DOCUMENT RESUME

ED 248 190 SP 024 016

TITLE Florida Teacher Certification Examination. Technical Report 1981-1982.
INSTITUTION Florida State Dept. of Education, Tallahassee.
PUB DATE 84
NOTE 57p.; For r ed documents, see SP 024 017-020.
PUB TYPE Reports - De. iptive (141) -- Statistical Data (110)

EDRS PRICE MF01/PC03 Plus ostage.
DESCRIPTORS Higher Educatio; Knowledge Level; *Measurement Techniques; *Minimum Competency Testing; Preservice Teacher Education; Quantitative Tests; Reading Tests; Standardized Tests; State Standards; *Teacher Certification; Teaching Skills; *Test Construction; Test Interpretation; Test Reliability; *Test Results; *Test Theory; Test Validity; Writing Skills
IDENTIFIERS *Florida Teacher Certification Examination

ABSTRACT

The Florida Teacher Certification Examination (FTCE) is based upon selected competencies that have been identified as minimal entry-level skills for prospective teachers. A description is given of the four subtests which make up the FTCE: (1) writing--essay on general topics; (2) reading--multiple choice "cloze" procedure on general education passages derived from textbooks, journals, and state publications; (3) mathematics--multiple choice questions on basic mathematics, simple computation, and "real world" problems; and (4) professional education--multiple choice questions on general education (personal, social, academic development, administrative skills, and exceptional student education). This technical report provides information on the test's creation and assembly, administrative procedures, and scoring and reporting. A table presents statistics on the number and percent of students (1981-82) passing all subtests, by education major or program. The psychometric characteristics of validity, reliability, item discrimination, and contrasting group performance of the FTCE are discussed. Appendices provide information on: (1) essential competencies and subskills tested; (2) mathematical illustrations of formulas; (3) security and quality control procedures; and (4) scoring the writing examination. (JD)

TEST BOOK NUMBER

0000

Print your full name here: _____
(last) (first) (middle)

TEST FORM CODE 8A

Florida Teacher Certification Examination

TECHNICAL REPORT 1981 — 1982

1982

State of Florida
Department of Education
Teacher Certification Section
Tallahassee, Florida 32301
Ralph D. Turlington, Commissioner
Affirmative action equal opportunity employer

Questions relative to this report or the Florida Teacher Certification Examination may be directed to the staff by calling 904/488-8198 or by writing:

Student Assessment Section
Department of Education
Knott Building
Tallahassee, Florida 32301

State of Florida
Department of Education
Tallahassee, Florida
Ralph D. Turlington, Commissioner
Affirmative action/equal
opportunity employer

FLORIDA: A STATE OF EDUCATIONAL DISTINCTION. "On a statewide average, educational achievement in the State of Florida will equal that of the upper quartile of states within five years, as indicated by commonly accepted criteria of attainment."

349-020784-MGP-200-1

FLORIDA TEACHER CERTIFICATION EXAMINATION

TECHNICAL REPORT

1981-1982

Student Assessment Section
Bureau of Program Support Services
Division of Public Schools
Department of Education
Tallahassee, Florida 32301

TABLE OF CONTENTS

INTRODUCTION

Background

Each applicant for an initial Florida teacher's certificate must pass the Florida Teacher Certification Examination (FTCE). The FTCE was established by Section 321.17 Florida Statutes and is administered by the Florida Department of Education.

The competencies that form the basis for the Florida Teacher Certification Examination were identified through a study conducted by the Council on Teacher Education (COTE)[1]. As a result of the study, twenty-three Essential Generic Competencies were established upon which to base the Examination and to form a part of the curricular requirements at Florida colleges and universities with approved teacher education programs. Later legislative action combined two of the competencies, numbers six and nineteen, and created an additional competency dealing with education for exceptional students.

An ad hoc task force convened by the Department of Education developed subskills for the identified competencies. The subskills were reviewed and critiqued by various individuals and organizations including a random sample of certified education personnel, statewide professional teacher organizations, and all colleges and universities with approved teacher education programs. The twenty-three Essential Generic Competencies and the subskills are listed in Appendix A.

Test item specifications were written for each subskill. Specifications are rules and parameters for writing test items to measure a particular subskill. They provide information such as the length of the stimuli, the mode of the stimuli (graph, problem situation, mathematical algorithms), the characteristics of the stem (question, statement completion), the characteristics of the correct answer, and the characteristics of the foils. The specifications also include detailed information about the content upon which the tests are based. The complete specifications are contained in the Florida Teacher Certification Examination Bulletin II: The General Education Subtests — Reading, Writing, Mathematics and in the Florida Teacher Certification Examination Bulletin III: The Professional Education Subtest. Copies are available from the Department for a nominal fee.

Passing scores for each subtest were recommended by a panel of judges, all of whom were either current or past members of COTE and who had been involved in the development of the Examination. The panel was made up of classroom

---

[1]COTE was a statutory advisory council appointed by the State Board of Education to advise the Commissioner of Education on all matters dealing with teacher education and certification. COTE was replaced by the Florida Education Standards Commission in 1980.

teachers, school administrators, teacher educators, and community representatives. Passing score recommendations were made to the Commissioner of Education for each subtest. These recommendations were adopted as a rule by the State Board of Education on July 30, 1980.

The operational tasks of preparing test forms, administering the tests, and scoring the answer sheets are completed through an external contract. The contract for these tasks was awarded to the University of Florida Office of Instructional Resources for the three administrations of the 1981-82 school year.

Periodically, contracts are issued for the development of additional test items. New items are needed to maintain a large pool of high quality and secure test items. A large item pool makes it possible to develop alternate forms of the test so that an examinee who retakes a subtest will receive a new set of questions.

All item development is subject to the restrictions of the item specifications. Test development contractors must provide intensive item reviews and conduct pilot tests of the items. Following this, the Department invites a panel of college and university educators to review the new items. This review consists of a critical reading of each item for possible bias, adequate subject content, and adequate technical quality. After the new items have been thoroughly reviewed and revised they are field-tested by imbedding them in a regular test form and administering them to a sample of examinees. The item difficulties are calibrated with latent trait techniques and equated to existing items. Later forms of the FTCE contain the new items.

Description of the Examination

The FTCE is administered three times a year at sites throughout Florida. The test takes an entire Saturday to complete. Examinees usually receive their results within one month. Examinees who fail any part of the FTCE may retake that portion at a subsequent administration. The FTCE is a written test composed of four subtests. The characteristics of the four subtests are summarized in Table 1.

TABLE 1

A Description of the Four Subtests
of the
Florida Teacher Certification Examination

Subtest	Competency Tested	Type of Question	Content	Scoring
Writing	2	Essay, writing production	General topics	Holistic scoring by trained experts
Reading	4	Multiple choice "cloze" procedure	General education passages derived from textbooks, journals, state publications	Objective
Mathematics	5	Multiple choice	Basic mathematics: simple computation and "real world" problems	Objective
Professional Education	6, 7, 9-18, 20-24	Multiple choice (problem solving application level)	General education (personal, social, academic development, administrative skills, exceptional student education)	Objective

The Writing Subtest is scored holistically (general impression marking) by three trained judges. The scoring criteria include an assessment of the following:

1. Using language appropriate to the topic and reader
2. Applying basic mechanics of writing
3. Applying appropriate sentence structure
4. Applying basic techniques of organization
5. Applying standard English usage
6. Focusing on the topic
7. Developing ideas and covering the topic

More detailed information on FTCE administrations is contained in the Florida Teacher Certification Examination Registration Bulletin. This booklet

is available free from Florida school district offices and from the Department of Education.

Four bulletins have been developed to provide information about the development of the FTCE. The subtest and item specifications have been published in Bulletin I: Overview, Bulletin II: The General Education Subtests -- Reading, Writing, Mathematics, and Bulletin III: The Professional Education Subtest. Bulletin IV: The Technical Manual describes the technical adequacy of the first examination. The first three bulletins were distributed to all Florida teacher education institutions and school system personnel offices in the fall of 1979. Bulletin IV, designed primarily for measurement professionals, was published in 1981. An overview of the coverage of the FTCE is provided in Appendix C of Bulletin IV. An annual Technical Report is produced to describe the psychometric characteristics of the three tests administered during each academic year. This report covers the 1981-1982 Examinations.

Rasch Calibration of Items

Calibration of items is conducted using Rasch methodology and the BICAL computer program. The Rasch model bases the probability of a particular score on two parameters, the person's ability and the item's difficulty. The model is expressed as:

$$p\ \{X_{vi} \mid B_v, \delta_i\} = \exp\ [X_{vi}\ (B_v - \delta_i)] \ /\ [1 + \exp\ (B_v - \delta_i)]$$

in which X_{vi} = a score

B_v = person ability

δ_i = item difficulty

Estimates of person ability and item difficulty are obtained using maximum likelihood estimation as described in Wright, Mead, and Bell (BICAL: Calibrating Items with the Rasch Model, 1980).

The process of obtaining item difficulties for new items involves field testing experimental items within regularly administered test forms. Multiple forms for each administration are comprised of sets of scored items in each form and different sets of experimental items. A subset of the scored items forms a common link between forms. The new items are calibrated to the same scale as the regular items. All items are then linked to the base scale of November 1980 by a linking constant. This linking constant is the difference between the average calibration values for the common items in November 1980 and their mean difficulty in the current administration. A description of this process can be found in Ryan (Item Banking, 1980).

Following each administration, the data are randomly divided into three sets of 700 candidates each. Candidates are assigned in sequential order to the appropriate data set. Calibrations are conducted on the data of the candidates in each set and the mean difficulty values across the data are calculated for each item.

TEST COMPOSITION, ADMINISTRATION, AND SCORING

Test Creation and Assembly

The items contained in the Department of Education item bank are calibrated and equated to the base scale established during the April 1980 field test. The items are given identification codes and detailed information on the item usage is maintained including the identification of the form on which each item was used, the difficulty value, item point-biserial correlation, and Rasch fit statistics for each item.

Each test form is designed to ensure that the items (a) fit the item specifications for the skill that they were designed to measure, (b) conform to the test specifications in number and type, and (c) represent a range of difficulty with a mean difficulty approximating zero logits.

A test blueprint is prepared for each form. Items are selected and subjected to content, style, and statistical reviews by the Office of Instructional Resources at the University of Florida and by the Florida Department of Education. Test items are screened for content overlap.

Placement of the items on the test is primarily a function of appearance and content. The order of the items is not related to their difficulty. Items are grouped together if they are similar in editorial style, directions, and question stems.

Experimental items are field-tested within each subtest but are not counted in a candidate's score. When multiple test forms are used, the core of regular (scored) items in each form remains the same for any administration. Test forms are spiralled so that each test center receives approximately the same number of each form. In this way, all experimental items are field-tested by at least 400 candidates who represent a cross-section of the people who take the Examination.

Once the form has been approved, the scoring key is verified. Staff members from the Department of Education, the Office of Instructional Resources, and three teachers from the public schools take the Examination. These persons are also asked to identify any ambiguous items or confusing directions.

Camera-ready copy is prepared by a test specialist and a graphic artist. Attention is paid to the proper placement of items to provide workspace where necessary. The camera-ready copy is again critiqued by the staff in the Department of Education and the Office of Instructional Resources. Corrections are made, the copy is sent to the printer, and a final check of the proof is made before the tests are printed.

Administration Procedures

Examination Dates, Times, and Locations

The FTCE is administered in the fall, winter, and summer of each year. Administration dates for 1981-82 were October 31, 1981; February 27, 1982; and July 10, 1982. Candidates were permitted to take all four subtests or any subtest previously not passed. Thirteen locations in the state were designated as testing areas. Specific sites within each area were selected as test centers. These centers were selected from the pool of established centers for the administration of standardized examinations. Designated test locations for the 1981-82 administrations were:

1. Pensacola
2. Tallahassee
3. Gainesville
4. Jacksonville
5. St. Petersburg
6. Tampa
7. Sarasota
8. Miami
9. Fort Myers
10. Orlando
11. Boca Raton
12. DeLand
13. Lakeland

All test centers were inspected to ensure that the rooms met the required specifications for lighting, seating capacity, storage facilities, air conditioning, and protection from outside disturbances. All facilities were able to accommodate handicapped candidates.

The test schedule is divided into morning and afternoon sessions. Testing time is fixed but allows adequate time for candidates to complete all sections of the Examination. Candidates may continue to the Reading Subtest after they finish the Mathematics section. The schedule for each subtest is listed below:

Writing	45 minutes	9:00 a.m. - 9:45 a.m.
Mathematics	70 minutes	10:00 a.m. - 11:10 a.m.
Reading	50 minutes	11:10 a.m. - 12:00 noon
Break	60 minutes	12:00 noon - 1:00 p.m.
Professional Education	150 minutes	1:30 p.m. - 4:00 p.m.

A security plan has been developed and implemented for the program. Refer to Appendix C for further information about security and quality control.

Special arrangements are made as necessary for handicapped candidates. A Braille version of the Examination is available. Typewriters or a flexible time schedule is permitted for handicapped candidates.

Test Manuals

Uniform testing procedures were established for use at all centers throughout the state. Documentation of the procedures is available in the Test Administration Manual for the program. The administration manual includes the following topics:

1. Duties of the Test Center Personnel
2. Receipt and Security of Test Materials
3. Admission, Identification and Seating of Candidates
4. On-site Test Administration Practices and Policies
5. Timing of Subtest Sections
6. Instructions for Completing Answer Documents
7. Special Arrangements for Handicapped Students

Additional information to candidates is found in several other sources. Candidates are notified about Examination requirements, locations, and procedures in the Registration Bulletin. Specific directions to candidates about the assigned test center and necessary supplies are printed on the Admission Ticket.

Scoring and Reporting

Scoring

The scoring process begins with a hand edit of the answer sheets, followed with the scanning of an initial set of sheets to verify the accuracy of the scanner, the key, and the scoring programs. The remaining sheets are scanned, and the data are divided into sets of 700 candidates. Three data sets are drawn using a systematic random sampling method for the calibration of items using the Rasch methodology. The items are adjusted to the base scale established by the April 1980 field test. A score table of equivalent raw scores to ability logits is calculated and used to determine the ability logits for the remaining candidates. Each person's score in ability logits is then transformed to a scale score with 200 as the minimum passing score. For a discussion of the procedures used to establish the cutting score see the technical discussion in Bulletin IV.

The essay is rated by three readers who use a four-point scale defined in State Board of Education rules. The resulting scores range from three to twelve points. The passing standard is set at six points. Details of the criteria for the rating of essays are available in Bulletin II.

Reporting

The reports generated for each administration include a candidate report and score interpretation guide, reports for institutions, and state-level reports.

Candidate reports indicate whether or not a test is passed; scaled scores are reported only for tests failed. Scores above the passing standard are not reported. However, candidates who fail one or more tests are provided their scale score for each subtest failed. A detailed analysis of performance is provided to individuals who fail the Professional Education Subtest.

The reports generated for the institutions and the state are listed below:

1. Number and Percent Passing for:

   a. Each subtest
   b. All four subtests
   c. Three, two, one or no subtests

2. Number, Percent Passing and Mean Scores for Each Subtest and the Total Examination by All Candidates and:

   a. First-time candidates
   b. Re-take candidates
   c. Vocational candidates
   d. Non-vocational candidates
   e. Florida candidates
   f. Non-Florida candidates
   g. Florida candidates from approved degree programs
   h. Florida candidates from non-approved programs
   i. Sex and ethnic categories

3. Number and Percent of Candidates by Florida Institutions and by Programs, Passing All Subtests and Each Subtest

4. Number and Percent of Candidates Passing All Subtests and Each Subtest by Program Statewide

5. Frequency Distribution for All Candidates for Each Subtest by Sex and Ethnic Category

6. Frequency Distribution for Each Subtest for Florida Institution

Statistical analyses of data are reported in the sections on the psychometric characteristics of the Examination.

TEST RESULTS FOR 1981-82

Results for the three test administrations[2] in this report are summarized in this chapter. The overall passing rates are shown in Table 2. As can be seen from the data there are no differences between first-time and all candidates for the first two administrations, and two percentage points higher performance for first-time takers in February 1982. The February 1982 administration was the first one that showed a large eff · · from retakers. This effect is not unexpected.

Table 2

Percent of Candidates Passing All Subtests
of the
Florida Teacher Certification Examination
August 1981 - February 1982

	First-Time Candidates	All Candidates
August 1981	80	80
October 1981	84	84
February 1982	86	84

Table 3 through 5 on the following pages show the number and percent passing each subtest and all subtests for: (a) approved program candidates; (b) non-approved program candidates; and (c) vocational technical candidates.

[2]The August 1981 administration was a part of the data for the 1980-81 Technical Report. It is also in this report because of a decision to make the Technical Report coincide with the same test administrations that are used in calculating the eighty percent report for approved programs. The eighty percent performance report is calculated from the summer, fall, and winter administrations.

Table 3

Florida Teacher Certification Examination Number and Percent Passing All Subtests and Each Subtest by Program Statewide for Approved Program Candidates

August 1981; October 1981; and February 1982 Administrations

PROGRAM	TOTAL EXAM			SUBTESTS READ			WRITE			MATH			PROFED		
	T	N	%	T	N	%	T	N	%	T	N	%	T	N	%
1 ADMINISTRATION/SUPERVISION	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	2	2	100.0
4 ART EDUCATION	81	70	86.4	82	78	95.1	81	78	96.3	84	76	90.5	82	80	97.6
5 BIBLE	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0
6 BUSINESS EDUCATION	59	50	84.7	60	54	90.0	59	56	94.9	60	58	96.7	68	64	94.1
12 EARLY CHILDHOOD ED.	24	21	87.5	24	22	91.7	24	23	95.8	24	23	95.8	24	24	100.0
13 VARYING EXCEPTIONALITIES	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
14 HEARING DISABILITIES	6	5	83.3	6	6	100.0	6	5	83.3	6	6	100.0	6	6	100.0
15 VISUAL DISABILITIES	10	8	80.0	10	10	100.0	10	9	90.0	10	9	90.0	10	10	100.0
17 MENTAL RETARDATION	175	154	88.0	177	168	94.9	175	169	96.6	177	162	91.5	175	175	100.0
18 SPEECH CORRECTION	17	17	100.0	17	17	100.0	17	17	100.0	18	18	100.0	17	17	100.0
20 ELEMENTARY EDUCATION	1401	1242	88.6	1413	1325	93.8	1412	1346	95.3	1429	1308	91.0	1405	1361	96.9
21 ENGLISH	129	124	96.1	129	128	99.2	129	129	100.0	129	125	96.9	129	127	98.4
22 GUIDANCE	22	20	90.9	22	22	100.0	22	22	100.0	23	21	91.3	22	22	100.0
23 HEALTH EDUCATION	20	18	90.0	20	20	100.0	20	19	95.0	20	19	95.0	20	20	100.0
24 VOCATIONAL HOME ECONOMICS	36	32	88.9	36	34	94.4	36	34	94.4	36	33	91.7	36	35	97.2
26 INDUSTRIAL ARTS	17	12	70.6	18	15	83.3	19	13	68.4	18	16	88.9	17	16	94.1
29 TRANSPORTATION	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
33 GRAPHIC ARTS	5	5	100.0	5	5	100.0	5	5	100.0	5	5	100.0	5	5	100.0
35 FRENCH	9	8	88.9	9	9	100.0	10	10	100.0	9	8	88.9	9	9	100.0
36 SPANISH	19	15	78.9	19	17	89.5	19	17	89.5	19	15	78.9	19	16	84.2
37 LATIN	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
38 GERMAN	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
43 EDUCATIONAL MEDIA SPECIALIST	20	17	85.0	20	20	100.0	20	20	100.0	20	17	85.0	20	20	100.0
44 MATHEMATICS	72	66	91.7	73	70	95.9	72	67	93.1	72	71	98.6	72	71	98.6
45 MUSIC EDUCATION	113	95	84.4	113	104	92.0	111	105	94.6	113	105	92.9	113	107	94.7
48 PHYSICAL EDUCATION	288	214	74.3	288	251	87.2	287	254	88.5	290	249	85.9	288	265	92.0
49 SCIENCE	9	9	100.0	9	9	100.0	9	9	100.0	9	9	100.0	9	9	100.0
50 CHEMISTRY	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0
51 PHYSICS	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
52 BIOLOGY	35	33	94.3	36	35	97.2	35	34	97.1	35	34	97.1	35	35	100.0
56 SOCIAL STUDIES	104	98	94.2	105	105	100.0	105	100	95.2	108	107	99.1	104	104	100.0
57 HISTORY	11	7	63.6	11	11	100.0	11	8	72.7	11	9	81.8	11	11	100.0
58 POLITICAL SCIENCE	6	4	66.7	6	6	100.0	6	5	83.3	6	5	83.3	6	6	100.0
60 SOCIOLOGY	3	0	0.0	3	2	66.7	3	1	33.3	4	2	50.0	3	0	0.0
62 SPEECH	4	4	100.0	4	4	100.0	4	4	100.0	4	4	100.0	4	4	100.0
63 VISITING TEACHER	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0
70 ADMINISTRATION, ADULT ED	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
74 PSYCHOLOGY	8	7	87.5	8	7	87.5	9	9	100.0	8	7	87.5	8	7	87.5
75 INSTRUMENTAL MUSIC	3	3	100.0	3	3	100.0	3	3	100.0	3	3	100.0	3	3	100.0
110 ADMINISTRATION	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
132 HUMANITIES	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
134 TECHNICAL EDUCATION	2	1	50.0	2	2	100.0	2	1	50.0	2	2	100.0	2	2	100.0
145 SPECIALIST IN SCHOOL PSY.	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
146 READING	6	6	100.0	6	6	100.0	6	6	100.0	6	6	100.0	6	6	100.0
153 MUSIC VOCAL	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
178 EARLY CHILD. ED/ELEM. ED.	109	94	86.2	109	101	92.7	110	104	94.5	113	100	88.5	110	109	99.1
184 ELEM/EMOTIONAL DISTURBANCE	5	4	80.0	5	4	80.0	5	5	100.0	5	5	100.0	5	5	100.0
185 ELEM/HEARING DISABILITY	21	20	95.2	21	21	100.0	21	21	100.0	21	20	95.2	21	21	100.0
186 ELEM/MENTAL RETARDATION	3	3	100.0	3	3	100.0	3	3	100.0	3	3	100.0	3	3	100.0
191 ELEM/MID SPEC. LEARN DISABIL	6	6	100.0	6	6	100.0	6	6	100.0	6	6	100.0	6	6	100.0
194 EMOTIONAL/MENTAL/LEARN/VARYING	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
195 EMOTIONAL/MENTAL/SPECIFIC	22	18	81.8	22	21	95.5	22	22	100.0	22	19	86.4	22	22	100.0
197 EMOTIONAL/SPECIFIC/VARYING	11	10	100.0	13	13	100.0	10	10	100.0	14	14	100.0	18	18	100.0
201 EMOTIONAL DISTURBANCE	64	56	87.5	65	62	95.4	65	60	92.3	65	62	95.4	64	63	98.4
202 SPECIFIC LEARN. DISABILITIES	165	156	94.5	166	164	98.8	165	161	97.6	168	163	97.0	165	163	98.8
205 HEALTH OCCUPATIONS ED.	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
206 ENGLISH/FRENCH	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
207 ENGLISH/SOCIAL STUDIES (7-9)	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0
210 ENGLISH/SPEECH	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0
215 FRENCH/SPANISH	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
224	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0
227 INDUSTRIAL ED/TECH ED	4	4	100.0	4	4	100.0	4	4	100.0	4	4	100.0	4	4	100.0
247 SCIENCE/MATH	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
255 JUNIOR HIGH/MIDDLE SCH/ENG/SOC	3	3	100.0	3	3	100.0	3	3	100.0	3	3	100.0	3	3	100.0
300 COMMUNICATION	2	1	50.0	2	2	100.0	2	2	100.0	2	1	50.0	2	2	100.0
302 EDUCATION	9	8	88.9	10	9	90.0	9	9	100.0	10	8	80.0	11	11	100.0
306 HEALTH SERVICES	3	3	100.0	3	3	100.0	3	3	100.0	3	3	100.0	3	3	100.0

Table 3 cont'd.

PROGRAM	TOTAL EXAM N	TOTAL EXAM %	SUBTESTS READ N	SUBTESTS READ %	WRITE N	WRITE %	MATH N	MATH %	PROF ED N	PROF ED %
370 LIBRARY SCIENCE	3	66.7	3	100.0	3	100.0	2	66.7	3	100.0
374 PUBLIC AFFAIRS & SERVICE	14	100.0	14	100.0	14	100.0	14	100.0	14	100.0
375 SOCIAL SCIENCES	4	75.0	4	100.0	3	75.0	3	75.0	4	100.0
379 SECONDARY SCIENCE	3	100.0	3	100.0	3	100.0	3	100.0	3	100.0
680 INDUSTRIAL EDUCATION	6	66.7	6	83.3	6	83.3	6	83.3	6	100.0
632 DISTRIBUTIVE EDUCATION	7	100.0	7	100.0	7	100.0	7	100.0	7	100.0

Table 4

Florida Teacher Certification Examination Number and Percent Passing
All Subtests and Each Subtest by Program Statewide
Non-Approved Program Candidates

August 1981; October 1981; and February 1982 Administrations

PROGRAM		TOTAL EXAM			SUBTESTS READ			WRITE			MATH			PROF ED		
		T	N	%	T	N	%	T	N	%	T	N	%	T	N	%
1	ADMINISTRATION/SUPERVISION	2	2	100.0	3	2	66.7	4	3	75.0	3	3	100.0	2	2	100.0
2	VOCATIONAL AGRICULTURE	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0
3	GEN AG	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0
4	ART EDUCATION	47	38	80.9	47	45	95.7	48	42	87.5	50	44	88.0	47	46	95.7
5	BIBLE	6	6	100.0	6	6	100.0	6	6	100.0	7	7	100.0	6	6	100.0
6	BUSINESS EDUCATION	28	18	64.3	30	24	80.0	31	28	90.3	29	23	79.3	28	24	85.7
9	BOOKKEEPING	17	16	94.1	17	16	94.1	17	16	94.1	17	17	100.0	17	16	94.1
12	EARLY CHILDHOOD ED.	13	12	92.3	13	13	100.0	13	13	100.0	13	12	92.3	13	13	100.0
14	HEARING DISABILITIES	8	8	100.0	8	8	100.0	8	8	100.0	8	8	100.0	8	8	100.0
15	VISUAL DISABILITIES	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
17	MENTAL RETARDATION	13	12	92.3	13	12	92.3	13	13	100.0	15	15	100.0	13	13	100.0
18	SPEECH CORRECTION	58	54	93.1	59	58	98.3	58	57	98.3	63	60	95.2	59	59	100.0
20	ELEMENTARY EDUCATION	251	187	74.5	257	209	81.3	257	217	84.4	257	204	79.4	256	220	85.9
21	ENGLISH	147	128	87.1	149	139	93.3	147	140	95.2	156	138	89.6	150	139	92.7
22	GUIDANCE	19	18	94.7	19	18	94.7	20	19	95.0	20	19	95.0	20	20	100.0
23	HEALTH EDUCATION	4	3	75.0	8	8	100.0	8	8	100.0	9	9	100.0	8	6	75.0
24	VOCATIONAL HOME ECONOMICS	10	7	70.0	10	9	90.0	10	10	100.0	11	8	72.7	10	9	90.0
25	GEN HOME ECONOMICS	10	7	70.0	10	8	80.0	10	8	80.0	10	7	70.0	10	8	80.0
26	INDUSTRIAL ARTS	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
31	JOURNALISM	19	15	78.9	19	18	94.7	19	17	89.5	20	15	75.0	19	18	94.7
35	FRENCH	8	7	87.5	8	8	100.0	8	8	100.0	8	7	87.5	8	8	100.0
36	SPANISH	26	15	57.7	27	17	63.0	27	17	63.0	28	21	75.0	28	20	71.4
38	GERMAN	4	4	100.0	4	4	100.0	4	4	100.0	4	4	100.0	4	4	100.0
43	EDUCATIONAL MEDIA SPECIALIST	17	14	82.4	17	14	82.4	18	17	94.4	18	16	88.9	17	17	100.0
44	MATHEMATICS	62	49	79.0	62	58	93.5	64	54	84.4	63	61	96.8	62	56	90.3
45	MUSIC EDUCATION	47	38	80.9	48	45	93.8	47	42	89.4	48	42	87.5	47	42	89.4
46	PHYSICAL EDUCATION	68	36	52.9	68	46	67.6	68	45	66.2	68	48	70.6	68	49	72.1
49	SCIENCE	10	8	80.0	10	9	90.0	10	8	80.0	10	9	90.0	10	10	100.0
50	CHEMISTRY	13	11	84.6	13	12	92.3	14	12	85.7	13	13	100.0	13	11	92.3
51	PHYSICS	4	3	75.0	4	4	100.0	4	3	75.0	4	4	100.0	4	4	100.0
52	BIOLOGY	63	57	90.5	64	60	93.8	63	61	96.8	64	60	93.8	64	61	95.3
56	SOCIAL STUDIES	37	25	67.6	38	31	81.6	39	31	79.5	40	32	80.0	37	31	83.8
57	HISTORY	63	47	74.6	63	58	92.1	63	59	93.6	63	51	81.0	63	57	90.5
58	POLITICAL SCIENCE	44	33	75.0	44	42	95.5	44	41	93.2	44	35	79.5	44	42	95.5
59	ECONOMICS	11	9	81.8	12	10	83.3	11	11	100.0	12	9	75.0	12	10	83.3
60	SOCIOLOGY	77	51	66.2	79	66	83.5	79	72	91.1	83	62	74.7	79	65	82.3
61	GEOGRAPHY	3	3	100.0	3	3	100.0	3	3	100.0	3	3	100.0	3	3	100.0
62	SPEECH	17	16	94.1	17	17	100.0	17	17	100.0	18	17	94.4	18	18	100.0
63	VISITING TEACHER	30	22	73.3	30	24	80.0	31	27	87.1	31	23	74.2	31	27	87.1
70	ADMINISTRATION, ADULT ED	2	1	50.0	2	1	50.0	2	2	100.0	2	2	100.0	2	2	100.0
78	PSYCHOLOGY	107	91	85.0	108	104	96.3	108	101	93.5	109	95	87.2	108	100	92.6
114	ADMINISTRATION	2	2	100.0	3	3	100.0	3	2	66.7	3	3	100.0	3	2	66.7
132	HUMANITIES	10	8	80.0	10	10	100.0	10	9	90.0	10	9	90.0	10	10	100.0
145	SPECIALIST IN SCHOOL PSY.	3	3	100.0	3	3	100.0	3	3	100.0	3	3	100.0	3	3	100.0
146	READING	3	3	100.0	3	3	100.0	3	3	100.0	3	3	100.0	3	3	100.0
147	VOCAL MUSIC	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
148	SCHOOL FOOD SERVICE	5	5	100.0	5	5	100.0	5	5	100.0	5	5	100.0	5	5	100.0
150	DRAMA	8	6	75.0	8	7	87.5	8	7	87.5	8	6	75.0	8	7	87.5
175	ADMIN-SUPERVIS/EMO DISTURB	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
178	EARLY CHILD. ED/ELEM. ED.	4	4	100.0	4	4	100.0	4	4	100.0	4	4	100.0	4	4	100.0
197	EMOTIONAL/SPECIFIC/VARYING	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
201	EMOTIONAL DISTURBANCE	11	10	90.9	11	10	90.9	11	10	90.9	12	11	91.7	11	10	90.9
232	SPECIFIC LEARN. DISABILITIES	7	5	71.4	7	7	100.0	7	7	100.0	7	6	85.7	7	5	71.4
315	FRENCH/SPANISH	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
319	GERMAN/SPANISH	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
347	SCIENCE/MATH	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	2	2	100.0
355	JUNIOR HIGH/MIDDLE SCH/ENG/SOC	18	18	100.0	18	18	100.0	18	18	100.0	18	18	100.0	19	19	100.0
356	JUNIOR HIGH/MIDDLE SCH/MATH/SC	0	0	0.0	1	1	100.0	1	0	0.0	0	0	0.0	0	0	0.0
357	HEARING DISABILITIES/MENTAL RC	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
359	BUSINESS,COMMERCE,MANAGEMENT	6	4	66.7	7	7	100.0	6	5	83.3	7	5	71.4	7	7	100.0
360	COMMUNICATION	2	2	100.0	2	2	100.0	2	2	100.0	3	2	66.7	2	2	100.0
361	COMPUTER & INFO SCIENCE	2	1	50.0	2	2	100.0	2	2	100.0	2	2	100.0	2	1	50.0
362	EDUCATION	13	10	76.9	14	12	85.7	15	14	93.3	16	14	87.5	13	10	76.9
364	FINE & APPLIED ARTS	2	1	50.0	2	2	100.0	2	2	100.0	3	1	33.3	2	1	50.0
365	FOREIGN LANGUAGES	3	2	66.7	3	2	66.7	3	2	66.7	3	2	66.7	3	2	66.7
366	HEALTH SERVICES	11	10	90.9	11	11	100.0	12	11	91.7	13	11	84.6	11	10	90.9
367	HOME ECONOMICS	1	0	0.0	1	1	100.0	1	1	100.0	1	1	100.0	1	0	0.0

Table 4 cont'd.

PROGRAM	TOTAL EXAM			SUBTESTS READ			WRITE			MATH			PROF ED		
	T	N	%	T	N	%	T	N	%	T	N	%	T	N	%
368 LAW	6	4	66.7	7	4	57.1	6	4	66.7	8	6	75.0	6	6	100.0
369 LETTERS	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
370 LIBRARY SCIENCE	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
373 PSYCHOLOGY	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0
374 PUBLIC AFFAIRS & SERVICE	7	5	71.4	7	5	71.4	5	5	71.4	7	6	85.7	7	6	85.7
375 SOCIAL SCIENCES	3	3	100.0	4	4	100.0	3	3	100.0	3	3	100.0	3	3	100.0
376 INTERDISCIPLINARY STUDIES	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
377 DEGREE VOCATIONAL	9	8	88.9	9	9	100.0	10	9	90.0	9	9	100.0	10	8	80.0
378 SECONDARY SCIENCE	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
869 INDUSTRIAL EDUCATION	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0

Table 5

Florida Teacher Certification Examination Number and Percent Passing
All Subtests and Each Subtest by Program Statewide for
Vocational Technical Candidates

August 1981; October 1981, and February 1982 Administrations

	TOTAL #	PASSED	% PASSED
TOTAL TEST	422	179	42%
READING	369	263	71%
MATH	379	273	72%
PROF. EDUCATION	354	265	75%
WRITING	361	247	68%

Table 4 cont'd.

PROGRAM		TOTAL EXAM			SUBTESTS READ			WRITE			MATH			PROF ED		
			N	%		N	%		N	%		N	%		N	%
368	LAW	6	4	66.7	7	4	57.1	6	4	66.7	8	6	75.0	6	6	100.0
369	LETTERS	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
370	LIBRARY SCIENCE	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
373	PSYCHOLOGY	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0	2	2	100.0
374	PUBLIC AFFAIRS & SERVICE	7	5	71.4	7	5	71.4	7	5	71.4	7	6	85.7	7	6	85.7
375	SOCIAL SCIENCES	3	3	100.0	4	4	100.0	3	3	100.0	4	4	100.0	3	3	100.0
376	INTERDISCIPLINARY STUDIES	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
377	DEGREE VOCATIONAL	9	8	88.9	9	9	100.0	10	8	80.0	9	9	100.0	10	8	80.0
379	SECONDARY SCIENCE	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0
889	INDUSTRIAL EDUCATION	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0	1	1	100.0

Table 5

Florida Teacher Certification Examination Number and Percent Passing
All Subtests and Each Subtest by Program Statewide for
Vocational Technical Candidates

August 1981; October 1981; and February 1982 Administrations

	TOTAL #	# PASSED	% PASSED
TOTAL TEST	422	179	42%
READING	369	263	71%
MATH	379	273	72%
PROF. EDUCATION	354	265	75%
WRITING	361	247	68%

PSYCHOMETRIC CHARACTERISTICS

The psychometric characteristics of validity, reliability, item discrimination, and contrasting group performance of the Florida Teacher Certification Examination (FTCE) will be addressed in this section. Knowledge of the psychometric characteristics of assessment tests is necessary for evaluating the tests.

Validity

Validity refers to the relevance of inferences that are made from test scores or other forms of assessment. The validity of a test can be defined as the degree to which a test measures what it was intended to measure. Validity is not an all-or-none characteristic, but a matter of degree. Validity is needed to ensure the accuracy of information that is inferred from a test score.

Specific types of validation techniques traditionally used to summarize educational and psychological test use -- criterion-related validity (predictive and concurrent), content validity, and construct validity -- are described in Standards for Educational and Psychological Measurement (APA, 1974, pp. 26-31). For the FTCE, the primary validity issue that must be addressed is the question of content validity. Content validity demonstrates that test behaviors constitute a representative sample of behaviors in a desired performance domain. The intended domain of the FTCE is that of entry-level skills as identified in the statute requiring the Examination as a basis for certification. This statute (231.17, F.S.) provides that

> Beginning July 1, 1980 ... each applicant for initial certification shall demonstrate, on a comprehensive written examination and through such other procedures as may be specified by the state board, mastery of those minimum essential generic and specialization competencies and other criteria as shall be adopted into rules by the state board.

The statute addresses only the status at certification and does not require that inferences be made from test scores to future success as a classroom teacher. No claims have been made with regard to measurement of specific aptitudes or traits, and no attempt has been made to establish relationships between the FTCE and independent concurrent or future criteria. It is only claimed that the test adequately measures the skills for which it was developed. The construct and criterion-related validation approaches are not appropriate to the validity issues related to development and use of the FTCE.

The content validity of the FTCE rests upon the procedures used to describe and develop test items and content areas. The intended coverage of the test was determined by a process involving professional consensus to (1) identify competencies which should be demonstrated as a condition for certification, and (2) identify subskills associated with each competency. The procedures by which the intended coverage was identified included surveys of the profession, reviews

by the Council on Teacher Education (COTE), reviews by the ad hoc COTE task force, and reviews by teachers and other professional personnel.

The general procedures used in test development were as follows:

1. The intended test coverage was identified and explicated. Competencies and subskills associated with each competency were identified and validated.

2. Test item specifications were developed and validated.

3. Draft items were written according to test item specifications and pilot-tested on a small sample of senior students preparing to be teachers.

4. The final item review consisted of (a) a review by a special panel comprised of classroom teachers, teacher educators, and administrators, and (b) item field-testing with seniors who were in teacher education programs. This was followed by another review by Department of Education staff. Items were subsequently placed in the item bank for future use.

5. Field-test data were reviewed by Department of Education staff. Items that did not perform well were deleted from the item bank or revised and field-tested again.

For the final item review process outlined in the fourth step, the items were divided by test area and reviewers were divided by area of expertise. The process included a review of item content, group differences in performance, and technical quality. Bulletin IV (pp. 13-17) contains further information about the development and review of test items.

In summary, the validity of the Examination has been well established as a result of (1) the extensive involvement of education professionals in the identification and explication of the necessary competencies and their associated subskills, (2) the precise item specifications which guided the item writers, and (3) the reviews of the items and the competencies/skills that they were designed to measure.

Reliability of Test Scores

Reliability refers to the consistency between two measures of the same performance domain. Although reliability does not ensure validity, it limits the extent to which a test is valid for a particular purpose. The main reliability consideration for the FTCE multiple-choice tests (Reading, Mathematics, and Professional Education) is the reliability of an individual's score. For the Writing test, a production writing sample, the reliability consideration is the reliability of the judges' ratings. The data in this section refer to the three FTCE administrations between October 1981 and July 1982. For information about field test reliability data, refer to Bulletin IV (1981).

Reliability of Multiple-choice Tests

A test score is comprised of a "true" score ("domain" score) and an "error" score. If an individual took several forms of a test, all constructed by sampling from the defined item domain, scores of the various test forms would not vary except as a result of random errors associated with item sampling errors and changes within an individual from one test to another such as attention, fatigue, or interest.

Reliability evidence is generally of two types: (a) internal consistency, which is essential if items are viewed as a sample from a relatively homogeneous universe; and (b) consistency over time, which is important for tests that are used for repeated measurement. For the FTCE, the primary reliability issue is that of internal consistency. Since one form of the test is administered to examinees at each administration, the reliability concern is that of consistency of items within that particular test (homogeneity of items). A test can be regarded as composed of as many parallel tests as the test has items, and every item is treated as parallel to the other items. In such a case, the appropriate reliability index is the Kuder-Richardson Formula 20 (KR-20) index. The KR-20 formula is shown in Appendix B.

The KR-20 index estimates the internal-consistency reliability of a test from statistical data on individual items. Separate KR-20 coefficients are calculated for the Reading, Mathematics, and Professional Education subtests for each FTCE administration. A high coefficient indicates that a test accurately measures some characteristic of persons taking it and means that the individual test items are highly correlated. The subtest KR-20 coefficients for the three 1981-1982 test administrations were above .78, indicating that the individual test items were highly consistent measures of the three subject areas assessed. Refer to Table 6 for the KR-20 coefficients.

Table 6

Kuder-Richardson Coefficients

	Math	Reading	Professional Education
October 1981	.88	.87	.83
February 1982	.84	.85	.79
July 1982	.87	.87	.81

Reliability of the passing standards for the objective tests is estimated with the Brennan-Kane (B-K) Index of Dependability. This index is an estimate of the consistency of test scores in classifying examinees as masters or nonmasters of the minimal performance standards. The high B-K coefficients of the tests (refer to Table 7) indicate that the candidates' scores are consistent with their classification as masters or nonmasters. Refer to Appendix B for the statistical formula for the B-K Index.

TABLE 7

Brennan-Kane Indices

	Math	Reading	Professional Education
October 1981	.94	.94	.96
February 1982	.96	.94	.96
July 1982	.95	.94	.95

Reliability of Scoring of the Writing Subtest

The major reliability consideration for the Writing test is the inter-judge reliability of ratings. The Writing test is a production writing sample that addresses one of two specific topics. The essays are rated independently by three judges with a referee to reconcile discrepant scores. Original reliability data were obtained from a study in which essays were written by 360 teacher education students at two universities. Raters were trained by the same procedures which are being used in the actual test administrations. The reliability of the scoring process is monitored at the University of Florida for each test administration. (Refer to Appendix D for additional information about the scoring of the Writing test.)

Two approaches are used to estimate the reliability. First, four indices of inter-rater agreement are computed. These four indices are: (a) percent complete agreement; (b) average percent of two of the three raters agreeing; (c) average percent agreement by pairs as to pass/fail; and (d) percent complete agreement about pass/fail. The second approach for reliability estimation is the calculation of coefficient alpha for the raters and the rating team. This coefficient indicates the expected correlation between the ratings of the team on this task and those of a hypothetical team of similarly comprised and similarly trained raters doing the same task. Field test inter-rater reliability data and coefficient alpha for the inter-rater reliabilities are reported in Tables 3.4 and 3.5 of Bulletin IV (pp. 22-23). Refer to Table 8 for rater reliability data for the 1981-19 FTCE administrations.

TABLE 8

Percentage of Rater Agreement for FTCE
Writing Test

		October 1981	February 1982	July 1982
Index 1 - % Complete Agreement		46.44	38.55	38.60
Index 2 - Average % Two of the Three Raters Agreeing		99.87	100.00	99.81
Index 3 - Average % Agreement by Pairs as to Pass/Fail		98.10	96.75	96.72
Index 4 - % Complete Agreement About Pass/Fail		97.15	95.12	95.08
Coefficient Alpha	Topic 1	.84	.87	.82
	Topic 2	.85	.85	.86

Examination of the reliability data for the Writing test indicates that the level of reliability achieved by the rating teams met acceptable standards for such ratings.

Discrimination

Item analysis for the FTCE includes examination of the items' capacity to differentiate between ability groups and the evaluation of response patterns to the individual items. The item analysis indices used are item difficulty level, item discrimination index, and point-biserial correlation coefficients.

Item difficulty level -- the percentage of examinees who answer each item correctly -- is calculated for each item. These percentages provide important information because items in the moderate range of difficulty differentiate relatively more examinees from each other than do extremely easy or extremely difficult items.

Related to the item difficulty level is the item discrimination index (see page 32) which is the extent to which each item contributes to the total test in terms of discriminating between the high and low achievers with regard

to the total test score. Any item that is below .20 on this index is evaluated for content and ambiguity of wording. Items that appear to be flawed are revised or eliminated. The ranges for item difficulty level and corresponding item discrimination indices are reported in Table 9.

The number and percent of examinees who select each alternative response (foil) were reported for each item in the multiple-choice tests. This foil analysis permits further evaluation of response patterns to the individual items and provides useful information about variations in response performance by different groups. These data are provided to the Department of Education staff and appropriate subcontractors and are not reported in this document.

Point-biserial correlation coefficients indicate the extent to which examinees with high test scores tend to answer an item correctly and those with low test scores tend to miss an item. While the item discrimination index is based on the performance of high and low achievers, the point-biserial coefficient includes the entire range of scores in the correlation, thereby indicating the item-total correlation or the extent to which an item score correlates with all other items measuring a particular subject area. Statistical formulas for these indices are listed in Appendix B.

TABLE 9

Frequency of Items within Specific Item Difficulty and Item Discrimination Ranges

MATH

Item Difficulty Ranges	Item Discrimination Range										
	.10 and below	.11-.20	.21-.30	.31-.40	.41-.50	.51-.60	.61-.70	.71-.80	.81-.90	.91-1.00	TOTAL
.81-1.00	15	33	27	14	9	0	0	0	0	0	98
.61- .80	0	0	3	4	12	7	2	0	0	0	28
.41- .60	0	0	0	0	2	2	0	0	0	0	4
.21- .40	0	0	0	0	0	0	0	0	0	0	0
.0- .20	0	0	0	0	0	0	0	0	0	0	0
TOTAL	15	33	30	18	23	9	2	0	0	0	130

READING

Item Difficulty Ranges	Item Discrimination Range										
	.10 and below	.11-.20	.21-.30	.31-.40	.41-.50	.51-.60	.61-.70	.71-.80	.81-.90	.91- 1.00	TOTAL
.81-1.00	139	50	19	6	0	0	0	0	0	0	214
.61- .80	0	0	3	7	6	5	0	0	0	0	21
.41- .60	1	0	0	1	1	2	0	0	0	0	5
.21- .40	0	0	0	0	0	0	0	0	0	0	0
.0- .20	0	0	0	0	0	0	0	0	0	0	0
TOTAL	140	50	22	14	7	7	0	0	0	0	240

PROFESSIONAL EDUCATION

Item Difficulty Ranges	Item Discrimination Range										
	.10 and below	.11-.20	.21-.30	.31-.40	.41-.50	.51-.60	.61-.70	.71-.80	.81-.90	.91-1.00	TOTAL
.81-1.00	32	42	28	2	0	0	0	0	0	0	104
.61- .80	2	16	24	27	15	1	0	0	0	0	85
.41- .60	3	1	17	18	4	1	0	0	0	0	44
.21- .40	1	1	2	2	0	0	0	0	0	0	6
.0- .20	0	0	1	0	0	0	0	0	0	0	1
TOTAL	38	60	72	49	19	2	0	0	0	0	240

Means and ranges for the point-biserial correlation coefficients for the three 1981-1982 administrations are reported in Tables 10 and 11.

TABLE 10

Mean Point-Biserial
Correlation Coefficients

	Math	Reading	Professional Education
October 1981	.38	.32	.27
February 1982	.37	.31	.25
July 1982	.40	.34	.25

TABLE 11

Point-Biserial Correlation Coefficients Between
Correct Item Response and Subtest Score

Range of Point-Biserial Coefficients	Math	Reading	Professional Education
.90-.99	0	0	0
.80-.89	0	0	0
.70-.79	0	0	0
.60-.69	1	0	0
.50-.59	18	1	0
.40-.49	38	55	14
.30-.39	42	93	64
.20-.29	23	78	100
.10-.19	8	12	50
Below .10	0	1	12
TOTAL ITEMS FOR 1981-82	130	240	240

The appropriateness of mean point-biserial correlation coefficients must be evaluated in the context of a particular testing program. According to A Reader's Guide to Test Analysis Reports (ETS, 1981), the mean biserial correlation will be higher when the examinee group represents a wide range of ability or knowledge or when the test items are very similar in content. Since the FTCE Reading and Professional Education tests are relatively easy, the scores were not greatly different. Thus, variability was reduced, and the point-biserial correlation coefficients were attenuated.

Contrasting Group Performance

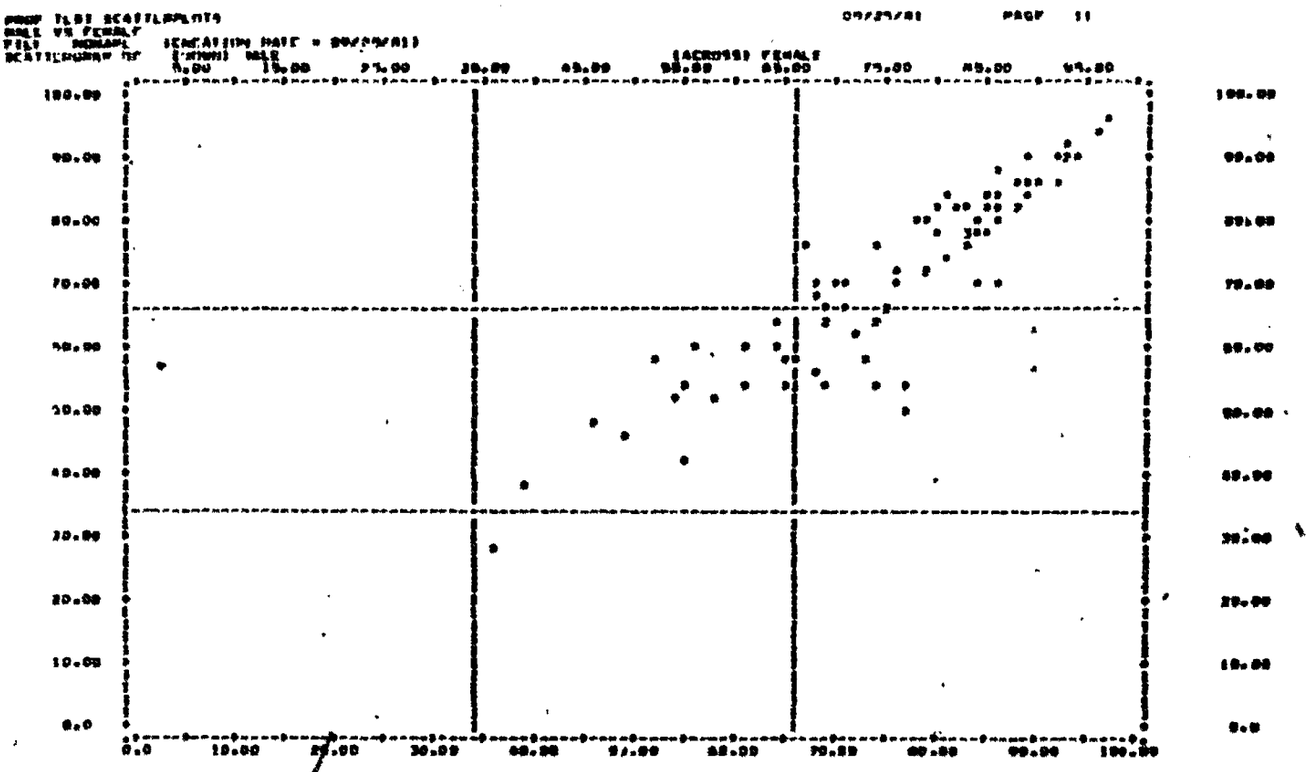
To the extent that scores on a test reflect group membership rather than the knowledge or skill that the test is designed to measure, the test is invalid. Although not all groups necessarily exhibit the same performance level in different areas of achievement, the procedure for analyzing contrasting group performance is to screen for any specific areas or items. Extensive review procedures were used during FTCE development to ensure that the Examination content was an accurate representation of candidate performance in terms of the competencies being evaluated. The procedure included (a) a series of reviews during the item development stage to screen for possibly offensive materials and for items that might invalidate examinee performance and (b) statistical analysis of field groups, ethnic groups, and program groups. These procedures are described in Bulletin IV (pp. 33-38).

After each FTCE administration, test content is examined for contrasting group performance. Score distributions and summary statistics (including mean, median, and standard deviation of the distribution and an index of skewness) are reported for each test. The content review for contrasting group performance includes (a) examination of scatterplots of performance on individual items and overall content by sex and ethnic category (male-female, black-white, white-hispanic, and hispanic-black), (b) analysis of performance by groups based on their test scores, and (c) individual item analysis by sex and ethnic category to screen for items that may discriminate negatively for a specific group.

Scatterplots

Scatter diagrams are graphic representations of the extent to which performance by two separate groups is related. Twelve scatterplots are produced for each FTCE administration, comparing performance by sex and ethnic category for each subtest. Entries that depart from the general pattern indicate that one group is performing differently from another group on specific items. In such cases, entries that depart substantially from the general pattern of other entries are reviewed for content that could account for differences in performance level. Items that are determined to be flawed during this review are revised or deleted from the item pool. An example of a scatterplot is illustrated by Figure 1.

Figure 1. Scatterplot of Percent of Examinees Who Answered Specific Items Correctly.[a]



[a]Performance for males is plotted on the vertical axis while performance for females is plotted on the horizontal axis.

Subtest Performance by Groups

The number and percentage of candidates who pass all subtests and individual subtests are reported by sex and ethnic designation after each FTCE administration. Table 12 displays these data.

TABLE 12

Number and Percent Passing All Subtests and
Each Subtest by Total Candidates and by Sex and Ethnic Designation[a]

First-Time Candidates

	All Candidates			Male			Female			White			Black		
	TOT	N	%	TOT	N	%	TOT	N	%	TOT	N	%	TOT	N	%
ENTIRE TEST	9616	8236	86	2101	1686	81	7515	6550	87	8305	7590	91	809	330	41
MATH	9647	8772	91	2110	1906	90	7537	6866	91	8325	7914	95	815	457	56
READING	9642	9008	93	2109	1900	90	7533	7108	94	8323	8069	97	814	549	67
PROF ED	9641	9287	96	2107	1980	94	7534	7307	97	8320	8231	99	815	618	76
WRITING	9635	9115	95	2107	1905	90	7528	7210	96	8320	8134	98	810	594	73

	Hispanic			American Indian/ Alaskan Native			Asian Pacific			Other		
	TOT	N	%	TOT	N	%	TOT	N	%	TOT	N	%
ENTIRE TEST	348	200	57	16	14	88	24	13	54	114	89	78
MATH	353	271	77	16	15	94	24	19	79	114	96	84
READING	351	255	73	16	16	100	24	21	88	114	98	86
PROF ED	352	299	85	16	15	94	24	21	88	114	103	90
WRITING	351	248	71	16	16	100	24	19	79	114	104	91

[a]Numbers in this table represent data from the three FTCE administrations (1981-1982) presented in this report.

Item Analysis by Sex and Ethnic Category

Separate item analyses -- including item difficulty levels, item discrimination indices, point-biserial correlations, foil analyses (alternative response choices), and KR-20 estimates of reliability -- are reported for each sex and ethnic category. The item analysis process includes the screening of the individual test items that may discriminate negatively for a specific group. When an outlying entry is identified on a scatter diagram, the item content is carefully reviewed to determine the necessity of deleting or revising the item. Foil analyses may also provide useful information with regard to contrasting group performance. Variations in response patterns by groups to different foils (alternative responses) may indicate the need for item revision.

The procedures described in this section -- including scatter diagrams, the analysis of subtest performance by groups, and item analysis by sex and ethnic category -- are used to ensure that scores obtained on the FTCE are accurate representations of the candidates' performance levels in terms of the competencies that are addressed and are not a reflection of membership in a specific sex or ethnic category.

SUMMARY

The Florida Teacher Certification Examination (FTCE) is an examination based upon selected competencies that have been identified by Florida educators as minimal entry-level skills for prospective teachers. In order to develop the Examination, the following tasks had to be accomplished: (a) planning; (b) writing and validation of test items; (c) field-testing the examination items; (d) setting passing scores; and (e) preparing for test assembly, administration, and scoring. The competencies (described in Appendix A) have been adopted by the Board of Education as curricular requirements for teacher education programs in the colleges and universities in Florida.

The FTCE consists of three objective tests (Reading, Mathematics, and Professional Education) and an essay test (Writing) that is scored by trained readers. The general test content is as follows:

Test	Content
Writing	One of two general topics
Reading	General education passages derived from textbooks, journals, state publications
Mathematics	Basic mathematics: simple computation, and "real world" problems
Professional Education	General education including personal, social, academic development, administrative skills, exceptional student education

Developmental items are included in the Examination along with regular test items. These developmental items are not counted in computing an individual's score.

The psychometric characteristics of validity, reliability, item discrimination, and contrasting group performance of the FTCE are described in this report. The validity of the examination has been well established as a result of (1) the extensive involvement of education professionals in the identification and explication of the necessary competencies and their associated subskills, (2) the precise item specifications which guided the item writers, and (3) reviews of the items and the competencies/skills that they were designed to measure. The reliability data indicate that the test items are consistent measures of the three subject areas and that the examinees' scores are consistent with their classification as masters or nonmasters of the minimal performance standards. The reliability data for the Writing test demonstrates that the scoring by the writing teams meets acceptable standards of consistency. Item analyses for the FTCE examine the power of the items to differentiate

between ability groups and evaluate response patterns to individual items. The indices that are used to monitor the differences between ability groups are item percent correct, item discrimination index, and point-biserial correlation coefficients. Additional item analysis procedures -- including scatter diagrams, the analysis of subtest performance by groups, and item analysis by sex and ethnic category -- are used. These procedures ensure that scores obtained on the FTCE are accurate representations of the candidates' performance levels in terms of the competencies that are addressed and are not a reflection of membership in a specific sex or ethnic category.

The FTCE is administered three times a year in selected locations throughout the state. Data from this report indicate that the percentage of candidates who passed the entire FTCE for the October 1981, February 1982, and July 1982 administrations were 84 percent, 84 percent and 85 percent, respectively. Examinees who do not pass all of the tests at one administration may retake the tests not passed at later scheduled testing dates.

APPENDIX A

Essential Competencies and Subskills

Essential Competencies and Subskills

1. Demonstrate the ability to orally communicate information on a given topic in a coherent and logical manner.
   a. Utilizes principles of simplicity and clarity in organization of oral presentation.
   b. Uses standard English in oral communication.
   c. Uses vocabulary suitable to the topic and audience.
   d. Speaks with a volume and pace that promote comprehension.
   e. Provides verbal and nonverbal cues to the organizational structure of the oral message.
   f. Provides relevant examples that illustrate oral content.

2. Demonstrate the ability to write in a logical easily understood style with appropriate grammar and sentence structure.
   a. Differentiates between formal and informal written English and demonstrates ability to use both forms.
   b. Uses language at the level appropriate to the topic and reader.
   c. Comprehends and applies basic mechanics of writing: spelling, capitalization, and punctuation.
   d. Comprehends and applies appropriate sentence structure.
   e. Comprehends and applies basic techniques for the organization of written material.
   f. Comprehends and applies standard English usage in written communication.

3. Demonstrate the ability to comprehend and interpret a message after listening.
   a. Accurately follows multi-step oral directions.
   b. Listens effectively for the organization, main idea, subordinate ideas and details of a message.
   c. Listens effectively in order to identify relevant and irrelevant information and propaganda techniques.
   d. Listens effectively in order to draw inferences.
   e. Summarizes the message after listening.
   f. Comprehends both standard and nonstandard English language used by students.

4. Demonstrate the ability to read, comprehend, and interpret, orally and in writing, professional material.
   a. Identifies and evaluates relevant professional material.
   b. Understands basic statistical terminology (such as mean, median, mode).
   c. Demonstrates literal reading skills (such as recognizing main idea, details, sequencing, comparison and contrast).
   d. Demonstrates interpretive reading skills (such as predicting outcome, drawing conclusions, making generalizations).
   e. Demonstrates critical reading skills (such as recognition of relevant and irrelevant information, propaganda techniques, and fallacies in reasoning).
   f. Produces a logical summary interpretation of the results of research in professional material.

5. Demonstrate the ability to comprehend and work with fundamental mathematical concepts.
   a. Adds, subtracts, multiplies and divides whole numbers, decimals and fractions.
   b. Demonstrates the meaning and use of fractions and percents.
   c. Represents and interprets data using charts, tables, graphs and maps.
   d. Solves measurement problems involving length, area, volume, capacity, weight, time and temperature, using U.S. customary and metric units.
   e. Applies mathematical skills to solve real world problems.
   f. Identifies geometric forms and relationships.

6/19. The ability to comprehend patterns of physical, social and academic development in students, including exceptional students in the regular classroom, and to counsel the same students concerning their needs in these areas.
   a. Demonstrates knowledge of basic principles of human growth and development, and awareness of individual differences in students.
   b. Obtains knowledge of students through classroom tests, teacher observations, and student records to contribute to understanding of student needs.
   c. Explains test data and other classroom evaluations to students in relation to their educational needs.
   d. Applies motivational techniques to encourage students to be achievement-oriented and goal-directed.
   e. Assists the student in relating achievements and interests to aptitude and ability.
   f. Demonstrates knowledge of alternative school and community resources for students who have special needs.
   g. Assists students in developing individual learning activities.
   h. Develops student awareness of career opportunities using school and community resources.

7. Diagnose the entry level knowledge and/or skills of students for a given set of instructional objectives using diagnostic tests, teacher observation and student records.
   a. Selects the specified knowledge or skill to be diagnosed, and determines the most appropriate method for conducting the diagnosis.
   b. Selects or constructs a test to diagnose student learning needs.
   c. Uses classroom observation techniques to diagnose student learning needs.
   d. Uses information from student records to diagnose student learning needs.
   e. Interprets results obtained from diagnostic tests, teacher observation and information from student records.

8. Identify long-range goals for a given subject area.
   a. Identifies state and district long-range goals.
   b. Formulates subject area goals consistent with state and district goals and student needs.

9. Construct and sequence related short-range objectives for a given subject area.
   a. Identifies knowledge, skills and attitudes to be attained for a subject area.
   b. Constructs or adapts short-range objectives for identified knowledge, skills and attitudes.
   c. Sequences short-range objectives consistent with commonly accepted principles of learning.

10. Select, adapt and/or develop instructional materials for a given set of instructional objectives and student learning needs.
   a. Determines desirable characteristics of materials based on objectives and student learning needs.
   b. Locates and evaluates available instructional material.
   c. Selects materials to assist students in mastering an objective.
   d. Demonstrates techniques for modifying materials to assist students in mastering an objective.

e. Determines materials to be developed based on existing resources and student needs.
f. Identifies and selects resources needed for materials development.
g. Designs and constructs materials based on instructional objectives, student needs and available resources.
h. After use, evaluates the effectiveness of instructional materials in accomplishing objectives, and revises accordingly.

11. Select/develop and sequence related learning activities appropriate for a given set of instructional objectives and student learning needs.
a. Comprehends basic principles of human growth and development.
b. Identifies conditions that affect learning.
c. Identifies alternative activities to achieve an objective.
d. Selects an appropriate learning activity to achieve an objective.
e. Combines appropriate learning activities into an instructional sequence.

12. Establish rapport with students in the classroom by using verbal and/or visual motivational devices.
a. Secures the attention of students through appropriate techniques.
b. Relates instructional objectives and activities to interests, capabilities and experiences of students.
c. Informs students about objectives, subsequent learning tasks and performance expectations.
d. Explains choices and limitations of possible learning activities.
e. Alters instructional strategies during learning activities based on student responses and other factors.
f. Relates students and teacher's experiences, thoughts and feelings to learning activities.
g. Uses reinforcement techniques to assist in student motivation.
h. Uses media to secure interest and maintain attention.
i. Uses student products and talent to secure interest and maintain attention.

13. Present directions for carrying out an instructional activity.
a. Selects appropriate means for presenting directions.
b. Secures attention of students for the purpose of giving directions.
c. Informs students of objectives, assessments and performance standards.
d. Informs students of the sequence and nature of learning activities to achieve the objectives.
e. Identifies materials for a learning activity and explains their use.
f. Determines if students understand directions.
g. Clarifies directions by responding to student questions.

14. Construct or assemble a classroom test to measure student performance according to criteria based on objectives.
a. Identifies uses of basic types of classroom tests and assessment techniques.
b. Identifies appropriate uses of norm-referenced and criterion-referenced testing.
c. Given an objective, specifies knowledge and skills to be assessed.
d. Selects appropriate assessment techniques to evaluate mastery of an objective.
e. Determines limitations, constraints and requirements for administering tests.
f. Constructs and identifies test items and tasks that evaluate mastery of an objective.
g. Identifies criteria for standards of performance.
h. Assembles test components including test items, directions and scoring keys.
i. Evaluates and/or revises tests on the basis of validity, reliability and student responses.

15. Establish a set of classroom routines and procedures for utilization and care of materials.
a. Involves students in developing classroom routines and procedures for utilization and care of materials.
b. Determines the type and amount of materials necessary to complete classroom assignments.
c. Organizes an effective system for placement and distribution of materials in the classroom.
d. Organizes and arranges a center that will serve as a focus of interest for student learning (such as a bulletin board, display table, or exhibit).
e. Identifies physical elements and arrangements in the classroom that directly affect learning.
f. Involves students in developing routines and procedures for physical movement in the classroom.
g. Arranges classroom furniture and equipment to accommodate selected teaching strategies.
h. Identifies approved procedures for movement of students in emergencies that can be anticipated.

16. Formulate a standard for student behavior in the classroom.
a. Identifies approved safety procedures and incorporates them into a standard for student behavior in the classroom.
b. Identifies and incorporates socially accepted norms (such as mutual respect, consideration of others, courtesy) into a standard for student behavior in the classroom.
c. Identifies characteristics of the student population (such as age and maturity) that need to be considered in formulating a standard for student behavior in the classroom.
d. Establishes a realistic standard of behavior that has potential for consistent application.
e. Identifies and incorporates state and local policies into a standard for student behavior in the classroom.

17. Identify causes of classroom misbehavior and employ a technique(s) for correcting it.
a. Identifies factors of the physical environment that affect student behavior.
b. Identifies social and emotional characteristics of the teacher that affect student behavior.
c. Identifies physical, social and emotional characteristics of the student that affect student behavior.
d. Identifies out-of-school factors that affect student behavior.
e. Identifies aspects of instructional procedures and techniques which affect student behavior.
f. Demonstrates effective techniques and strategies for managing student behavior.
g. Uses selected verbal and nonverbal techniques for enforcing and modifying student behavior.
h. Identifies and uses school and community resources for assistance in modifying student behavior.
i. Obtains and utilizes parental assistance for modifying student behavior.

18. Identify and/or develop a system for keeping records of class and individual student progress.
a. Constructs a system for recording individual student knowledge and skills progress in a subject area.
b. Identifies methods for reporting individual student progress in knowledge and skills in a subject area.
c. Identifies methods for recording class progress in knowledge and skills in a subject area.
d. Identifies methods for reporting class

progress in knowledge and skills in a subject area.

e. Demonstrates knowledge of the laws and policies governing the content and use of student records.

19. (See Section #6)

20. Identify and/or demonstrate behaviors which reflect a feeling for the dignity and worth of other people including those from other ethnic, cultural, linguistic and economic groups.
    a. Creates a learning environment in which students express themselves openly and honestly.
    b. Assists students in understanding that individual differences enable each person to make unique contributions to the group effort.
    c. Demonstrates awareness of cultural differences in dress, beliefs and practices.
    d. Establishes an environment for positive communication and interaction between students from different socio-cultural backgrounds.

21. Demonstrate instructional and social skills which assist students in developing a positive self-concept.
    a. Exhibits behavior in the classroom that is empathetic, positive and reinforcing.
    b. Assists students in initiating self-directed learning.
    c. Assists students in understanding their needs, motives, experiences and individual value and dignity.
    d. Selects and uses curriculum materials in accordance with the abilities and mastery levels of individual students.

22. Demonstrate instructional and social skills which assist students in interacting constructively with their peers.
    a. Establishes an environment that permits students to cooperate and share ideas and materials.
    b. Assists students in applying constructive criticism in response to each other's work.
    c. Establishes a learning environment designed to assist students in exhibiting positive interpersonal traits (such as mutual respect and cooperation).
    d. Uses techniques that assist students in examining their values, attitudes and beliefs.

23. Demonstrate teaching skills which assist students in developing their own values, attitudes and beliefs.
    a. Assists students in understanding the need to explore alternative solutions to problems.
    b. Establishes teaching strategies that allow students to make choices based on clearly defined consequences.

24. Beginning July 1, 1982, the ability to recognize and be aware of the instructional needs of exceptional students.
    a. Identifies the characteristics of exceptional students that have implications for modifying the learning environment.
    b. Demonstrates awareness and appropriate use of educational programs, support services, personnel and other resources available to meet the needs of exceptional students.
    c. Demonstrates the ability to identify and appropriately refer students who may be in need of exceptional student education.
    d. Demonstrates awareness of the roles of the parent, teacher, and other professional personnel as members of the educational team responsible for planning, implementing, and evaluating the exceptional student's program.
    e. Demonstrates the ability to recognize and/or use alternative instructional strategies to implement that portion of the exceptional student's program for which the teacher has responsibility.
    f. Identifies and/or selects effective techniques and strategies for facilitating integration and social acceptance of exceptional students.

APPENDIX B

Mathematical Illustrations of Formulas

The following formulas were used in the calculation of statistics for the FTCE:

(a) Point-Biserial Correlation

$$r_{pb} = \frac{m_s - m_u}{\sigma} \left[\sqrt{pq} \right]$$

where r_{pb} = point biserial correlation coefficient

m_s = mean total score of examinees answering item right

m_u = mean total score of examinees answering item wrong

σ = standard deviation of total score for entire group

p = proportion of examinees getting item right

q = $1 - p$

(b) Kuder-Richardson Formula 20 Reliability Coefficient

$$KR_{20} = \frac{I}{I - 1} \left[1 - \frac{\Sigma pq}{\sigma_y^2} \right]$$

where I = number of items/questions (any omitted questions not included)

p = proportion of examinees getting item right

q = $1 - p$

σ_y^2 = variance of the total score

(c) Standard Error of Measurement

$$\sigma_s = \sigma_x \sqrt{1 - r_{xx}}$$

where σ_s = standard error of measurement

σ_x = the standard deviation of total scores

r_{xx} = the reliability coefficient

(d) Item Discrimination Index

$$\frac{R_u - R_\ell}{\frac{1}{2}(T)}$$

where R_u = number of students in high score range (i.e., the upper 27%) who answered the item correctly

R_ℓ = number of students in low score range (i.e., the lower 27%) who answered the item correctly

T = total number in the upper and lower groups

(e) Coefficient Alpha

Coefficient alpha is used as an estimate of the inter-rater reliability of Writing test scores. This coefficient indicates the expected correlation between the ratings of the team on this task and those of a hypothetical team of similarly comprised trained raters doing the same task.

$$r_{kk} = \frac{k}{k-1} - \left[1 - \frac{\Sigma\sigma_i^2}{\sigma_y^2} \right]$$

where r_{kk} = coefficient of reliability (alpha)

k = number of test items

$\Sigma\sigma_i^2$ = sum of the variances of each item

σ_y^2 = variance of the examinees' total test score

(f) Brennan-Kane Reliability

$$\hat{m}(c) = 1 - \frac{1}{n_i - 1} \left[\frac{X_{PI}(1 - X_{PI}) - S^2(X_{PI})}{(X_{PI} - C)^2 + S^2(X_{PI})} \right]$$

where n_i = number of items

X_{PI} = grand mean over n_p persons and n_i items

$S^2(X_{PI})$ = sample variance of persons' mean scores over items

that is, $\frac{\text{SS persons}}{n_p} \div n_i$

APPENDIX C

Security and Quality Control Procedures

SECURITY

A security and quality control plan has been developed and implemented for the program. Components of the security plan include:

1. Controlled, limited access to all examination materials;
2. Shredding of developmental materials and used booklets;
3. Strict accounting of all materials and of persons working with test items at the testing agency and test centers.

A signed security document is obtained from every individual who has access to the examination materials. The security document contains an agreement that the individual will not reveal -- in any manner to any other individuals -- the examination items, paraphrases of the examination items, or close approximations to the examination items. Only persons who have a "need to see" the items because of their work on the project are allowed to view any parts of the examination.

Test Security During the Administration

During the production phases of this project all typing and reproduction are done by persons who have security clearances. All materials are signed out when they are removed from locked storage and checked in when they are returned. One person is assigned responsibility for the secure files while all work is in process; this person is able to account for all materials at all times. Material that needs to be revised and unusable materials are not placed in wastebaskets but are kept in a locked file for special destruction.

The following plan has been implemented to ensure rigorous security of all materials during actual examination administration. Materials remain in secure storage at the test centers until the morning of the test date. If multiple rooms are used at a center, each room is assigned blocks of materials that must be signed for by a room supervisor, the only person who has access to the room supply. Test books and materials are never left unguarded. Candidates are assigned seats by center personnel. The seating arrangements minimize the possibility of a candidate seeing the papers of other candidates. Books are distributed by the room supervisor and proctors. Each booklet is handed to the examinees individually and the examinees sign a receipt for the booklets by serial number. Immediately after distribution, an inventory is taken to ensure that the sum of the distributed and unused books equals the number of books assigned to the testing room. Any discrepancy is reported to the center supervisor and immediate steps are taken to reconcile the discrepancy and locate the missing material. Every such incident is reported to the Project Manager, and appropriate action is instituted to prevent further occurrences and to recover any missing materials.

Candidates cannot leave the room during a test session except for an emergency. If a candidate must leave the room, materials are delivered to the room supervisor or proctor and held until the candidate's return. No materials

may be removed from the test room at any time. Only one candidate may leave the room at a time. Provision of a break between subtests reduces the need for candidates to leave during a test session. At the end of a session all test books are collected and accounted for before collection of the answer documents. After the answer documents are accounted for, all candidates in the room are dismissed. Upon return to their original seats, candidates are reidentified by test administration personnel before the distribution of materials for the next subtest. During breaks and the lunch period, all materials are either locked in secure storage or are placed under direct supervision of test administration personnel. All used and unused materials are returned to locked storage immediately after test administration.

Quality Control

To ensure quality control during the scoring and reporting process, the following procedures are used:

1. Each answer document is checked for proper coding and marking in response areas;

2. Computer edit programs are used to check for valid program codes on the registration forms and for matching names and social security numbers on the registration and scoring files;

3. Test data are used to verify the accuracy of all scoring and reporting programs;

4. Sample data are drawn prior to scoring from each administration to screen for key, printing, or procedural errors;

5. Random answer documents are hand-scored during the scanning process to verify proper operation of the scanner;

6. A complete review of all procedures -- which includes hand-checking a sample of test data -- is completed by members of the University of Florida, Office of Instructional Resources and the Department of Education before printing the candidate score reports;

7. Analyses of the holistic scoring process are conducted. This review addresses the overall reliability of the ratings, the distribution of scores, and number of refereed scores for each reader. Specific procedures for quality control during the holistic scoring process are documented in the Procedural Manual for Holistic Scoring;

8. The accuracy of the calculations for the institutional and state reports are hand-verified.

APPENDIX D

Scoring the Writing Examination

HOLISTIC SCORING OF THE WRITING SUBTEST OF THE FLORIDA TEACHER CERTIFICATION EXAMINATION

The Writing Subtest

The Writing Subtest was designed to assess a candidate's ability to write in a logical, easily understood style with appropriate grammar and sentence structure. The subskills to be measured are:

a. Uses language at the level appropriate to the topic and reader;
b. Comprehends and applies basic mechanics of writing: spelling, capitalization, and punctuation;
c. Comprehends and applies appropriate sentence structure;
d. Comprehends and applies basic techniques for the organization of written material;
e. Comprehends and applies standard English usage in written communication.

The candidate is given a choice between two topics on which to write an essay during the 45-minute examination period. This essay should demonstrate the competency and subskills specified above. The essay or writing sample is scored holistically by at least three trained and experienced judges.

The Process of Holistic Scoring

Holistic Scoring Defined

Holistic scoring or evaluation is a process for judging the quality of writing samples. It has been used for many years by professional testing agencies for credit-by-examination, state assessment and teacher certification programs.

Essays are scored holistically, that is for the total, overall impression they make on the reader, rather than for an analysis of specific features of a piece of writing. Holistic scoring assumes the skills which make up the ability to write are closely interrelated and that one skill cannot be separated from the others. Thus, the writing is viewed as a total work in which the whole is something more than the sum of the parts. A reader reads a writing sample quickly, once. He or she obtains an impression of its overall quality and then assigns a numerical rating to the paper based on judgments of how well it meets a particular set of established standards.

The Reader

The key to effectiveness of the holistic scoring process is the readers who must make valid and reliable judgments. Readers must bring to the process experience in teaching and grading English compositions. In addition, they must

be willing to undergo training in holistic scoring which demands they set aside personal standards for judging the quality of a writing sample and adhere to standards which have been set for the examination. The goal for the reading of the Writing Subtest of the Florida Teacher Certification Examination is to rate a large number of essays according to their overall competence in a consistent or reliable manner according to previously established standards based on a set of defined criteria. By undergoing a set of training procedures a group of experienced teachers of composition can develop a high level of consistency in making judgments about the quality of a group of essays.

The Criteria

The criteria established to score the essays for the Florida Teacher Certification Examination are listed below. They were developed to accommodate specific conditions imposed by the Writing Subtest:

(1) They reflect those characteristics widely accepted as indicative of good writing;
(2) They can be translated into operational descriptions of levels of competence;
(3) They reflect the general competency statement and subskills identified by the Council on Teacher Education.

Specific Criteria for Evaluation of Essays

1. Rhetorical Quality

1.1 Unity: An ordering and interdependence of parts producing a single effect: completeness.

1.2 Focus: Concentration of a topic; the presence of a "center of gravity."

1.3 Clarity: Lucidity of expression; lack of ambiguity and distortion.

1.4 Sufficiency: Appropriate depth and breadth of expression to meet the writer's purposes and the demands of the particular topic.

2. Structural and Mechanical Quality

2.1 Organization: Consistent and coherent integration and connection of parts.

2.2 Development: Appropriate and sufficient exposition of ideas; use of detail, examples, illustration, comparisons, etc.

2.3 Paragraph and Sentence Structure: Appropriate form, variety, logic, relatedness of and among structural units.

2.4 Syntax: Appropriate ordering of words to convey intended meaning.

3. Observance of Conventions in Writing

3.1 Usage: Appropriate use of language features: inflections, tense, agreement, pronouns, modifiers, vocabulary, level of discourse, etc.

3.2 Spelling, Capitalization, Punctuation: Consistent practice of accepted forms.

The relationship between the subskills and the scoring criteria is illustrated in the figure below.

ESSENTIAL COMPETENCIES: Demonstrate the ability to write in a logical, easily understood style with appropriate grammar and sentence structure.	RHETORICAL				STRUCTURAL				CONVENTIONAL	
	1.1 Unity	1.2 Focus	1.3 Clarity	1.4 Sufficiency	2.1 Organization	2.2 Development	2.3 Structure	2.4 Syntax	3.1 Usage	3.2 Mechanics
a. Use language appropriate to the topic and reader.			X	X						
b. Apply basic mechanics of writing.									X	X
c. Apply appropriate sentence structure.								X	X	X
d. Apply basic techniques for organization.	X	X			X	X	X			
e. Apply standard English usage.								X	X	X

Operational Descriptions

The operational descriptions based on the scoring criteria reflect the four levels of competency which the readers are to assign each of the essays they read. Each reader will independently score or rate a paper on a scale of 1 to 4, with 4 being the highest rating. The descriptions which follow are an attempt to express clearly and precisely the general, overall impressions a reader has in terms of the criteria when he or she reads essays of varying

quality. The four levels or quality of competence could be expanded or decreased. However, for the task of scoring the Writing Subtest, it provides enough degrees of distinction to be meaningful yet manageable for large scale testing.

4. The essay is unified, sharply focussed, and distinctively effective. It treats the topic clearly, completely, and in suitable depth and breadth. It is clearly and fully organized, and it develops ideas with consistent appropriateness and thoroughness. The essay reveals an unquestionably firm command of paragraph and sentence structure. Syntactically, it is smooth and often elegant. Usage is uniformly sensible, accurate, and sure. There are very few, if any, errors in spelling, capitalization, and punctuation.

3. The essay is focussed and unified, and it is clearly if not distinctively written. It gives the topic an adequate though not always thorough treatment. The essay is well organized, and much of the time it develops ideas appropriately and sufficiently. It shows a good grasp of paragraph and sentence structure, and its usage is generally accurate and sensible. Syntactically, it is clear and reliable. There may be a few errors in spelling, capitalization, and punctuation, but they are not serious.

2. The essay has some degree of unity and focus, but each could be improved. It is reasonably clear, though not invariably so, and it treats the topic with a marginal degree of sufficiency. The essay reflects some concern for organization and for some development of ideas, but neither is necessarily consistent nor fully realized. The essay reveals some sense, if not full command, of paragraph and sentence structure. It is syntactically bland and, at times, awkward. Usage is generally accurate, if not consistently so. There are some errors in spelling, capitalization, and punctuation that detract from the essay's effect if not from its sense.

1. The essay lacks unity and focus. It is distorted and/or ambiguous, and it fails to treat the topic in sufficient depth and breadth. There is little or no discernible organization and only scant development of ideas, if any at all. The essay betrays only sporadically a sense of paragraph and sentence structure, and it is syntactically slipshod. Usage is irregular and often questionable or wrong. There are serious errors in spelling, capitalization, and punctuation.

Training of Readers

The training of readers for the Writing Subtest of the Florida Teachers Certification Examination consists of three steps:

Step 1: Acquiring information about the examination and holistic scoring process.

Step 2: Reading and scoring essays which have been selected as good examples of the various levels of competence in writing. The practice essays have been scored by experienced readers and annotated in accordance with the operational descriptions. By reading, scoring and discussing the essays, the readers practice until they consistently give the same ratings to essays as the experienced readers.

Step 3: Reading and scoring a sample of the actual Writing Subtests which have been selected and scored prior to the training session. These samples will serve as the standards for the scoring of the examination and will include essays which represent each of the competency levels. As in Step 2, the emphasis will be on each reader to assign scores which agree with those established earlier by the experienced readers. This step occurs immediately before the actual scoring session and often is repeated during the session to ensure continued consistency or reliability of assigned scores or ratings.

Setting the Standards

Prior to Step 3 in the training, standards for the Writing Subtest are established. The Chief Reader, who is responsible for conducting the holistic scoring, and his assistants, the Assistant Chief Reader and the Table Leaders, select, at random, a sample of papers from the total group of essays written on a particular topic. These papers are read and scored independently by each person. Results are compared and consensus is reached for the identification of four papers. Each becomes a standard for one of the four competency levels. Additional papers are chosen to be used in Step 3 of the training procedures. This process is repeated for the second topic of the Writing Subtest.

The Scoring Session

The scoring session begins immediately after Step 3. Readers are assigned to tables in groups of four or five. The number of readers and the number of tables are determined by the number of essays to be scored. Each table of readers is also assigned a Table Leader. The Table Leader's primary task is to continually monitor the scoring process and consult with readers as questions or "problem" papers arise. The Table Leader is an experienced reader who has helped set the standards.

Each reader is given a set of papers to read, rate and mark the score. The identity of the writer is not known to the reader. The papers range, on the average, from 200 to 400 words in length, and each can be read and scored holistically in approximately two minutes. As the scoring of a set of papers is completed by a reader, a clerk collects and returns the paper to an operation table. The scores given by the reader are covered, and the papers are

redistributed to another set of folders and delivered by the clerk to a second table of readers. This procedure continues until each paper has been read by three different readers. Each reader reads, judges and scores at his own pace. Scoring sessions are approximately three hours long, with ten minute breaks each hour. Usually there are two scoring sessions for each day of holistic reading.

After a paper has been scored by three different readers, the scores are examined at the operations table. If one of the scores varies from any other by two levels or more (ex. 3-3-1), the paper is sent to the Chief Reader or Assistant Chief Reader who serves as referee. This person assigns a rating which replaces the discrepant score. Papers whose original ratings are 1-2-3 or 2-3-4 are refereed and scored as follows:

Rating of 1-2-3

(a) A referee rating of 1 will replace the 3, resulting in a score of 4
(b) A referee rating of 2 will replace the 1, resulting in a score of 7
(c) A referee rating of 3 will replace the 1, resulting in a score of 8

Rating of 2-3-4

(a) A referee rating of 2 will replace the 4, resulting in a score of 7
(b) A referee rating of 4 will replace the 2, resulting in a score of 11
(c) A referee rating of 3 will replace the 2, resulting in a score of 10

All initial scores of 5 will be refereed. If any paper is refereed and a discrepancy still occurs, the essay is submitted to a new team of readers until consistency is obtained.

The three scores are then added together for a total score. Thus the lowest score possible is a 3, the highest, 12.

Final Steps

After the reading sessions are completed, Table Leaders evaluate the performance of Readers. The Chief Reader evaluates the Table Leaders. Readers are asked for comments and suggestions for improving training and scoring procedures.

Two approaches for reliability estimation are the percentage of rater agreement and the calculation of coefficient alpha for the raters and the rating team, which indicates the expected correlation between the ratings of the team and those of a hypothetical team of similarly comprised and similarly trained raters doing the same task. The four indices that represent rater agreement are: (a) percent complete agreement; (b) average percent of two of the three raters agreeing; (c) average percent agreement by pairs as to pass/fail; and (d) percent complete agreement about pass/fail. These data are reported in Table 8 of this report.