
| **DEMONSTRATES ENTHUSIASM FOR TEACHING, LEARNING, AND THE SUBJECT** | | |
| 8.24 Communicates personal enthusiasm | 1 2 3 4 | 5 |
| 8.25 Stimulates learner interest | 1 2 3 | 4 5 |
| 8.26 Conveys the impression of knowing what to do and how to do it | 1 2 3 4 | 5 |
| **HELPS LEARNERS DEVELOP POSITIVE SELF-CONCEPTS** | | |
| 9.27 Demonstrates warmth and friendliness | 1 2 3 4 | 5 |
| 9.28 Demonstrates sensitivity to the needs and feelings of learners | 1 2 3 4 | 5 |
| 9.29 Demonstrates patience, empathy, and understanding | 1 2 3 4 | 5 |
| **MANAGES CLASSROOM INTERACTIONS** | | |
| 10.30 Responds [illegible] to learners and their behavior | 1 2 3 4 | 5 |
| 10.31 [illegible] interpersonal relationships | 1 2 3 4 | 5 |
| [illegible] | 1 2 3 4 | 5 |
| [illegible] | 1 2 3 4 | 5 |

APPENDIX D
NORTH CAROLINA
TEACHER PERFORMANCE APPRAISAL INSTRUMENT

*INSTRUCTIONS*
1. *Based on the evidence from observation and discussion, the evaluator is to rate the teacher's performance with respect to the 33 basic elements of teaching listed below.*
2. *The evaluator is encouraged to add pertinent comments at the end of each major function.*
3. *The teacher is provided an opportunity to react to the evaluator's ratings and comments.*
4. *The evaluator and the teacher must discuss the results of the appraisal and any recommended action pertinent to it.*
5. *The teacher and the evaluator must sign the instrument in the assigned spaces.*
6. *The instrument must be filed in the teacher's personnel folder.*

---

*Teacher Name* _____

School _____

Rating Scale
(Please Check)

| Performs Unsatisfactorily | Needs Improvement In Performance | Meets Performance Expectations | Exceeds Performance Expectations | Superior Performance | Not Applicable |
|---|---|---|---|---|---|

*The following are Broad Program Functions. They refer to planning, operating, and updating the instructional program as a total program extending over the school year.*

A. Major Function: Planning the Program

1. Contributes as requested to the development of annual objectives for the school. ☐ ☐ ☐ ☐ ☐ ☐

2. Develops an annual instructional plan that includes the formulation of objectives, strategies, timelines, and evaluation procedures consistent with annual school objectives. ☐ ☐ ☐ ☐ ☐ ☐

Comments _____

---

B. Major Function: Overseeing the Program

1. Applies curriculum scope, sequence, continuity, and balance in carrying out the annual instructional plan. ☐ ☐ ☐ ☐ ☐ ☐

2. Implements learning strategies that address the needs identified in the annual instructional plan. ☐ ☐ ☐ ☐ ☐ ☐

3. Uses appropriate evaluation methods to determine whether the annual instructional plan is working. ☐ ☐ ☐ ☐ ☐ ☐

4. Makes changes in the annual instructional plan when evaluation indicates a need, and seeks advice and assistance if needed. ☐ ☐ ☐ ☐ ☐ ☐

Comments _____

---

C. Major Function: Updating the Program

1. Renews competence and keeps up with advances in child growth and development and uses this knowledge to improve the instructional program. ☐ ☐ ☐ ☐ ☐ ☐

2. Renews competence and keeps abreast of new knowledge, research, and practice in subject area(s) and applies this knowledge to improve the instructional program. ☐ ☐ ☐ ☐ ☐ ☐

Comments _____

---

Note on Rating Scale: Any one of three rating scales was suggested for use in 1982-83. The five-point scale is illustrated here. The four-point scale is obtained by omitting the "Superior Performance" category. The three-point scale is obtained by omitting the "Superior Performance" category and the "Performs Unsatisfactorily" category.

*The following are Particular Technical Functions. They refer to the means by which the teacher adapts the broad program functions to lessons and units of study on a daily basis.*

| | Rating Scale (Please Check) | | | | | |
|---|---|---|---|---|---|---|
| | Performs Unsatisfactorily | Needs Improvement In Performance | Meets Performance Expectations | Exceeds Performance Expectations | Superior Performance | Not Applicable |
| D. Major Function: Managing Daily Instruction | | | | | | |
| 1. Prepares daily lesson plans, makes classroom presentations, conducts discussions, encourages practice, and corrects student work in a manner that demonstrates subject area competence. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 2. Correlates subject matter to students' interests, needs, and aptitudes. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 3. Uses resources, materials, and enrichment activities that are related to the subject(s). | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 4. Employs instructional methods that are appropriate to the instructional objectives. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 5. Involves students, parents, and others as needed to help ensure that students keep up with daily lessons. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Comments _____ | | | | | | |
| E. Major Function: Differentiating Instruction | | | | | | |
| 1. Identifies students' strengths and weaknesses in relation to objectives to determine if grouping is required because of differing skill levels. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 2. Groups students as needed for effective teaching and learning. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 3. Uses the school's media center to support and supplement instructional activities. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 4. Provides instructional activities that aid students in becoming independent learners. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Comments _____ | | | | | | |
| F. Major Function: Individualizing Instruction | | | | | | |
| 1. Monitors individual student achievement of objectives as teaching occurs. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 2. Provides individual students with prompt feedback on their progress and provides necessary remediation. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 3. Adjusts instruction to objectives and individual student needs on a daily basis. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 4. Arranges to have appropriate materials and equipment available to satisfy individual needs. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Comments _____ | | | | | | |
| G. Major Function: Supervising | | | | | | |
| 1. Manages the daily routine so that students know what they are to do next and are able to proceed without confusion. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 2. Keeps student talk and movement at a level that lets each student attend to his or her instructional task without interruption. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 3. Maintains a pleasant working atmosphere that does not stifle spontaneity and warmth. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Comments _____ | | | | | | |

*The following are Indirect Facilitating Functions. They refer to a moderately related set of activities that do not involve direct teaching between teacher and student, but have important effects on the success of that direct teaching. Non-Instructional Duties refer to the teacher's essential role in the logistics of administering a program to a large social group of several hundred students in a limited space.*

Rating Scale (Please Check)

| Performs Unsatisfactorily | Needs Improvement In Performance | Meets Performance Expectations | Exceeds Performance Expectations | Superior Performance | Not Applicable |
|---|---|---|---|---|---|

H. Major Function: Human Resources

1. Uses student talent as a resource in instructing, developing materials, and operating equipment.
2. Makes appropriate use of volunteers and resource teachers with special skills and knowledge.
3. Makes use of appropriate community resources to extend classroom learning.
4. Makes effective use of other professional personnel to improve instruction and classroom management.

Comments _____

I. Major Function: Human Relations

1. Shows respect for the worth and dignity of all students.
2. Is aware of and encourages respect for cultural differences.
3. Establishes rapport with parents.

Comments _____

J. Major Function: Non-Instructional Duties

1. Carries out non-instructional duties as assigned or as a need is perceived.
2. Adheres to established laws, rules, and regulations.

Comments _____

Evaluator's Summary Comments _____

Teacher's Reactions to Evaluation _____

Evaluator's signature and date

Teacher's signature and date

Signature indicates that the written evaluation has been seen and discussed.

# APPENDIX E

## NORTH CAROLINA EVALUATION COMPARISONS

| | Regular Certification | Provisional Certification | Mean | Standard Deviation | Scales Used |
|---|---|---|---|---|---|
| tate Evaluations (N=146) | | | | | |
| Type A | 3 | 0 | 3.8 | .15 | (2) Does not meet standard<br>(3) Meets minimum standard<br>(4) Exceeds standard<br>(4) Superior attainment<br>(0) Not applicable |
| Type B | 39 | 15 | 3.1 | .32 | (2) Performs unsatisfactorily<br>(2) Needs improvement in performance<br>(3) Meets performance expectations<br>(4) Exceeds performance expectations<br>(0) Not applicable |
| Type C | 30 | 10 | 3.5 | .35 | (2) Performs unsatisfactorily<br>(2) Needs improvement in performance<br>(3) Meets performance expectations<br>(4) Exceeds performance expectations<br>(4) Superior performance<br>(0) Not applicable |
| Type D | 28 | 21 | 3.3 | .33 | (2) Needs improvement in performance<br>(3) Meets performance expectations<br>(4) Exceeds performance expectations<br>(0) Not applicable |
| lean Score | | | 3.3 | | |
| istrict Evaluations (N=146) | | | | | |
| (82% agreement between 2 raters using a 5-point scale) | | | 3.4 | .70 | |
| ll Evaluations--Mean Score = 3.3 | | | | | |