An assessment of teacher competency testing: a case study of the teacher certification test (TCT) in the state of Georgia

Tanya Dekova Young
Atlanta University
AN ASSESSMENT OF TEACHER COMPETENCY TESTING: A CASE
STUDY OF THE TEACHER CERTIFICATION TEST (TCT)
IN THE STATE OF GEORGIA

A DEGREE PAPER
SUBMITTED TO THE FACULTY OF ATLANTA UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF PUBLIC ADMINISTRATION

BY
TANYA DEKOVA YOUNG

DEPARTMENT OF PUBLIC ADMINISTRATION

ATLANTA, GEORGIA
JULY 1988
ABSTRACT

PUBLIC ADMINISTRATION

Young, Tanya DeKova B.S., Georgia State University, 1985

An Assessment of Teacher Competency Testing: A Case Study of the Teacher Certification Test (TCT) in the State of Georgia

Advisor: Professor George O. Kugblenu

Degree Paper dated July 1988

The primary purpose of this paper is to examine the use of standardized tests to assess teacher competency in the State of Georgia. The tests, which are most commonly known as the Teacher Certification Tests (TCT), are designed to measure individuals' knowledge of subject matter for the grade levels and fields for which they seek certification.

This study is significant because of the increasing public demand for accountability in education in this country. In an effort to respond to public pressure to improve the quality of education in Georgia, state legislators adopted a mandatory competency testing policy for its teachers, a provision in the Quality Basic Education Act of 1985. Currently, the TCT is the sole basis for assessing teacher competence in Georgia. Only those individuals who pass the examination are considered to have demonstrated a minimum knowledge of their subject area and are deemed competent to teach in the Georgia public school systems.
The major problem that has surfaced since the TCT was introduced has been the differential passing rates for minority teachers and failure by some veteran teachers to obtain passing scores on the examination. This has given rise to allegations by minority groups and the Georgia Association of Educators (GAE) that the TCT is racially biased and not job related.

This paper addresses the following issues associated with teacher competency testing in Georgia: racial bias, job-relatedness, and predictive validity. The major findings of this study reveal that the TCT is not racially biased; however, evidence exists to support the claim that the test is not sufficiently job related or a valid predictor of job performance.

The primary data for this study were obtained from interviews with Susan Lacetti, staff writer for the Atlanta Journal and Constitution; Michael Kramer, Legal Counsel for the Georgia Association of Educators (GAE), and various teachers who have already taken the test. Sources of secondary data that were utilized included court documents, books, journals, magazines, and unpublished materials.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>The Purpose of the Study</td>
<td>5</td>
</tr>
<tr>
<td>II. STATEMENT OF THE PROBLEM</td>
<td>6</td>
</tr>
<tr>
<td>III. REVIEW OF THE LITERATURE</td>
<td>8</td>
</tr>
<tr>
<td>An Historical Overview</td>
<td>8</td>
</tr>
<tr>
<td>&quot;Test&quot; Defined</td>
<td>11</td>
</tr>
<tr>
<td>A Legal Overview</td>
<td>13</td>
</tr>
<tr>
<td>IV. METHODOLOGY</td>
<td>16</td>
</tr>
<tr>
<td>V. ANALYSIS OF THE PROBLEM</td>
<td>17</td>
</tr>
<tr>
<td>A Question of Racial Bias</td>
<td>20</td>
</tr>
<tr>
<td>A Question of Job-Relatedness</td>
<td>23</td>
</tr>
<tr>
<td>A Question of Predictive Validity</td>
<td>26</td>
</tr>
<tr>
<td>Testing versus Performance Evaluation</td>
<td>28</td>
</tr>
<tr>
<td>Testing versus Teacher Education Reform</td>
<td>29</td>
</tr>
<tr>
<td>VI. CONCLUSIONS</td>
<td>32</td>
</tr>
<tr>
<td>VII. RECOMMENDATIONS</td>
<td>34</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>36</td>
</tr>
</tbody>
</table>
I. INTRODUCTION

Public demand for accountability in education in the United States has increased considerably over the past twenty years. Moreover, public pressure to improve the quality of education in public school systems initiated educational reform nationwide. The reform movement has been characterized by a significant change in approaches to teacher certification in these school systems. The two phases of this movement included the implementation of approved teacher education programs in colleges and universities; and the introduction of competency testing as part of teacher certification requirements.¹

Prior to the 1960s, few states required accreditation of teacher education programs or testing of any kind for individuals entering the teaching profession. The majority of states certified pre-service teachers solely on the basis of successful completion of a teacher education program. The State of Georgia was no exception. It was not until the mid-1970s that the Georgia Department of Education (GDOE), with the assistance of the National Evaluation Systems (NES) of Amherst, Massachusetts, developed the first customized, criterion-referenced testing program. The Georgia program served as a model for other states attempting to develop and administer criterion-referenced programs.

Consequently, implementation of similar programs and/or the use of some kind of competency testing as part of certification procedures occurred in approximately 38 states.\footnote{Ibid., p. 1.}

The significant increase in efforts of state legislatures and boards of education to reform the standards for selecting and certifying teachers can be attributed, in part, to a series of commission reports providing evidence of declining educational standards and overall poor educational performance of our nation's students. Many of these reports were generated by the College Board, the National Science Board (NSB), and the National Commission on Excellence in Education (NCEE). The NCEE, in particular, declared in a 1983 report that America was "a nation at risk" and a "tide of mediocrity" has been the prevailing trend. The NCEE further addressed the failure of American education and made some of the following observations:

1. International comparisons of student achievement, completed a decade ago, reveal that on 19 academic tests American students were never first or second and, in comparison with other nations, were last seven times.

2. Some 23 million American adults were functionally illiterate by the simplest tests of everyday reading, writing, and comprehension.

3. About 13 percent of all 17 year-olds in the United States can be considered functionally illiterate. Functional illiteracy among minority youth may run as high as 40 percent.

4. Average achievement of high school students of most standardized tests is now lower than 26 years ago when Sputnik was launched.
5. The College Board's Scholastic Aptitude Tests (SAT) demonstrate a virtually unbroken decline from 1963 to 1980. Average verbal scores fell over 50 points and average mathematics scores dropped nearly 40 points.

6. Average tested achievement of students graduating from college is also lower.³

Based upon these observations and others, the NCCE made numerous recommendations for improving the quality of education in the United States. One such recommendation by the Commission was that persons preparing to teach should be required to demonstrate an aptitude for teaching, and to demonstrate competence in an academic discipline.⁴

Because teacher incompetence is frequently cited by citizens' groups, parents, and legislators as a major contributing factor to declining academic performance and test scores of students on standardized tests, the public support for teacher testing has been overwhelming. It is estimated that more than 85 percent of those surveyed in recent national polls, including the Gallup poll (1984), supported the idea that teachers themselves should be required to pass examinations for subject areas in which they teach. Therefore, the national response to the public's mandate for excellence in education has been the adoption of mandatory competency testing for teachers in most states.

The impetus for the teacher competency testing movement began primarily in the Southern states, spreading from Georgia, Louisiana, 


⁴ Ibid., p. 78.
and Florida to the West and Northeast. In 1978, Georgia recognized the need to reform the method of certifying its public school teachers and was the first state in the country to introduce a customized, criterion-referenced testing program. This program, mandated by the Georgia Department of Education, was applicable to only prospective teachers seeking certification for employment in Georgia public schools, K-12. Prior to the implementation of this program, there were no appropriate content knowledge tests or acceptable existing instruments to assess on-the-job performance of Georgia's public school teachers. Since the program's inception on September 1, 1978, applicants seeking initial certification have been required to pass written examinations which measure minimum content knowledge of the subject areas for which they intend to teach.  

Beginning on July 1, 1986, applicants for certification renewal are required by state law to pass a written criterion-referenced test before they can be recertified. Prior to this law, veteran teachers were recertified upon application with no legislative provision for examinations. However, with the enactment of the Quality Basic Education Act (QBE) in 1985, all applicants for renewable certificates, including experienced teachers were subject to mandatory testing requirements. Georgia is only one of three states (including Arkansas

and Texas) which now require veteran teachers to pass state-administered examinations for recertification.

The first statistics on the performance of the prospective teachers who took the Georgia teacher certification tests were made available by the State Department of Education in 1983. Of the 22,261 testees, 34 percent of the black candidates passed, compared to 87 percent of the white candidates. More recent statistics provided yearly by the Department of Education reveal a similar pattern of performance by minorities. "Of the 14,944 teachers who took the test between September 1985 and June 1986, 26 percent of those who failed were black compared with 4 percent who were white." In 1987, of the 327 who failed the TCT, 244 (75 percent) were black compared to 68 (21 percent) who were white, and 15 (4 percent) who were of other minority groups. The trend indicates that a disproportionate number of black pre-service and in-service teachers are being excluded from the teaching profession because of their inability to obtain passing scores on the Georgia teacher competency tests.

The purpose of this paper is to examine the use of standardized tests to assess teacher competency in the State of Georgia.

7 Ibid.
II. STATEMENT OF THE PROBLEM

One of the most complex tasks confronting policy-makers and legislators today is ensuring the quality of education in public schools. In an effort to respond to public demand for accountability in education, legislators across the nation have ordered tests of students and teachers in particular, to ensure that minimum competence in subject areas is exhibited by these groups.

Georgia, like the majority of states that implemented teacher certification testing programs, began testing preservice teachers, only, for initial certification; and more recently, veteran teachers to determine whether they should remain in the teaching profession. The primary aim of the certification testing program is to ensure that only those individuals who meet specified proficiency levels in their subject area content are permitted to teach. The test itself, which is most commonly known as the Teacher Certification Test (TCT), is designed to test individuals' knowledge of the subject matter for grade levels and fields for which they seek certification. Only those individuals who obtain a passing score are considered to have demonstrated a minimum knowledge of their subject area and are deemed competent to teach in the classroom.

The major problem that has surfaced since the TCT was introduced in the State of Georgia has been the differential passing rates for minority teachers and failure by some veteran teachers to obtain passing
scores. This has given rise to allegations by minority groups and the Georgia Association of Educators (GAE) that the test is racially biased and not job-related. Thus, the following questions persist:

1) Is the TCT racially biased?

2) How related is any pencil-and-paper test to what an individual actually does on the job?

3) Does obtaining a passing score on the TCT predict success as a teacher and thereby enhance the quality of education in public school systems?

4) Should the TCT be the sole basis for assessing teacher competence as is the practice in the State of Georgia?

These issues associated with teacher competency testing in Georgia are addressed in the analysis section of this study.
III. REVIEW OF THE LITERATURE

The literature reviewed for this study focuses on the historical and legal developments in the use of standardized tests in the United States. The widespread use of tests has generated much public interest and continues to be a controversial matter.

An Historical Overview

A review of the literature reveals that use of tests in America has had a diverse history and one that ultimately reflects our social and political history as well. At one time, many Americans, including some state legislators, naively accepted the widespread use of tests because they were simply considered apropos in society. From this attitude toward testing, society moved to another period in which they were shown to some degree not to be reflective of things compatible with our social principles. Currently, society is experiencing a competency testing movement in which there is a greater acceptance of tests "because we seek salvation in a time of doubt about our schools, our teachers, and indeed, about ourselves."9 By and large, society has rejected the notion that tests measure what people "are" and has progressed to the level of advocating the use of tests to measure what people "do" (i.e., to measure basic professional, academic, and on-the-job skills).

Testing to assess a teacher's competence is not new. One of the first state-administered examinations for prospective teachers occurred in 1849 in state-supported schools in Oregon. During this period, individuals who applied for teaching jobs very seldom had the more traditional credentials from colleges and universities which exist today. Because of the constricted labor economy in this state, many of these individuals sought teaching positions and the state was forced to hire them without the necessary credentials. In turn, this led to testing individuals to determine who was qualified to teach.° It would appear, then, that addressing the quality of teachers through testing has been a recognized practice rooted in public policy. However, according to Vlaanderen:

Testing for teacher certification ... has not been widely applied since colonial times, when oral and written examinations were administered by school boards and later by county superintendents. When state school systems began to emerge about 1825, certification required only the satisfactory completion of a prescribed curriculum at approved institutions. "Normal schools" supplied the curriculum; their graduates were presumed to be adequately equipped to teach schools. This approach, dominant by 1910, is being challenged [in the 1980s] ... by the teacher competency testing movement.11

Public interest in ensuring competence has, for the most part, extended beyond an interest in the quality of education to assessing competence in labor forces as well. This has been particularly true of such traditional professions as law, medicine, and other classes of

---

10 Ibid., p. 5.

practitioners in which examinations are used as a means of guaranteeing competence and screening out individuals from other backgrounds or the poorly trained. Even more fundamental than assessing competence of practitioners in various professions is that of testing job applicants as part of the selection process used by private/public employers. According to Robert Guion:

> The addition of testing to the selection process is based on the belief that something needs to be improved within the organization and that the testing, through improved selection, can affect that improvement.  

Testing as part of the employment selection process in the public sector, which was a by-product of legislation, occurred as early as the 1880s. With the passage of the Civil Service Act of 1883, the United States Civil Service System was established in an effort to address the public demand for accountability in government regarding selection of its employees. The enactment of this law was primarily in response to abuses of the patronage system in which case many unqualified or inexperienced persons were appointed to public sector jobs as reward for party loyalty, political support, and friendship. This is often referred to as the "spoils system." The consensus of the public was that candidates for government jobs should be selected on the basis of merit and competence. This led many public entities to adopt merit systems which include the use of standardized tests as part of the employment selection process. Today, candidates for employment or promotion to

---

government positions, subject to open competition, continue to be tested and ranked on the basis of test results.

"Test" Defined

The term "test" varies in meaning among laypersons, practitioners, and professionals. It has been defined by Byham as a systematic measure, or sample of behavior that is constructed to relate to a particular attribute being studied. A test, as defined by the Federal Equal Employment Opportunity Commission is "any paper-and-pencil or performance measure used as a basis for any employment decision" (e.g. eligibility for hire, transfer, promotion, membership, training, referral or retention).

This definition includes, but is not restricted to, measures of general intelligence; mental ability and learning ability; specific intellectual abilities; mechanical, clerical and other aptitudes; dexterity and coordination; knowledge and proficiency; occupational and other interests; and attitudes, personality, or temperament. The term "test" includes all formal, scored, quantified or standardized techniques of assessing job suitability including specific qualifying or disqualifying personal history or background requirements, scored interviews, biographical information blanks, interviewers' rating scales, scored application forms, etc.

Tests serve as a refinement on our observation because they tend to be objective; present a standard set of conditions in which measure-


15 Ibid.
Test usage has primarily three essential functions: 1) diagnosing needs, 2) benchmarking qualifications to "do," and 3) protecting against false credentials. While these functions of measurement are critical to reasonably accurate assessment of the needs and strengths as well as performance capacity of individuals within a society, consideration must be given to the limitations of tests themselves. These limitations include the following:

1) Tests are fallible. No test is entirely free of error.

2) Tests are only one source of information about an individual.

3) Tests may discriminate against disadvantaged individuals.

In spite of the most well-intended use of tests to assess the ability of an individual to perform or to attest competence, there continues to be issues and problems associated with testing. These problems are most commonly present in selection for employment or for admission to educational institutions, criterion-referenced testing, licensing, and certification, test bias, and competency testing. The controversies range from the exclusionary nature of tests to the improper use and misinterpretation of their results. Other technical issues associated with testing include validity, reliability, fairness

---

17 Daves, The Uses and Misuses of Tests, p. 5.
and accuracy. Moreover, tests themselves are imperfect and some have been found to be seriously flawed. Testing, therefore, has been subjected to extensive scrutiny, criticisms, and debate which has led to a high level of litigation over the years. By the early 1960s, the usage of tests proliferated to unprecedented levels. Typically, there was no corresponding evidence of their validity for the jobs for which they were administered. Moreover, it was clear that these tests were having an exclusionary effect on minorities and some of them had a discriminatory effect on certain groups of people.

A Legal Overview

The first landmark case on the issue of testing involved Griggs v. Duke Power Company in 1971. The plaintiffs in this case claimed that the educational and testing requirements of the Duke Power Company were discriminatory because many blacks were excluded from certain jobs on the basis of nonvalidated requirements. The court discovered that Duke Power had not conducted research to validate these requirements and standards; and determined that standards not validated against job success are unlawful. The court also defined business necessity as job relatedness and determined it to be the only lawful reason for testing in employment decisions. The U.S. Supreme Court ruled that the company administered intelligence test (Wonderlic Personnel Test), which blacks failed at a higher rate than whites, was not job related. While there are still many undecided legal issues, this decision helped to

\[19\] Ibid., p. 17.
clarify the application of Title VII of the Civil Rights Act to testing, and established two major points:

1) It is not enough to show a lack of discriminatory intent if the selection tool results in a disproportionate effect that discriminates against one group more than another or continues a past pattern of discrimination.

2) The employer has the burden of proving that an employment requirement is directly job-related. This case represented the first major challenge to employment tests and has been precedent-setting for other cases in the lower courts. Thus, the ruling has had a considerable effect on all private and public employment. For example, the Office of Personnel Management (OPM) administered the Professional and Administrative Career Examination (PACE) for a number of years to persons possessing unspecialized degrees to determine whether they qualified for a civil service appointment. PACE is a federal civil service professional entrance examination. A study by the General Accounting Office (GAO) revealed that 99 percent of all blacks failed PACE compared with 84 percent of all whites. The GAO found that OPM's validation procedures did not conform to the Uniform Guidelines on Employee Selection Procedures. PACE, which has been discontinued, came under fire from minority groups who sued on the

\[\text{References:}\]


grounds that it was discriminatory. As a result, in 1980, the OPM phased out the test and substituted a quota plan for hiring minorities.

For teacher certification testing, the most significant legislation was the passage of the 1972 amendment (Public Law 92-261) to the Civil Rights Act that struck out the exemption for educational personnel in public institutions, which extended the provisions of EEOC beyond private industry to state and local government agencies. . . . The 1972 amendment paved the way for later litigation (e.g. United States v. State of North Carolina, 1975) that successfully challenged the National Teacher Examination (NTE) as a teacher selection test. 22

Teacher certification testing will be increasingly subject to litigation as well because of the adverse impact on the large number of minority candidates. Therefore, test standards will be particularly important because of their functions, namely:

1) to provide the test developer and test user with guidelines for their own professional work

2) to provide information concerning what should and should not be expected from tests and test usage so that there is a yardstick or standard against which tests and test practices can be evaluated by the professional community.

3) to be used by the courts in consideration of cases alleging malpractice or improper treatment; and cases involving discrimination or adverse impact in employment selection. 23

Meeting test standards will ensure that the necessary technical information is made available for use in both legal and policy debate.

---


23 Daves, The Uocs and Misuses of Tests, p. 15.
IV. METHODOLOGY

This study employs descriptive analysis to examine the use of standardized tests to assess teacher competence. Primary data for this study were collected from interviews principally with staff writer Susan Lacetti, who covers the area of education in public schools for the Atlanta Journal and Constitution. Others interviewed were Faye Fox, President of the Georgia Association of Educators (GAE), and Attorney Michael Kramer, Legal Counsel for the GAE, and various educators in the Georgia public school system.

Unstructured interviews were conducted in person with Susan Lacetti and eight teachers who took the Teacher Certification Test (TCT). Faye Fox and Michael Kramer, who were interviewed by telephone, were selected because of their key roles in initiating the federal suit filed by the Georgia Association of Educators.

Secondary data were obtained from various sources such as court documents, books, magazines, and journals.
V. ANALYSIS

There are four major types of teacher tests administered in Georgia:

1. Admission tests are administered to prospective teachers prior to admission to a teacher education program, typically within a college or university. These tests usually assess basic reading, writing, and arithmetic skills.

2. Certification tests are required either after exiting a teacher program or before initial certification. These criterion-referenced tests usually assess knowledge of content in specific subject areas.

3. Recertification tests are administered to practicing classroom teachers as a condition for recertification. The same type of certification tests that are administered to prospective teachers are given to veteran teachers.

4. Performance assessments are administered to beginning teachers, usually during their first year of teaching while holding a provisional certificate.\(^{24}\) This on-the-job assessment is known as the Teacher Performance Assessment Instrument (TPAI) and measures eight teaching competencies considered essential for certification. Prior to 1980, Georgia was the only state to formally assess teacher performance as part of the certification process.

\(^{24}\)Rudner, What's Happening in Teacher Testing, p. 3.
Teacher certification testing normally covers one of the five domains: 1) basic skills, 2) general knowledge, 3) knowledge of specific content areas (e.g. elementary education, music, English), 4) professional knowledge and skills, and 5) teacher performance (e.g. Georgia TPAI).

Georgia teacher certification tests were developed to measure job-related skills and minimum content knowledge of subject areas. There are 28 tests for specific certification fields. Each test is a criterion-referenced test employing the multiple-choice format. The test assesses an individual's competence with regard to a specific set of performance objectives. These performance objectives measure content that practicing Georgia educators most frequently use in teaching students in the classroom. The test is based on the assumption that a minimal command of the subject matter is necessary in order to teach effectively.

The Georgia TCT is the result of a collaborative effort of the Georgia Department of Education, National Evaluation Systems (NES), and ad hoc committees composed of Georgia teachers from public schools and colleges. Typically, such a panel involved in the validation process which is based on legal precedent, addresses the following issues:

1. Opportunity to learn: Does each test question focus on material that at least 90 percent of the teacher education students would have had an opportunity to learn.

2. Relative emphasis: Is the relative emphasis given to content topics in the test the same as, more than, or less than the emphasis given the topic in the curriculum.

3. Coverage: Are there any major content topics not covered in the tests that are part of the curriculum.
4. Content appropriateness: How would the overall similarity between test content and the teacher education curriculum test be described.

5. Job relevance: What is the degree to which the content of the test is relevant to the job of the beginning practitioner—crucial, important, questionable, not relevant? 25

Having been scrutinized by a panel for item-objective content match, accuracy, lack of bias, minimal competency standards, and content validity, the Georgia teacher certification tests were validated to be job related and administered for the first time to prospective teachers in 1978.

The two complaints filed by the Georgia Association of Educators with the Equal Employment Opportunity Commission question the validity and the job-relatedness of the TCT which was validated in this manner. After the first complaint was filed, the Georgia Board of Education contracted the Center for Educational Research and Evaluation at the University of North Carolina in Greensboro to address such questions as bias, validity, job-relatedness, and standard setting. On December 20, 1985, the consultants for the Center for Educational Research and Evaluation declared in a report entitled "A Preliminary Evaluation of the Georgia Teacher Certification Program" that

Due to serious questions about the procedures used... there is inadequate evidence to support the claim that the Georgia TCT is sufficiently valid for either certification of teachers or other education personnel. 26

25Ibid., p. 35.

While plans are under way to correct flaws and remove doubt concerning its validity, the Board of Education still requires applicants for certification and recertification to pass a test of questionable validity.

A Question of Racial Bias

Minority groups and the GAE continue to protest the use of the TCT, alleging that it is racially biased and not job related. In October 1986, the GAE filed a class action suit against the State and the State Board of Education under Title VII of the Civil Rights Act of 1964 and the Fourteenth Amendment to the U.S. Constitution which requires governments to give all people "equal protection" under the laws.27

The controversy surrounding the equity and validity issues inherent in the TCT is not new. A similar case which challenged teacher testing and raised questions of equity and fairness occurred in 1977 (United States v. South Carolina).28 South Carolina's use of the National Teacher Examination (NTE), which was developed by the Educational Testing Service (ETS), had an adverse impact on blacks. Consequently, the State investigated the test and found that it did not meet legal requirements. The test was later validated to meet legal requirements and cut off scores were set in a systematic, empirical manner. The action resulted in validation and approval of some of the NTE tests for use in South Carolina.

27 Ibid.

Another case similar to the TCT involved Georgia's use of the NTE and standard setting. Prior to the introduction of the TCT in 1978, the NTE was the most widely used test for prospective teachers in Georgia; and it, too, has received a great deal of attention. In the case of the GAE v. Jack P. Nix, a federal district court found Georgia's use of the NTE to be irrational and arbitrary.\(^\text{29}\) Georgia required teachers to obtain very high scores on the NTE for eligibility for a six-year certificate that entitled them to more pay. Georgia failed to provide sufficient evidence that the cut off score (i.e., the minimum score designated as passing) was not arbitrarily established. The procedures used to establish the passing score were not empirical and resulted in a cut off score that bore no systematic relationship to successful job performance. In this case, the law prohibits use of a test in such a manner.

The consultants for Georgia, however, found in some instances that the cut off scores for the TCT appeared to have been established administratively without stated rationale or expert judgment. The cut off scores for this test are particularly significant because they are being used to make career determining decisions about Georgia's pre-service and veteran teachers. That is, those applicants or teachers who do not pass the test are denied certification and therefore ineligible to teach in Georgia public schools. Attorneys for the GAE have argued that termination of those teachers who fail the test should be postponed because of the questionable validity of the test. In addition, they

argue that state officials have not demonstrated that the passing scores are appropriate to the alleged purposes for which the TCT is being used.

Teacher testing programs are required by law to be equitable and fair. In order to reduce the incidence of litigation, public administrators and/or policy makers will need to ensure the validation of all tests used to determine any employment benefit; be cautious about using a test for a purpose other than that recommended by the test developer; and carefully review the use of any test that has an adverse impact on minority groups. If the test is clearly job related and adverse impact results from its use, then users of the test must provide evidence that the test is valid for its intended use. Finally, it is essential to use other appropriate criteria in addition to the test for decisions relating to promotion, identification of meritorious teachers, or similar purposes. Moreover, there must be a consistent effort by policy makers and administrators to ensure the proper use of standardized tests in assessing teacher competence.

According to the consultants of the Center for Educational Research and Evaluation, there is a significant disparity in the passing rates of minority and white educators who took the TCT. Although state officials ensured that test questions were evaluated by experts, other safeguards, such as field testing were not used in developing the TCT. While consultants suggest that there may have been negligence on the

---

part of state officials in that the TCT was not developed, administered and used in accordance with generally accepted professional standards, examination results refute the claim that the TCT is racially biased. The Department of Education reported that the passing rate is 83 percent for blacks and 98 percent for whites. This indicates a standard of fairness according to the Guidelines of the Equal Employment Opportunity Commission which stipulate a minority pass rate of only 80 percent of the majority pass rate on tests.  

A Question of Job-Relatedness

The Georgia teacher certification tests, like any pencil-and-paper tests, are subject to job-relatedness studies to identify the desired skills; and item classification studies to evaluate relative emphasis. Such studies provide content related evidence to evaluate whether items on a proposed or existing test assess skills related to a particular job.  

In developing and validating a test, the user must first ensure that proposed content is related to the job. Procedures in performing such a study will vary depending upon whether or not the user is developing or selecting a test; the domain of the test; and its purpose. It is, therefore, important for users of any standardized test to note some of the issues which are raised concerning job-relatedness studies which include the following:

---

1. the results of the job relatedness study may differ depending on whether respondents are asked to rate the importance of job elements for (a) the job of teaching (b) professionally acceptable practice for entry level teachers, and (c) effective teaching. Goals of the program must be clearly defined so that proposed content can be properly evaluated.

2. because the job description is the basis for the test plan and the test plan is the basis for developing test items, the job description must be comprehensive.

3. because job-relatedness studies rely on expert judgment, results can differ depending on the sample surveyed.

4. simple guidelines exist for selecting the criteria for determining whether a job element is sufficiently job related to be included in the test plan.  

Essentially, the job relatedness studies involve task definition, questionnaire development, sampling, and statistical analysis. When using any test, particularly a pencil-and-paper test, to determine an applicant's qualifications or level of competence for a job, the user must provide data to show that: the test standard is related to the job standard. Secondly, the applicants who score above the test standard will perform at or above the job standard; and that applicants who score below the test standard will perform below the job standard.

Meeting this challenge is particularly important because of the major assumptions which have been advanced by many of Georgia's policy makers and administrators who advocate teacher testing. Primarily it

33Ibid., p. 136.
is assumed that: 1) Teachers who pass the tests are more effective in the classroom than those who do not pass; 2) The quality of candidates in teacher education programs will improve as a result of implementing testing programs. Galambos, in his paper, "Testing for Certification and Recertification," states further that:

The rationale for testing prospective teachers is based on the assumption that one cannot teach what one does not know. While knowing something does not guarantee that the ability is present to transmit the knowledge to students, it is a necessary, if not sufficient, qualification for teaching.34

With increasing litigation concerning job relatedness of teacher tests, such assertions must eventually be verified by empirically relating performance on the test to actual performance on the job.

According the the Center's report on the TCT, questions unrelated to various fields of certification appeared on the TCT, and therefore, called into question whether the test is sufficiently job-related.35 Moreover, the consultants raised "serious questions" about the procedures used to determine job-relatedness. Thus, the State of Georgia and the Board of Education have not demonstrated that the pencil-and-paper instrument (i.e., TCT) tests the essential competencies of teaching; and that there is a correlation between scores and teacher performance on the job.

---


A Question of Predictive Validity

One of the questions that persists concerning the TCT is whether it has been validated as a predictor of job performance.

Validity is germane to any teacher testing effort. It refers to the extent to which an instrument measures what it is intended to measure. Most validation efforts for teacher certification focus on content validity. This process generally has two components: 1) determining whether the test content reflects significant aspects of the job and measures those aspects proportionally (usually assessed through job analysis); and 2) determining whether test items accurately measure that job's content. For example, in attempting to validate the Georgia TCT, experts in each content field made dichotomous judgments about test items.\(^{36}\) Judgments were obtained about the match among the objectives of the test and the items used to measure those objectives.

Some researchers and measurement specialists have argued that content validity is insufficient criterion for establishing the validity of a test. Madaus and Pullin also caution that any inferences made from scores of a test validated in this manner about an "individual's ability to teach successfully are at present unsubstantiated—and may, in fact be indefensible."\(^{37}\)

---


\(^{37}\)George F. Madaus and Diana Pullin, "Teacher Certification Tests: Do They Really Measure What We Need to Know," Phi Delta Kappan (September 1987): 33.
Nonetheless, the inference that is made from an individual's score on the TCT is considered predictive. The assessment is made that an individual does or does not have a command of the content and/or principles and techniques to be minimally successful in the classroom. It has been argued, however, that such an inference requires criterion-related evidence which is very seldom gathered. This evidence enables test developers to evaluate whether or not test scores are related to teacher effectiveness, a criterion which is often difficult to operationalize. Test developers have two forms of criterion-related validation available to them: (1) concurrent validation which examines the relationship between the test scores of practicing teachers and current performance; and 2) predictive validation which examines the relationship between the test scores of prospective teachers and future performance. Use of either form of test validation and the proper use of a well-constructed test can provide a better basis for making personnel decisions about preservice and in-service teachers.

While Georgia and the Board of Education have attempted to establish content validity, predictive validity for the TCT has been generally unsubstantiated. Only scant evidence exists to show direct relationship between scores on standardized tests and on-the-job competence primarily because suitable and stable criteria are often difficult to operationalize in empirical studies examining links between teacher competency tests and teacher performance. Yet, it is still debated whether content-related evidence is a sufficient basis to justify validity of the inference made from scores on the TCT. Moreover, the TCT
has been used widely as an instrument to assess competence and to promote excellence among teachers. This may be due, in part, to the fact that the courts have indicated a willingness to accept teacher testing as a valuable assessment device. However, judicial approval of teacher testing does not necessarily mean it is wise public policy, but only that the courts have found that teacher testing procedures meet the constitutional requirements for equity and fairness in those cases which raised these issues.

Testing versus Performance Evaluation

Teacher testing measures knowledge, only. It is not synonymous with evaluation, nor is it a substitute for a performance evaluation. An evaluation is a more comprehensive process by which to determine an individual's achievement, competence, and/or performance on the job. Although on-the-job skills are the most appropriate domain in teacher certification, Georgia is only one of approximately 10 states that require assessment of these skills. However, this performance assessment is made of Georgia's beginning teachers only.

In a 1984 report to Governor Joe Frank Harris, entitled "Priority for a Quality Basic Education," from the Governor's Education Review Commission, the Commission stated that "personnel evaluation is not adequately addressed in Georgia's schools." The Commission reported that many of the experienced professionals are infrequently evaluated or not evaluated at all; and few are evaluated against student outcomes.

---

Moreover, evaluation instruments were found to be insufficiently diagnostic to provide a basis for an improvement plan. Finally, a survey of the teachers administered by the Commission revealed that principals are weak in classroom observation skills, conducting observation conferences, and developing improvement plans.

In light of this significant information, it is reasonable to assume that teacher quality is only one of many factors that have contributed to the mediocre performance of Georgia's public schools. At present, the Georgia teacher competency test is the sole basis for assessing teacher competence. Essentially, personnel decisions are made on the basis of test scores without tempering the results with personal judgment.

It would appear that Georgia opted for teacher competency tests as a quick solution to the complex problem of improving the quality of education in the public schools. Georgia policymakers, however, have engaged in adopting a more comprehensive package of educational reform that goes well beyond testing. With the enactment of the Quality Basic Education Act (QBE) in 1985, issues such as personnel evaluations, curriculum reform, and salaries have been addressed. However, the testing provision in QBE is no substitute for personnel evaluations or quality in teacher preparation programs.

**Testing versus Teacher Education Reform**

Teacher preparation programs traditionally have not played a significant role in preventing unqualified individuals from entering the teaching profession. Differences among educational institutions
and grading processes in college classes reduce the usefulness of course grades as a standardized measure of teacher competence. Many states have responded to the lack of standardization in college graduation requirements with mandatory teacher testing of their prospective teachers.

Examination results show that many minority candidates who failed the TCT in Georgia graduated from predominantly black institutions. In 1983, a number of these approved programs in the University System of Georgia were placed on probation by the State because less than 80 percent of their graduates passed the examination. It is reasonable to expect that such action may very well force colleges and universities to carefully review teacher education programs and/or improve their quality, and thereby strengthen their role in producing talented, well-educated minority teachers.

While minority teachers constitute approximately 12.5 percent of the national teaching force, it is estimated that by 1990, minority representation could be reduced to 5 percent if this trend of disparate racial impact continues unabated in this country. There will be sociological costs of reducing the number of minority teachers in public schools, especially since minority student enrollment is as high as 50

---


40 Ibid.
percent or more in some Georgia public school systems (e.g., Atlanta) and steadily increasing for others. While schools provide the principal focus of role models for students during their formative years, the benefits of having successful minority role models clearly will be lost.

Ibid.
VI. CONCLUSIONS

The GAE and minority groups have raised the legal questions of equity and fairness regarding the use of the TCT. They have made allegations that the test is racially biased and not job related. The pass rate on the TCT for blacks is 83 percent. Based on the fact that EEOC guidelines stipulate a minority pass rate of only 80 percent of the majority pass rate on tests, it is reasonable to conclude that the TCT is not racially biased. Evidence, however, exists to support the claim that the test is not sufficiently job related. Overall, predictive validity for the TCT has generally been unsubstantiated; that is, there is little or no empirical data to support the contention that a passing score on the TCT guarantees success as a teacher or assures teacher effectiveness. Therefore to reduce the incidence of litigation, administrators would be well-advised to refrain from making personnel decisions solely on the basis of test scores without tempering those results with personal judgment (e.g., performance evaluations and/or other appropriate evaluation criteria).

While the TCT is flawed in the areas of job-relatedness and predictive validity as reported, it is also important to recognize that tests are imperfect. Yet, some testing is needed in a society where public interest in ensuring competence is so great. However, in order to make proper use of these instruments, it is crucial to understand what tests can and cannot do. The TCT is, in fact, a test of knowledge. Testees
are basically required to demonstrate a minimal command of the subject matter in which they intend to teach. In this sense, testing can contribute to the quality of education, but it cannot constitute the total effort of achieving excellence in education. A more comprehensive package is needed, such as curricular reform; review of teacher salaries to attract talented minority teachers; and performance evaluations of all teachers. When legislatures and boards of education address and take decisive action on these matters, then the goals of the education reform movement will be realized, not only in the State of Georgia, but across the nation as well.
VII. RECOMMENDATIONS

Teacher certification testing, as part of a total education reform package, can contribute to improving the quality of education in the Georgia public school systems if corrective action is taken to review testing procedures to meet recognized test standards and legal requirements. The following recommendations are made based upon the review of the literature and the finding of this study:

1. Georgia state officials should review the current procedures used to develop the TCT to ensure compliance with test standards and legal requirements.

2. The TCT should be modified to reflect job relatedness of all test questions, with each item being field tested before appearing on tests.

3. Because results from the TCT are used to make career determining decisions, cut-off scores should be set in a systematic, empirical manner; and other appropriate evaluation criteria (e.g. student outcomes, instructional objectives, informal/systematic observations, as well as performance evaluations) used in addition to the test for personnel decisions.

4. The Georgia Department of Education should offer state-funded remedial programs to teachers who fail the test.
5. Finally, further research in concurrent and predictive validation relative to the TCT should be conducted.
VIII. BIBLIOGRAPHY


Fox, Faye, President of the Georgia Association of Educators. Telephone interview, Atlanta, Georgia, 18 February 1988.


Kramer, Michael, Legal Counsel for the GAE. Telephone interview, Atlanta, Georgia, 19 February 1988.


"Pass or Fail, Teachers Call Test Poor Judge of Competence." Atlanta Journal and Constitution, 30 August 1987, Sec. B, 1.


Staff writer for the Atlanta Journal and Constitution, Atlanta, Georgia, Interview, 22 February 1988.


