THE EFFECTS OF TEACHER TRAINING AND EXPERIENCE IN A CLASSROOM DECISION-MAKING SIMULATION

David Michael Shannon
Reading, PA

B.S. Ed., Kutztown University, 1986

A Dissertation Presented to the
Graduate Faculty of the University of Virginia
in Candidacy for the Degree of
Doctor of Philosophy

Department of Education

University of Virginia
August, 1990

Donald W. Ball, Ph.D.
Committee Co-Chair

Donald M. Medley, Ph.D.
Committee Co-Chair

Michael S. Caldwell, Ph.D.

Carolyn M. Callahan, Ph.D.

David L. Clark, Ph.D.

ABSTRACT

The Effects of Teacher Training and Experience in a Classroom Decision-making Simulation

The purpose of this study was to verify and extend the findings obtained during the field test of the Simulation Test of Interactive Teaching Competencies-Hays (STITC-H) test. This test simulated the limited time allowed for making decisions in the classroom by projecting written descriptions of typical classroom situations on a videotape. Alternative teacher responses were presented, one at a time, and subjects were given five seconds to indicate whether each course of action was appropriate or inappropriate.

The specifications for the STITC-H were developed using a group of indicators of successful teaching specified by the Virginia Beginning Teacher Assistance Program (BTAP) and each situation was approved by a panel of experts which included teachers, administrators and college faculty.

This study was designed to answer three major research questions which attempted to verify whether teachers learn the research findings on effective teaching from their professional training and from their classroom teaching experiences and whether this knowledge is reflected in their performance on the Simulation Exercise in Classroom Decision-Making. It was hypothesized that: 1) university students with training in education would score higher than students without training in education, 2) those subjects

with classroom teaching experience (student teachers and teachers) would score higher than those without classroom teaching experience (non-education and education students) and 3) experienced teachers would score higher than student teachers.

Regarding the overall test performance of subjects, two of the three hypotheses were rejected at the .05 level. A significant difference between the overall performance of experienced teachers and student teachers was not found. All three hypotheses were rejected upon the examination of specific group profiles across the eleven competencies. Evidence in support of construct validity was establihed as this test discriminated among subjects with varying degrees of training and experience in education.

It was recommended that this test be refined, administered to other groups, correlated with other measures such as the NTE and GRE as well as future classroom observations. If this simulation exercise can be used to predict teacher performance, it could then be used a both a screening device within teacher education programs and local school divisions and to identify strengths and weaknesses of practicing teachers so that the appropriate inservice training could be scheduled and other remedial services provided.

ACKNOWLEDGEMENTS

The preparation and completion of this final document could not have been accomplished without the help, encouragement and support of many people.

Among these people I'd like to thank the members of my dissertation committee - Donald Medley, Don Ball, Carolyn Callahan, Michael Caldwell, and David Clark - for their continued feedback and constructive ideas throughout this study. My special thanks to Don Medley, who worked so patiently and diligently with me. I'd also like to extend my appreciation to the Commonwealth Center for the Advancement of Teaching for its' efforts in recruiting pre-service teachers to participate in this study.

The support staff from the Department of Educational Studies, the Bureau of Educational Research, the Department of Curriculum and Instruction, and the Office of Teacher Education have certainly made the entire process a much easier and pleasant one. Special thanks to Dave Drucker and Cynthia Nicholson and for their assistance in the production of the simulation exercise used in this study.

Lastly, I would like to acknowledge all my colleagues and friends in the Bureau of Educational Research, Department of Educational Studies, and other program areas who have been a critical factor not only in the dissertation process, but have been constant throughout the past four years at the University.

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

Recently, the competence of teachers, especially beginning teachers, has been sharply questioned. The decline of students' test scores below the public's expectations have led the public to believe that part of this problem can be linked to teacher preparation. Therefore, in response to public opinion, state legislatures and state departments of education have mandated that teachers be tested to assure their competence in the classroom. Such teacher testing is expected to result in more effective teaching practices and, in turn, return the level of student achievement to an acceptable level in the eye of the public.

State legislatures began to mandate the testing of teachers in 1964. Since this time, the movement to test teachers before admission to teacher education programs or state certification has practically encompassed the United States. State mandates to test teachers were established in three states by 1977, increased to 12 in 1980; 22 in 1981; 28 in 1982; 30 in 1983; and 38 in 1987, with 8 of the remaining 12 states drafting proposals. (Association of Teacher Educators, 1988; Eisenberg & Rudner, 1988; Sandefur, 1985).

Teacher assessment has been proposed because of its

potential benefits to many people. Teacher educators would be able to screen applicants prior to admission and by assessing the same students at the completion of the program estimate the extent to which their curriculum prepared teachers. For teachers, assessment might increase the status of the profession resulting in more respect and admiration by the public. Local education agencies would be able to screen teacher candidates as well as identify areas for the professional development of practicing teachers. Finally, the elevation of the teaching profession would result in more favorable reactions by the general public because of the increase in student achievement associated with more successful teachers.

Competency testing of teachers has received support from both the general public and teachers. In the 1986 Gallup Poll, 89% of the public favored such tests before hiring teachers and 85% favored tests for experienced teachers (Gallup Poll: Public Opinion, 1986; Gallup, 1986). In 1984, 63% of the classroom teachers polled believed that prospective teachers should be required to pass a state competency test before becoming certified (Gallup, 1984). Further support has been expressed by the National Education Association (NEA) (Futrell & Robinson, 1986), American Federation of Teachers (AFT) (Shanker, 1985), National Commission on Excellence in Education (1988), the Holmes Group (Holmes Group, 1986), Council of Chief State School

Officers (CCSSO), and the Carnegie Forum on Education and the Economy (cited in Rudner, 1987).

Despite the general support for the testing of teachers, two major questions must be answered before such an assessment program can be established. First, it must be determined what is to be assessed. What does a teacher need to know in order to be judged as effective or competent? Second, specifications must be established as to how to collect the information and determine the passing level. The determination of what comprises competence is necessary before any assessment can take place. It is this definition that serves as the foundation on which a valid test must be built. This definition can be as general as the basic skills of reading, writing and arithmetic. However, the competencies specifically related to teaching must include the knowledge, skills, and professional values which are most applicable to successful teaching (Medley, 1982). These competencies are much more difficult to identify and to measure.

Traditionally, the assessment of teachers has been conducted with standardized paper and pencil tests or supervisory ratings of classroom performance. State legislatures and boards of education have generally relied on paper and pencil tests. For certification purposes, most states and teacher preparation institutions use the National Teacher Examination (NTE) or some state standardized exam.

Educational Testing Service (ETS) reported in November of 1987 that thirty-two states used a nationally standardized test for certification purposes or admission to teacher education programs and one state used a state developed test while six states used a combination of state and national tests (Association of Teacher Educators, 1988).

Despite certain advantages, existing methods of teacher assessment have not been very successful. Pencil and paper tests are relatively cheap, quick, easy to administer and readily available. They also appear objective and fair because each teacher is faced with the same task and the personality of the teacher is not a factor in determining the outcome. These tests have been criticized primarily because they lack face and predictive validity. The task of choosing the one "best" answer on a multiple-choice test is much different from reacting to situations in the classroom for which there may not be a "best" alternative (Medley, 1978; Pottinger, 1979).

Quirk, Witten, and Weinberg (1973) have compiled a review of the predictive validity studies conducted during the first 30 years of the NTE's existence. Quirk, et al. (1973) reported a median correlation of .05 between NTE scores and ratings by college supervisors and principals during student teaching. Correlations between NTE scores and principal ratings during the first year of teaching were not much better having a median value of .11. Two

primary reasons for this poor prediction were offered by Quirk, et al. First, the NTE only measures a sample of the important qualities related to successful teaching, most of which have more to do with planning strategies and class management rather than subject matter knowledge. The second reason noted concerns the unreliability of the rating scales used for on-the-job evaluations.

Given these conditions, there appears to be a great need for a valid measure of teacher competence which would be relatively inexpensive, quick, and easy to administer and score. Hays (1988) developed such a test of teacher competencies to meet this need. The specifications for the Simulation Test of Interactive Teaching Competencies-Hays (STITC-H) were developed using a group of indicators of teacher competence specified by the Virginia Beginning Teacher Assistance Program (BTAP) (Bureau of Educational Research, 1984; Hays, 1988). Thirteen of the fourteen competencies identified by BTAP were related to interactive teaching and were used in the construction of the STITC-H. The fourteenth competency, Planning, was identified as a preactive function of the teacher and thus was not included. A list of the definitions and indicators for each of these thirteen competencies is included in Appendix A.

The competencies included in the BTAP program were derived from a review of the literature on teacher effectiveness. This review consisted of a meta-analysis of

five reviews of the research literature on effective teaching (Brophy & Good, 1986; Crawford & Robinson, 1983; Florida State Department of Education, 1980; Smith, 1983; Weber & Roff, 1983). This meta-analysis enabled the BTAP developers to draw upon the findings from hundreds of individual research studies (BER, 1983). This research base was used as the basis for both the construction and scoring of the STITC-H.

Hays (1988) found that experienced teachers (n=46) performed significantly higher than student-teachers (n=30) in five of the thirteen competencies as well as total test score. Both of these groups (n=76) who were trained in teacher education programs scored significantly higher than students without training in education (n=31) on the total test and nine of the thirteen competencies. These findings suggest that what this test measures is something which teachers learn from their training or from their experience as teachers. Also, because of the test's research base, it may be seen as an indication of the degree to which teachers apply research findings to teaching problems under the conditions simulating interactive decision-making found in the STITC-H.

What are the reasons for the initial success of the STITC-H? Do both teacher training and experience play critical roles in the development of teachers' knowledge and ability to apply research findings to classroom situations?

It seems reasonable to believe that training would serve a significant role in the development of this knowledge. Teacher educators have begun to recognize the value of teacher effectiveness research and to develop ways of integrating it into the teacher education curriculum (Ross & Kyle, 1986). When teachers are given this type of information, they tend to increase their use of these behaviors (Needles, 1980). Professional education courses have also been identified as the primary source of important teaching behaviors by practicing teachers (Hoffman and Roper, 1985; Clark, Smith, Newby and Cook, 1985).

The evidence in support of teaching experience appears less conclusive. Generally, studies have indicated that experienced teachers exhibit behavior more consistent with the literature on teacher effectiveness than beginning teachers (Ayers, 1980; Adams, 1982; Fagot, 1978; Osborne, 1985; Housner & Griffey, 1985). However, others suggest that experienced teachers tend to select those specific behaviors which are most practical and immediate, which are not always consistent with the knowledge base (Huberman, 1985; Lortie, 1975) and teachers have not been the best judges of effective teaching behaviors identified in the research (Coker, Medley, Soar, 1980). This may be because teachers tend to work in non-collegial environments and lack access to existing bodies of research (Rosenholtz & Smylie, 1984).

A test that measures this knowledge and ability to apply research findings to classroom situations would be of great value for both teacher evaluation and program evaluation purposes. Teacher educators would be able to assess their students to provide specific feedback directly to them as well as obtain information regarding the strengths and weaknesses of their training program. Local education agencies would be able to screen applicants as well as assess experienced teachers.

Statement of the Problem

This study aimed to verify and extend the findings obtained during the field test of the STITC-H test. The major emphasis in this study was to examine the construct validity of the Simulation Exercise in Classroom Decision-Making. In establishing construct validity, it was theorized that this test was a measure of generic knowledge about research-based interactive teaching competencies. It was proposed that this generic knowledge may be learned through university teacher-training programs, the reading of the research literature on effective teaching, and experiences as classroom teachers.

Therefore, it was hypothesized that: 1) scores of university students with training in education differ from scores of students without training in education, 2) scores of subjects with classroom teaching experience (student teachers and teachers) would differ from those of subjects

without classroom teaching experience (non-education and education students) and 3) scores of experienced teachers would differ from scores of student teachers.

The following research questions were addressed in this study.

1. Do scores on the Simulation Exercise in Classroom Decision-Making reflect differences in the ability to apply teacher effectiveness findings to teaching problems between subjects with professional training and those without such training?
2. Do scores on the Simulation Exercise in Classroom Decision-Making reflect differences in the ability to apply teacher effectiveness findings to teaching problems between subjects with classroom teaching experience and those without experience?
3. Do scores on the Simulation Exercise in Classroom Decision-Making reflect differences in the ability to apply teacher effectiveness findings to teaching problems between experienced classroom teachers and student-teachers?

The STITC-H test has been modified for use in this current study. This process required a series of item analyses to examine the internal consistency of each of each of the competencies. As a result of this process, the test was reduced from 65 to 42 problems, consisting of a total of

188 items which can be administered within an hour. The format of the test has also been changed from the original synchronized tape-slide presentation to a single videotape. The revised version will be entitled A Simulation Exercise in Classroom Decision-Making.

In order to address the three questions stated above, a two factor design with repeated measures on one factor was used to analyze scores of four groups of students and teachers from the state of Virginia. These groups included experienced teachers (T), pre-service teachers who had completed student teaching (ST), pre-service teachers who had not yet completed student teaching (PreST), and undergraduate students who were not enrolled in education training programs (NoED). Orthogonal comparisons were used to test for the significance of each of the three hypotheses.

Limitations of the Present Study

First, it should be recognized that the competencies on which this instrument was constructed do not cover the entire range and complexity of competence contained within teacher effectiveness research literature. Therefore this study was conducted within the constraints of the twelve competencies identified and no attempts were made to generalize beyond these limits. Second, this study did not deal directly with teaching behavior, only with teachers' indications of how they would behave in specific classroom

situations, which can best be regarded as a reflection of their perceptions of effective teaching practices.

CHAPTER 2

Review of Literature

This literature review examined research related to three major questions. The first section investigates the extent to which teachers' perceptions of effective teaching are consistent with the research on effective teaching. The second section examines the degree to which teachers' actual behavior follows the practices outlined in the literature and the relationship between teachers' behavior and their amount of training and classroom experience. The last section reviews the of the literature pertaining to validity of existing tests of teachers' professional knowledge.

Teachers' Perceptions and Teacher Effectiveness

Research studies conducted to determine which teaching behaviors might lead to increased student achievement have varied a great deal both in terms of their design and the results reported from them. As these studies have accumulated, several reviews, summaries and lists of effective teaching behaviors have been compiled (Brophy & Good, 1986; Crawford & Robinson, 1983; Gage, 1978, Medley, 1977, Smith, 1983; Weber and Roff, 1983). These reviews have served as the basis for state assessment systems such as the Beginning Teacher's Assistance Program (BTAP) (Bureau of Educational Research, 1984) and the Florida Performance Measurement System (Florida State Department of Education, 1983).

Although there appears to be a common core of research findings pertaining to teacher effectiveness, there is little evidence that practicing teachers' perceptions are in agreement with these findings. It seems reasonable to think that experienced teachers may be in the best position to identify effective teaching behavior. Although this seems sensible, discrepancies between teachers' perceptions and research findings exist.

Marchant and Bowers (1988) investigated the extent to which teachers of various subjects, grade levels and years of experience agreed with a list of research-based effective teaching behaviors. The Teaching Behaviors Questionnaire, a list of fifty statements was used and teachers (N=300) were asked to indicate the extent to which they agreed with each statement using a five-point Likert-type scale. This questionnaire was based on an updated list of "teacher should" statements from Gage (1978), using Brophy and Good (1986) as a major resource in the revision process.

It was found that elementary teachers and teachers with little experience (6 or less years) were in higher agreement with the research-based teaching behaviors than secondary teachers and teachers with many years experience (25 or more). There were no significant differences found between groups in the areas of subject matter, SES, class size, student achievement level, and education level of teacher.

The general decline in agreement as years of experience

increased might be due to a "burn-out" factor or a change in the historical perspective of effective teaching which is presented in teacher preparation programs. Perhaps the best explanation for the higher agreement for elementary teachers is that most of the process-product research studies were conducted at lower grade levels (Medley, 1977) and may not generalize to higher grade levels.

A second reason could be a matter of training. Greater emphasis is placed on the teaching-learning process within the elementary preparation programs. Elementary teachers may be more oriented toward effective teaching behaviors as a result of the nature of their preparation. One study conducted by Thompson (1981) suggested that elementary teachers placed more emphasis on teaching methods which advance the social growth of the student while secondary teachers emphasize intellectual, analytical skills.

In addition to being better trained, elementary teachers are also seen as more effective than secondary teachers. Jandes, Murphy, & Sloan (cited in Marchant & Bowers, 1988) concluded that Illinois elementary teachers, principals and superintendents perceived their schools over 60% higher in overall effectiveness than their secondary counterparts. Richardson (1985) found that both teachers and principals perceived elementary teachers as more effective in classroom management than secondary teachers.

In general, teachers have not been the best judges of

effective teaching. When teachers were asked to identify teaching behaviors they perceived as effective, only about half of the behaviors identified were found to be consistent with the research findings (Coker, Medley, Soar, 1980).

What are the reasons for these discrepancies with the research findings? Some findings indicate that teachers simply tend to select behaviors that are practical and immediately useful instead of applying the concepts or principles revealed by research studies on teacher effectiveness (Huberman, 1985; Richardson, 1985; Lortie, 1975). These teachers tend to rely on experience, experimentation, and fellow colleagues' "tricks of the trade" as their primary sources of professional knowledge and skill acquisition. It has also been suggested that teachers tend to work in non-collegial environments and lack access to existing bodies of research (Rosenholtz & Smylie, 1984).

Teachers' Behavior and Experience

Research regarding the comparisons between novice and experienced teachers indicates that they have different perceptions of classroom situations and the students in those classrooms (Clark and Peterson, 1986). Calderhead (1981) made some distinctions between experienced and novice teachers regarding their understanding and interpretation of classroom situations describing beginning teachers as more narrow in their assessments of what was going on in the

classroom. Berliner (1986) suggested the development of a schema as a distinguishing factor between novice and expert teachers. As a result, novice teachers tended to pay attention to a greater variety of cues, see and describe the situation in a literal sense whereas experienced teachers have the ability to zero in on the critical cues, make sense of the situation more quickly, and draw proper inferences. Experienced teachers have also been reported as being more frequently aware of student performance cues such as student errors, deficient responses and attention level (Fogarty, Wang, & Creek, 1983).

Other studies of expert teachers have revealed the importance of routines in the teaching process. Over the course of a 3 1/2 month period, both expert teachers and novices were observed, videotaped, and interviewed in order to capture the routines and functions that they perform in their classrooms (Leinhardt & Greeno, 1986). Expert teachers were selected from a group of teachers whose students exhibited the greatest amount of academic growth over a five year period and novices were chosen from a pool of student teachers. The expert teachers were described as having a large repertoire of well scripted routines which enabled them to open the instructional period more efficiently, review the assigned homework with less difficulty, spend less time lecturing and more in guided practice, and complete the lesson with fewer management

problems. Another study (Brooks & Hawke, 1985), reported by Berliner (1986), supported the importance of routines for seventh grade math teachers in opening an instructional period. These routines regarding classroom management, support and permitted teacher-learner exchanges are typically established during the first few days of school (Leinhardt, Weidman, & Hammond, 1987).

If experienced and beginning teachers have different perceptions of the classrooms and the teacher effectiveness literature, it follows that they may respond differently in various classroom situations. The attitudes that teachers have toward teaching can affect their behaviors and ultimately the achievement of their students. Studies investigating the behavior of teachers and teaching experience have been done on both a longitudinal (Adams, 1982; Ayers, 1980; Sandefur, 1969) and cross-sectional basis (Fagot, 1978; Fogarty, 1983; Housner & Griffey, 1985; Osborne, 1985; Pearce, 1988; Young, 1979).

Generally these studies have indicated that teachers with more experience demonstrate behavior more closely associated with that identified by the process-product research literature (Ayers, 1980; Adams, 1982; Fagot, 1978; Osborne, 1985; Housner & Griffey, 1985). However, others suggest that experienced teachers tend to select those specific behaviors which are most practical and immediate, which are not always consistent with the knowledge base

(Huberman, 1985; Lortie, 1975).

Ayers (1980) used the Classroom Observation Record (COR) and reported that teachers increased the time-on-task and their general affective approach over a five year period. Using the same instrument, Adams (1982) reported that teachers became more organized, empathetic, and their students became more alert and confident during a six year period from student teaching through their fifth year of teaching. Both Ayers (1980) and Adams (1982) identified an increase in the degree of student verbal behavior as teachers became more experienced.

Sandefur (1969), using a modified version of the COR, concluded that teachers became more friendly, understanding, taught with more originality, were judged to be more poised, confident, mature, responsive, optimistic and demonstrated a greater breadth in their teaching from the time of student teaching and the end of their first year. In another study involving student teachers, Young (1979) found that the lessons taught by teachers in their first year (n= 9) and teachers with two or more years of experience (n = 13) were more rapidly-paced and interactive than those taught by student teachers (n = 17). Student teachers tended to reprimand students more often than the other two groups of teachers.

More specific changes pertaining to reinforcement and questioning skills were identified by Fagot (1978) as

teachers with 3 or more years and some training in early childhood were reported as directing students more often, responding more often to children's questions, asking more questions, and making more favorable comments to a group of preschool children than teachers with less experience and no training. Osborne (1985) revealed that experienced teachers were more task-oriented, made fewer demands for attention, and provided more instruction when interacting with a group of children exhibiting distractible behavior than those with no formal teaching experience.

Using videotaped lessons of teachers in small groups, Fogarty, et al. (1983) reported that experienced teachers, with an average of 10 years experience, were more likely to respond to student cues with an instructional response whereas novices were more likely to use a management approach. Experienced teachers were also able to attend to a greater number of these cues and considered a greater variety of options, primarily instructional, when making classroom decisions. Housner and Griffey (1985) conducted stimulated recalls with eight inexperienced and eight experienced physical education teachers. Experienced teachers were more likely to focus their attention on individual student performance cues and respond with feedback and instruction while inexperienced teachers concentrated on the class as a whole failing to recognize the need for individual attention.

Pearce (1988) investigated the performance of teachers with various degrees of experience on the fourteen competencies included in the Virginia Beginning Teacher's Assistance Program (BTAP). Teachers with less than two years experience scored highest on nine of the fourteen competencies, teachers with 3 to 5 years experience scored highest in four and the most experienced group, those with 11 or more years, scored highest on one competency. The competencies relating to classroom management (consistent rules, awareness, and reinforcement) defined the significant discriminant function. This function revealed that the group of teachers with 11 or more years experience performed significantly lower than the other three groups with less experience.

With the exception of Ayers (1980), Adams (1982), and Pearce (1988), studies pertaining to teacher performance and amount of classroom teaching experience have focused on specific teaching behaviors and were confined to a specific time and teaching context. These studies also failed to provide a consistent definition of what is meant by an experienced teacher. Experienced teachers have been identified as having as few as one year (Sandefur, 1969) of teaching to eleven or more years of teaching experience (Pearce, 1988). A great deal of variability exists regarding sample size, ranging from as little as 8 to as many as 644, complicated by the fact that these samples were

not randomly selected.

Given these limitations, some consistent patterns regarding the changes in teaching behavior of teachers with varying degrees of experience have been established. Experienced teachers were consistently identified as being more positive and effective than novice teachers. Their behavior was also described as being more stable. Finally, experienced teachers failed to demonstrate, to any great extent, negative behavior such as reprimanding and losing control of the class while beginning or novice teachers were associated with such behavior.

One possible reason for these differences in the performance of novice and experienced teachers may be that teachers may learn different behaviors from different sources. Some studies suggest that teachers rely primarily on their own teaching experiences, experimentation, and fellow colleagues, tending to abandon or at least move away from their professional preparation and student teaching experiences (Clark, et al, 1985; Huberman, 1985; Lortie, 1975). Other studies report that professional education courses and student teaching experiences were of most importance in the development of teaching skills (Clark, Smith, Newby and Cook, 1985; Hoffman & Roper, 1985; Hoste, 1985).

It is important for teachers at all levels of teaching experience to be kept informed of current research knowledge

from which applicable teaching behaviors and skills may be derived. Unless professionals remain current with new knowledge provided by research studies, their competence can only be in a transitory state (Andrews, 1980). Measures of such professional knowledge or competence currently exist in part to determine the extent to which what teachers learn on the job and through their professional training is consistent with current research-based pedagogical knowledge.

It is therefore important to have a measure of such knowledge to monitor teacher training programs and diagnose experienced teachers' knowledge and ability to apply research findings to classroom situations. Existing statewide programs use tests which claim to measure this knowledge. Is there any evidence that they do? This will be discussed in the following section.

Tests of Professional Knowledge

Thirty-eight states currently mandate the testing of teachers' knowledge of subject-matter or pedagogy for admission to teacher preparation programs or for state certification (Association of Teacher Educators, 1988). Of these 38 states, 32 use one form or another of the NTE. Other nationally standardized tests used by some states include the Scholastic Aptitude Test (SAT), California Achievement Test (CAT), and the California Basic Educational Skills Test (CBEST). State-developed tests are currently

being used in eight states including Alabama, Arizona, Connecticut, Florida, Indiana, Pennsylvania, South Carolina, and Washington. The content of these state-developed tests consists of basic skills, general knowledge, professional knowledge, and subject-matter knowledge.

The state of North Carolina has required the NTE as a prerequisite for certification since 1964. Georgia began developing plans for the assessment of teachers in 1970 followed by Florida in 1978 and the states of Arkansas, South Carolina, and Tennessee in 1979. The number of states mandating teacher competency testing grew to 12 in 1980: 22 in 1981; 28 in 1982; 30 in 1983; and 38 in 1987, which 8 of the remaining 12 states drafting proposals. (Association of Teacher Educators, 1988; Eisenberg & Rudner, 1988; Sandefur, 1985).

Most validity studies of tests of professional knowledge have been predictive or concurrent in nature. These validation studies have been done in efforts to establish relationships with other measures of professional knowledge or teacher performance in the classroom. A few studies have been undertaken to relate test scores to teacher training and teaching experience.

Results from three statewide samples have indicated that teachers from liberal arts programs have generally scored higher than those from teacher education programs (SREB, 1984). Two studies examining the Weighted Common

Examination Total (WCET) scores from the NTE were conducted in the states of Louisiana and North Carolina. In both cases, teachers from liberal arts programs scored higher than teachers from teaching training programs. One exception was found on the Specialty Area Test in Elementary Education on which those trained teachers performed higher in the state of Louisiana. A study conducted in Georgia used the Georgia Teacher Certification Tests as a basis for comparison. These results also indicated that teachers with a bachelors degree in the arts and sciences scored slightly higher than teacher education graduates, with the exception of communicative arts (SREB, 1984).

Predictive and Concurrent Validity of the NTE

Although the NTE has been in existence since 1940 and has been used heavily in the assessment and certification process of teachers, estimates of its predictive and concurrent validity have generally been very low. An extensive review of predictive and concurrent validity studies conducted during the first thirty years of the NTE was compiled by Quirk, et al. (1973). Quirk, et al. (1973) reported a median correlation of .05 between NTE WCET scores and ratings by college supervisors and principals during student teaching. Correlations between NTE scores and principal ratings during the first year of teaching were not much better having a median value of .11.

Chiu (1989) examined sixteen predictive validity

studies which reported a total of 168 bivariate correlations of the WCET with classroom performance ratings by principals, students and with low-inference observational systems. The median correlation was .15. A review of three predictive validity studies (79 bivariate correlations) using the Core Battery from the revised NTE tests resulted in a median of .09.

Chiu also studied the relationship between NTE tests and the fourteen competencies included in the Virginia Beginning Teachers Assistance Program (BTAP). Chiu's efforts concentrated on the new NTE exams which were revised in 1982. Four simple correlations were computed between the standardized sum of 16 BTAP competency scores and the Core Battery tests. Two significant, but low correlations were found between BTAP total score and 1) Professional Knowledge (PK) scores (r= .125) and 2) the difference between Professional Knowledge (PK) and General Knowledge (GK) scores (r = .209). In addition, a multiple correlation of .224 was found between BTAP total score and a weighted combination of the Core Battery tests.

Two primary reasons for the poor predictive ability of the NTE were offered by Quirk, et al (1973). First, the NTE only measures a sample of the important qualities related to successful teaching, most of which have more to do with planning strategies and class management rather than subject matter knowledge. Second, the ratings scales used for on-

the-job evaluations are very unreliable.

Validity of State-Developed Tests

A number of states have attempted to develop their own tests of teachers' competence, measuring knowledge of basic skills and of subject matter areas, because the NTE tests do not fit their certification system. Eight states have developed written tests for admission or certification purposes while 16 others are in the planning stages (Association of Teacher Educators, 1988).

Of the eight state tests already developed, only four require the completion of a test of pedagogy or professional knowledge. Those states which include a test of professional knowledge include Arizona, Florida, Pennsylvania, and Washington. No empirical evidence of predictive or concurrent validity was found in the literature regarding these state-developed tests.

Other Tests of Professional Knowledge

Other tests reported in the literature have a tendency to assess more specific professional knowledge as listed by Harris (1981). These tests include the Behavior Objective Writing Skills Test, the Curry-Geis Syllabification Skills Test, the Illinois Tests in the Teaching of High School English, Inventory of Teacher Knowledge of Reading, and the Multiple-Choice Items for a Test of Teacher Competence in Educational Measurement. Three tests were found that measured professional knowledge and skills in more general

contexts. These included the Creative Teaching Dilemma (Riley, 1982), the Preactive Decision Exercises (McNergney, Medley, Aylesworth, & Innes, 1983; Saunders, Herbert & Chronister, 1989), and the Teaching Situation Reaction Test (Frve, 1972; Murray, 1969).

Preactive Decision Exercises

The Preactive Decision Exercises (PreDEx) was developed as a pencil and paper test designed to measure teachers' planning abilities (McNergney, et al., 1983) and has been revised as a computer simulation (Saunders, et al., 1989). It contains a set of planning problems that teachers may face in their classrooms and provides a number of alternative courses of action which a teacher might take. Each teacher is asked to indicate whether or not each alternative is congruent or incongruent with their professional knowledge and philosophy. Curriculum outlines, class schedules, a school philosophy statement, and student records are available for teachers to examine before making a decision about each alternative.

PreDEx offers a number of advantages as a measure of professional knowledge. PreDEx claims to simulate the actual classroom more closely than multiple-choice tests do because its format permits more than one congruent or incongruent course of action for each situation. Another advantage pertains to the flexible scoring system used by PreDEx. In recognition that there is no general consensus

regarding the proper way to plan or what type of knowledge is necessary to plan effectively, a scoring system was incorporated that could be adjusted to meet the philosophies and curriculums of different teacher training programs. Finally, some preliminary evidence of the test's discriminant validity has been presented as experienced teachers have performed better than inexperienced teachers and non-education students (Saunders, et al., 1989).

Teaching Situation Reaction Test

The Teaching Situation Reaction Test (TSRT) is a forced-choice pencil and paper test developed in 1957 by Duncan and Frymier (Frye, 1972). Since that time the TSRT has been refined, expanded, rekeyed, and undergone several other revisions. The fourth edition of the TSRT consisted of 48 classroom situations to which the examinee would respond by ranking four possible solutions. The situations included on the test are related to the areas of planning, classroom management, and teacher-pupil relationships.

Several predictive validity studies have been conducted and significant relationships with supervisory ratings have resulted in five of six studies involving practicing teachers and two of three studies involving pre-service teachers (Murray, 1969). The TSRT has also been correlated with a number of attitude tests and scales, correlating positively with the factors of objectivity r= .13), empathy (r= .18), and control (r= .28) (Murray, 1969). The TSRT has

also shown a positive relationship with students' grade point averages and change in acceptance of others.

Creative Teaching Dilemma

The Creative Teaching Dilemma was developed by John Riley in 1982 to measure creative teaching abilities based in part on the work in creativity by Torrance. The test is comprised of a series of teaching dilemmas for which the examinees must gather additional facts, identify the problem(s), generate possible solutions and state a final solution. The test is scored according to two criteria: fluency and flexibility. Fluency is described as the number of alternative solutions generated and flexibility is defined as the ability to identify a variety of types of problems and generate solutions from different points of view.

The Creative Teaching Dilemma does recognize the existence of more than one acceptable answer but does so in a very subjective way. The credit given for the number of solutions presents little difficulty but, more specific guidelines are needed to measure the quality of these solutions. Riley (1982) reports 100 percent agreement on fluency scores among five judges but, only 60% agreement regarding the flexibility scoring.

In a predictive validity study, the Creative Teaching Dilemma was reported as correlating .39 with the classroom management component of the Georgia Teacher Assessment

Instrument (Riley, 1982). Correlations have ranged from .31 to .60 with the Torrance Tests of Creative Thinking Demonstration Form.

Simulation Test of Interactive Teaching Competencies - Hays

The Simulation Test of Interactive Teaching Competencies - Hays (STITC-H) was developed by Hays (1988) as a synchronized audiotape-slide test of generic teaching competencies (see Appendix A). Sixty-five problems depicting classroom situations are presented and the teacher is required to respond to four possible courses of action as either "appropriate" or "inappropriate" for each problem. To simulate the speed with which decisions are made in the classroom, the examinee is given only five seconds to respond to each alternative.

Hays (1988) found that subjects with teacher preparation training (n=76) scored significantly higher than those without training (n=31) on nine of thirteen competencies as well as on the total test score. These competencies included Academic Learning Time, Clarity of Structure, Individual Differences, Evaluation, Learner Self Concept, Meaningfulness, Reinforcement, Close Supervision, and Awareness. The influence of teaching experience was also revealed as experienced teachers (n=46) scored significantly higher than student teachers (n=30) on the total test and on the competencies of Clarity of Structure, Individual Differences, Affective Climate, Meaningfulness,

and Close Supervision. These findings suggest that this test does measure some things which teachers learn from their training or from their experience as teachers.

Summary of Professional Knowledge Tests

Multiple-choice tests offer many advantages over other measures in that they are easy to administer, easy to score, objective, and relatively inexpensive. The NTE and the four state-developed tests offer an additional advantage in providing national (NTE) and state norms. Multiple-choice tests also have their limitations. Generally, correlations between pencil and paper tests and classroom performance have been very low.

The Creative Teaching Dilemma and the STITC-H are the only tests of interactive professional knowledge found that recognize the existence of more than one "correct" answer. The PreDEx test does this for planning situations in which teachers generally have more time and information on which to base their decision. In addition to simulating the reality of the classroom by providing the option for more than one correct response, the STITC-H restricts the time period in which teachers had to make an interactive teaching decision, as in the actual classroom.

Summary

Several inconsistencies with the research findings of effective teaching have been revealed among groups of teachers. Specifically, elementary teachers and teachers

with little teaching experience were in higher agreement than secondary teachers and teachers with many years experience (Marchant & Bowers, 1988).

Experienced teachers have consistently been identified as being able to establish routines early in the school year (Berliner, 1986; Leinhardt, et al., 1987; Shavelson, 1983) that enabled them to be more positive and effective (Adams, 1982; Ayers, 1980; Sandefur, 1969), more interactive and responsive to students (Fogarty, et al, 1983; Housner & Griffey, 1985; Osborne, 1985; Young, 1979) than novice teachers.

The differences in the perceptions and behaviors of teachers may be a function of training or simply due to years of teaching experience. Teachers' behaviors have been attributed to both professional education courses (Clark, Smith, Newby and Cook, 1985; Hoffman & Roper, 1985; Hoste, 1985) and general teaching experiences, experimentation, and colleagues (Clark, et al, 1985; Huberman, 1985; Lortie, 1975).

States have taken an increased responsibility in the assessment of teachers, generally relying on the NTE or other state-developed tests. The support for the validity of these tests has been scarce as relationships with supervisory ratings, and performance-based observational systems have been very low (Chiu, 1989; Quirk, et al., 1973). Other tests have been constructed for assessing the

professional knowledge of teachers and have failed to offer sufficient evidence of validity. One test, STITC-H, has established some credibility as a measure of teachers' competence (Hays, 1988) and preliminary findings have suggested that this test measures something which teachers learn from their training or from their experience as teachers. The original STITC-H test, as well as the revisions made for its use in this study, will be discussed in the next chapter outlining the methods and procedures of the study.

CHAPTER 3

Methods

The purpose of this study was to examine the construct validity of a revised STITC-H test called the Simulation Exercise in Classroom Decision-Making. In establishing construct validity, it was hypothesized that this test was a measure of generic knowledge of research-based findings about teacher competence. It was further hypothesized that this generic knowledge may be learned through university teacher-training programs, the reading of the research literature on effective teaching, and experiences as classroom teachers.

Therefore, it was hypothesized that: 1) the scores of university students with training in education would differ from scores of students without training in education, 2) the scores of those subjects with classroom teaching experience (student teachers and teachers) would differ from those of college students without classroom teaching experience (non-education and education students) and 3) the scores of experienced teachers would differ from student teachers.

Instrumentation

A revised version of the Simulation Test of Interactive Teaching Competencies-Hays (STITC-H) was used as a measure of teachers' ability to identify courses of action pertaining to specific classroom teaching situations which

are congruent with research findings. The original STITC-H consisted of 65 problems depicting situations that occur in the actual classroom. These problems were developed by Hays and colleagues and then reviewed by a panel of teachers and administrators before the final selection of items was made (Hays, 1988). Each problem was presented in both a visual and auditory format using a slide projector and a cassette player. Four possible alternative courses of action were presented on audiotape following each problem and the examinee was given five seconds to judge whether each alternative was appropriate or inappropriate.

Originally, the STITC-H contained 260 such alternatives to the 65 problems, scored on thirteen competency keys. Items were retained for use in the present study if they met two main criteria: 1) they contributed to the internal consistency of their respective competency, and 2) reflected the content which that competency proposed to measure. The structure of the STITC-H imposed one additional restriction in this refinement process. Because each problem included four possible alternatives, items were evaluated in groups of four.

Cronbach's coefficient alpha for each of the thirteen competencies included on the original STITC-H is shown on Table 1. These internal consistency estimates were used as the basis for further refinement of the test. It was from this point that items were evaluated to determine whether

they would be retained, revised or eliminated. The reduction process that followed involved several decisions which will now be explained.

Table 1

Reliability Estimates for Original STITC-H

Competency	Number of Items	Alpha
Academic Learning Time	29	.67
Accountability	26	.40
Clarity of Structure	16	.65
Individual Differences	22	.48
Evaluation	16	.41
Consistent Rules	15	.55
Affective Climate	22	.73
Learner Self-Concept	17	.55
Meaningfulness	17	.55
Questioning Skills	16	.12
Reinforcement	31	.71
Close Supervision	16	.38
Awareness	17	.48

The first decision was to eliminate the competency of Questioning Skills. This was done for two reasons. First, as indicated on Table 1, the reliability coefficient (alpha) for this competency was extremely low (.12) indicating a lack of internal consistency. The second reason dealt with the nature of teacher's questioning behavior. Questioning skill is a context dependent behavior and very difficult to depict in the present format of the test. As a result of the elimination of this competency, 16 items comprising four problems were removed from the test.

The second stage in the item reduction process was to eliminate all those items that correlated negatively with their assigned competency. During this stage, items were evaluated in sets of four, as they appeared on the STITC-H. For example, if three of the four items comprising a given problem correlated either negatively or very low ($r < .10$), the problem was eliminated. Otherwise, only the specific items were eliminated, leaving some problems with less than four alternatives. At the conclusion of this stage, the test consisted of 142 items covering 47 problems.

These 142 items were then carefully examined for content. The list of competency definitions and indicators found in Appendix A was used as a reference during this stage. Several items were targeted during this stage and were either tested on other competency keys or eliminated. Items which resulted in negative or low correlations after

the revised item analyses were also removed at this time. It was during this time that additional items were written to be included on the test. These new items were targeted toward those competencies and problems which needed the greatest assistance.

The final form of the test contains 42 problems with a total of 188 items (See Appendix B for selected items). Of these 188 items, 125 had been retained from the original STITC-H and 63 were developed to be field tested in this study. The final reliability estimates based on the 125 retained items are presented in Table 2. This table also indicates the number of new items written for each competency. Further item analysis will be necessary using the data obtained from the sample in this study to determine the final reliability estimates for each of the twelve competencies.

Table 2

Reliability Estimates After Item Reduction

Competency	Original Items	Alpha	New Items	Total
Academic Learning Time	15	.69	3	18
Accountability	12	.50	10	22
Clarity of Structure	9	.72	2	11
Individual Differences	8	.51	11	19
Evaluation	6	.40	7	13
Consistent Rules	7	.48	4	11
Affective Climate	17	.82	0	17
Learner Self-Concept	11	.72	11	22
Meaningfulness	7	.58	6	13
Reinforcement	15	.69	2	17
Close Supervision	12	.58	0	12
Awareness	6	.32	7	13
TOTAL	125	.92	63	188

Data Collection

Sample

Four groups of subjects participated in the simulation exercise. These four groups included 1) experienced practicing teachers (T), 2) pre-service teachers who have completed student teaching (ST), 3) pre-service teachers who had not yet to begin student teaching (PreST), and 4) university students from major areas other than education (NoED). Pre-service teachers without student teaching experience were not included in the field test of the STITC-H. Their inclusion in this study provided for further discrimination among levels of experience in education, a direct comparison with those students not enrolled in education courses as well as additional opportunities for future predictive validity studies.

The samples of pre-service teachers (PreST and ST) were selected with the assistance of the Commonwealth Center for the Education of Teachers from three cooperating institutions affiliated with the Commonwealth Center. These institutions will be referred to as School A, School B, and School C throughout the remainder of this study. Students who had not yet student taught (PreST) were sampled from pre-service teacher education courses at participating institutions. Only those students who had completed at least four or more professional education courses were considered for this sample. The sample of student teachers

(ST) was also selected from participating colleges and universities affiliated with the Commonwealth Center. These students had either completed their student teaching experiences during the fall of 1989 or were currently completing their student teaching experience. Students selected from specific institutions remained intact and subgroup comparisons were made before any further analysis was considered.

Practicing teachers were selected from school divisions. Teachers were sampled from both elementary and secondary schools and the test was administered during faculty meetings and days of early dismissals. These teachers had varying degrees of classroom teaching experience, ranging from 2 to 28 years with both a mean and median of 15 years. The influence of experience was examined within this group of practicing teachers before comparisons were made with other groups in this study.

The sample of undergraduate students majoring in areas other than education (NoED) was selected primarily from a counseling course offered by the School of Education at the University of Virginia. Over 150 students were enrolled in this class with over 90% of them being undergraduates majoring in areas other than education. Comparisons were made across all levels of undergraduate experience within the NoED group before making comparisons with the PreST group of students.

Each subject was asked to complete a background questionnaire after completing the test (see Appendix C). Information gathered from this instrument included age, sex, educational background, teaching experience, and items regarding the nature of the simulation exercise. Each person was also given the opportunity to receive a profile sheet of his test results by including his name and mailing address on this questionnaire.

Test Administration

The Simulation Exercise in Classroom Decision Making was given to the groups of subjects identified above during March and April of 1990. The test was given primarily in two settings: 1) classrooms at those colleges participating in the study, and 2) faculty meeting rooms at participating school divisions. Tests were administered to groups of approximately twenty-five at a time using a video cassette player and several monitors placed throughout the room.

Instructions were given on the video tape, both visually and orally. A copy of these instructions can be found in Appendix B. Subjects recorded their answers to the 188 items on an opscan sheet. It took approximately 60 minutes to administer this test, including instructions and the background information form.

Scoring

The answer sheets were scored using an optical scanning machine. Each alternative was judged as either appropriate

or inappropriate based on support from the research. Scores were compiled for both the total test and specific competencies. Each intact group of students and teachers participating in this study will be given the opportunity for a debriefing session with the study's Principal Investigator. During this session, an overview of the study's results will be presented.

Data Analysis

The analysis of data took place in three phases. The first phase was to examine the internal consistency of each of the twelve competencies comprising the instrument. Secondly, the sample was refined as the four groups were examined for their homogeneity. Finally, the construct validity of the instrument was evaluated as three major hypotheses were tested.

Coefficient alpha was calculated to estimate the internal consistency of each of the twelve competencies. A two factor design with one repeated measure was used in the second and third phases.

Reliability

The primary means of assessing the reliability of this test was that of internal consistency. Estimates of internal consistency were made with the calculation of Cronbach's coefficient alpha. These calculations were made for the total test score and each specific competency score. The obtained alpha is an indication of the extent to which a

given set of items measure the same factor (competency) and can be presented as evidence of both internal consistency and construct validity. Kane (1986) recommends a criterion of .50 for criterion referenced tests. The 63 new items to be field tested in the current study were added to the existing competencies and coefficient alpha was calculated for each competency. Items which contributed to these keys were retained while others will be further examined, revised, rewritten, or discarded.

Preliminary Analysis

Descriptive statistics were calculated for each of the groups participating in this study. These statistics included measures of central tendency and variability such as means, standard errors of the mean, standard deviations and ranges. These measures were calculated for the total test as well as each competency and were used to compile individual profiles. Each individual was also given the opportunity to receive an individual profile of his or her performance upon request. These individual profiles included the number and percentage of items correct and an indication of their performance relevant to each of the four groups participating in this study.

The homogeneity of each of the four main groups comprising this study's sample was examined using a two factor design with one repeated measure (Winer, 1971). This design examined overall group performance as well as

patterns over the twelve competencies. This design is illustrated in Figure 1. Any subgroup found to be heterogenous with the other subgroups was identified as a separate group and not combined with the other subgroups for the overall analysis.

Groups	Subjects	Competencies C1	...	C4	...	C8	...	C12
Group 1	S_1	...	S1	...	S1	...	S1
	:		:		:		:
	:		:		:		:
	:		:		:		:
Group 2	:		:		:		:
	:		:		:		:
	:		:		:		:
	:		:		:		:
Group 3	:		:		:		:
	:		:		:		:
	:		:		:		:
	:		:		:		:
Group 4	Sn	...	Sn	...	Sn	...	Sn

Figure 1 - Two Factor ANOVA with One Repeated Measure

Validity

Two types of validity were assessed in this study; content validity and construct validity. Content validity was most heavily evaluated when the STITC-H test was constructed. The specifications for the original STITC-H were constructed from the indicators of successful teaching specified by the Beginning Teachers Assistance Program (BTAP). The research basis for these items has been maintained throughout the revision of the STITC-H test. The revisions were also made in accordance with the BTAP competencies and indicators that appear in Appendix A.

Evidence of construct validity will be provided by affirmative answers to each of the three research questions included in this study. Construct validity was examined separately within elementary and secondary grade levels. The three questions posed in this study are listed below.

1. Do scores on the Simulation Exercise in Classroom Decision-Making reflect differences in the ability to apply teacher effectiveness findings to teaching problems between subjects with professional training and those without such training?
2. Do scores on the Simulation Exercise in Classroom Decision-Making reflect differences in the ability to apply teacher effectiveness findings to teaching problems between subjects with classroom

teaching experience and those without experience?

3. Do scores on the Simulation Exercise in Classroom Decision-Making reflect differences in the ability to apply teacher effectiveness findings to teaching problems between experienced classroom teachers and student-teachers?

Each of these three questions was answered in terms of total test score and the scores of specific competencies. It the test measures teachers' knowledge of research findings in teacher competence, the scores should be affected by the training, teaching experience, or by both. Such differences will therefore be taken as evidence of construct validity.

A two factor design with one repeated measure was used to test the construct validity of the test in the current study (Winer, 1971). This design is often referred to as a mixed model with one between-subjects factor (group membership) and one within-subjects variable (competencies) (Myers, 1979). Scores on each competency were standardized to a mean of 50 and a standard deviation of 10, to make them comparable. This had the effect of fitting eleven means in the ANOVA - instead of the usual grand mean, and eliminated the main effect of B (competencies). A summary of the components of this design is presented in Table 3.

Table 3

ANOVA Summary Table

Source of variation	df	E(MS)	F-ratio	
Between Subjects	na - 1			
Groups (A)	a - 1	$o^2_e + no^2_a$	MSa / MSs/a	
Subjects w/ groups (S/A)	a(n-1)	o^2_e		
Within Subjects	na(b-1)			
Competencies (B)	b - 1	$o^2_e + no^2_b$	MSb / MS bxs/a	*
Group X Competencies Interaction A X B)	(a-1)(b-1)	$o^2_e + no^2_{ab}$	MSab / MS bxs/a	
Comp X Subjects w/ groups (B X S/A)	a(n-1)(b-1)	o^2_e		

* omitted - see text

Table 4

ANOVA Summary Table (After Partitioning)

Source	df	E(MS)	F-ratio
Between Subjects	n(4- 1)		
Groups (A)	3		
Contrast 1 (training)	1	$o^2e + no^2_{a1}$	$\frac{MSa_1}{MSs/a}$
Contrast 2 (classroom exp)	1	$o^2e + no^2_{a2}$	$\frac{MSa_2}{MSs/a}$
Contrast 3 (teaching exp)	1	$o^2e + no^2_{a3}$	$\frac{MSa_3}{MSs/a}$
Sub w/groups (S/A)	4(n-1)	o^2_e	
Within Subjects	n4(12-1)		
Group X Competencies Interaction (A X B)	33		
Contrast 1 (training)	11	$o^2_e + no^2_{a1b}$	$\frac{MS\ a_1\ x\ b}{MS\ b\ x\ s/a}$
Contrast 2 (classroom exp)	11	$o^2_e + no^2_{a2b}$	$\frac{MS\ a_2\ x\ b}{MS\ b\ x\ s/a}$
Contrast 3 (teaching exp)	11	$o_{2e} + no^2_{a3b}$	$\frac{MS\ a_3\ x\ b}{MS\ b\ x\ s/a}$
Comp X Sub w/groups (B X S/A)	4(n-1)(11)	o^2_e	

The principal analysis in this study used the design shown in Table 4, which incorporates three preplanned contrasts. This design produced six F-tests, three pertaining to the overall performance of the groups and three pertaining to the patterns across the twelve competencies. In order to answer the three specific questions in this study, this design was employed three times using the groups specified in each of the three orthogonal contrasts. This was necessary to partition the variance explained by group membership and determine the proportion of variance that each contrast independently explained. F-tests were then applied to determine the significance of each contrast. This process was repeated for the A X B interaction to determine whether differences exist among group profiles over the twelve competencies. Independent t-tests were applied within each of the three orthogonal group contrasts to determine the specific competencies on which the groups differ.

Hypotheses to be Tested

Six hypotheses were tested from the design shown in Table 4. Three hypotheses pertain to differences in overall group performance and three pertain to different group profiles over the twelve competencies. A difference in profiles means that group differences vary for different competencies, which implies that not all such differences are zero. These six specific hypotheses are presented

below.

Overall Group Differences

1. 0^2_{a1} (training) = 0

2. 0^2_{a2} (classroom experience) = 0

3. 0^2_{a3} (teaching experience) = 0

Group Profile Differences

4. 0^2_{a1b} (training X competencies) = 0

5. 0^2_{a2b} (classroom exp. X competencies) = 0

6. 0^2_{a3b} (teaching exp. X competencies) = 0

Summary

The analysis of the data has been described in three phases using two primary samples of subjects, elementary and secondary. First, reliability estimates were calculated for each of the twelve competencies using coefficient alpha as the primary measure of internal consistency. The second phase examined the homogeneity of the four major groups of subjects. The last phase regarded the construct validity of the simulation exercise as six hypotheses were proposed to be tested. A two factor design with one repeated measure was used in the second and third phases of the data analysis process. The rejection of any one of the hypotheses can be presented as evidence of construct validity of some part of the test.

CHAPTER 4

Results

The analysis of the data took place in three phases. The first phase examined the internal consistency of each of the twelve competencies which comprised the Simulation Exercise in Classroom Decision-Making. Secondly, the four major groups of subjects which constituted the sample in this study were examined for homogeneity and refined accordingly. Finally, the construct validity of the instrument was evaluated as three major hypotheses were tested. The results from each of these phases will be presented in this chapter.

Reliability

The internal consistency of the total test as well as each of the twelve competencies was estimated using Cronbach's coefficient alpha. Items from the original STITC-H test were retained on their respective competency scales and coefficient alphas were recalculated using data collected in the current study. The 63 additional items, written for this study, were assigned to specific competencies and revised alpha coefficients were calculated. The calculation of these reliability estimates took place in several steps.

First, alpha coefficients for the twelve competencies and total test were calculated using the 125 original items. The second step was to examine these initial internal

consistency estimates and eliminate specific items which were not supporting the overall competency. As a result of this second step, 14 items were eliminated which were found to have negative correlations with their respective competency total score. Third, the additional 63 items were correlated with the revised competency total scores from step two. One at a time, these new items were added to their assigned competencies in order of descending magnitude. Those items found to have negative item-total correlations were not included in this process. Forty out of the 63 items were found to contribute to specific competencies and were retained on the final scales.

Upon the completion of this process, the content of the items from each competency was reviewed and examined for consistency with the BTAP indicators of the respective competency. These specific BTAP indicators can be found in Appendix A. Clarity of Structure was eliminated at this time because only a few core items remained which failed to define the competency. Eleven of the twelve competencies were preserved having alpha coefficients ranging from .33 (Close Supervision) to .60 (Accountability) with a mean of .45. A summary of this process is presented on Table 5.

Table 5

Summary of Reliability Estimates

Competency	Step 1		Step 2		Step 3	
	Items	Alpha	Items	Alpha	Items	Alpha
Academic Learning Time	15	.43	12	.51	14	.56
Accountability	12	.48	12	.48	22	.60
Clarity of Structure	9	.20	8	.27	X	X
Individual Differences	8	.29	7	.32	10	.37
Evaluation	6	.12	4	.23	11	.41
Consistent Rules	7	.28	7	.28	11	.37
Affective Climate	17	.50	16	.55	16	.55
Learner Self-Concept	11	.33	11	.33	15	.37
Meaningfulness	7	.30	6	.33	11	.49
Reinforcement	15	.42	13	.47	13	.47
Close Supervision	12	.15	9	.29	8	.33
Awareness	6	.34	6	.34	11	.45
TOTAL	123	.69	111	.69	143	.64
Average		.32		.37		.45

Preliminary Analysis

Sample

Data were collected from four groups of subjects. These four groups included 1) university students not majoring in Education, 2) university students majoring in Education, 3) student teachers, and 4) experienced practicing teachers. The intention was to analyze the data within both elementary and secondary samples. However, the response within the secondary sample, particularly from student-teachers, was not sufficient to construct the four groups described above and thus perform the specified data analysis. As a result, this study utilized data exclusively from an elementary sample.

Demographic Composition of Elementary Sample

The demographic composition of each of the four sampled groups is summarized on Tables 6, 7, and 8. Table 6 provides background information regarding the age and gender of the subjects. Table 7 summarizes the educational background and undergraduate status of the subjects. Lastly, Table 8 outlines the training and experience in the field of education for each group of subjects. The information provided on these tables served as the basis for further examination of the homogeneity of each group.

Table 6

Age and Gender of Elementary Sample

	*NoED n=127	PreST n=52	ST n=47	Teachers n=36
	n (%)	n (%)	n (%)	n (%)
Age				
under 19	2 (1.6)	1 (1.9)	1 (2.1)	0 (0)
19-21	86 (67.7)	32 (61.5)	2 (4.3)	0 (0)
22-30	37 (29.1)	18 (34.6)	40 (85.1)	7 (19.4)
31-40	0 (0)	1 (1.9)	3 (6.4)	12 (33.3)
over 40	0 (0)	0 (0)	0 (0)	17 (47.2)
Missing	2 (1.6)	0 (0)	1 (2.1)	0 (0)
Gender				
Female	78 (61.4)	46 (88.5)	43 (91.5)	29 (80.5)
Male	47 (37.0)	5 (9.6)	3 (6.4)	5 (13.9)
Missing	2 (1.6)	1 (1.9)	1 (2.1)	2 (5.6)

* NoED - Non-education students
PreST - Education Students (pre- student teaching)
ST - Student teachers
Teachers - Experienced teachers

Table 7

Education and Status of Elementary Sample

	*NoED n=127	PreST n=52	ST n=47	Teachers n=36
	n (%)	n (%)	n (%)	n (%)
Education Level				
Working on Bachelors	123 (96.9)	49 (94.2)	28 (59.6)	0 (0)
Completed Bachelors	2 (1.6)	2 (3.8)	2 (4.3)	9 (25.0)
Working on Masters	0 (0)	0 (0)	15 (31.9)	10 (27.8)
Completed Masters	0 (0)	0 (0)	1 (2.1)	9 (25.0)
Masters plus	0 (0)	0 (0)	0 (0)	7 (19.4)
Missing	2 (1.6)	1 (1.9)	1 (2.1)	1 (2.8)
Undergraduate Status				
1st Year	0 (0)	0 (0)	0 (0)	**
2nd Year	17 (13.4)	6 (11.5)	0 (0)	
3rd Year	97 (76.4)	16 (30.8)	0 (0)	
4th Year	3 (2.4)	26 (50.0)	5 (10.6)	
5th Year	2 (1.6)	1 (1.9)	25 (53.2)	
Missing data	2 (1.6)	3 (5.8)	17 (36.2)	

* NoED - Non-education students
PreST - Education Students (pre- student teaching)
ST - Student teachers
Teachers - Experienced teachers

** Undergraduate Status does not apply in these cases.

Table 8

Training and Experience of Elementary Sample

*	NoED n=127	PreST n=52	ST n=47	Teachers n=36
	n (%)	n (%)	n (%)	n (%)
Number of Education Courses				
One or fewer	106 (83.5)	0 (0)	0 (0)	0 (0)
Two	21 (16.5)	2 (3.8)	0 (0)	0 (0)
Three	0 (0)	7 (13.5)	0 (0)	0 (0)
Four	0 (0)	2 (3.8)	1 (2.1)	0 (0)
Five or more	0 (0)	40 (76.9)	46 (97.9)	31 (86.1)
Missing data	0 (0)	1 (1.9)	0 (0)	5 (13.9)
Highest Level of Classroom Teaching Experience				
None	127 (100)	3 (5.8)	0 (0)	0 (0)
Observations	0 (0)	14 (26.9)	0 (0)	0 (0)
Mini-teaching (prior to ST)	0 (0)	35 (67.3)	0 (0)	0 (0)
Student Teaching	0 (0)	0 (0)	47 (100)	0 (0)
Practicing Teacher	0 (0)	0 (0)	0 (0)	36 (100)
Missing data	0 (0)	0 (0)	0 (0)	0 (0)

* NoED - Non-education students
PreST - Education Students (pre- student teaching)
ST - Student teachers
Teachers - Experienced teachers

Gender

As Table 6 shows, each of the four groups was comprised of over 80% females with the exception being the group of non-education students which was 61% female. Because the non-education group had a greater number of males relative to the other three groups, comparisons were made between males and females before any further analyses were considered. It was revealed that females consistently scored higher than males indicating a lack of homogeneity within the non-education group. Overall, females obtained a mean of 52.9, significantly higher than males, who obtained an overall mean of 45.0 ($F_{1,123} = 25.62$, $p < .001$). The competency profiles for males and females were also found to be significantly different ($F_{10,1230} = 2.03$, $p < .05$). The means for both males and females are presented in Table 8 while the group profiles are displayed on Figure 2. Given this finding, it was decided to use only females for all future comparisons.

Table 9

Non-Education Student Test Performance

	Males (n=47)		Females (n=78)	
	Mean	SD	Mean	SD
Competency				
Academic Learning Time	47.9	(9.4)	51.4	(10.2)
Accountability	50.5	(11.4)	49.5	(9.2)
Individual Differences	46.3	(10.4)	52.0	(9.3)
Evaluation	48.1	(11.2)	50.9	(9.1)
Consistent Rules	46.9	(11.6)	51.8	(8.6)
Affective Climate	46.5	(10.1)	52.1	(9.4)
Learner Self-Concept	48.2	(12.0)	51.0	(8.6)
Meaningfulness	46.1	(11.6)	52.4	(8.2)
Reinforcement	45.7	(10.9)	52.7	(8.1)
Close Supervision	49.2	(10.5)	50.7	(9.7)
Awareness	49.2	(10.9)	50.4	(9.6)
TOTAL	45.0	(11.5)	52.9	(7.7)

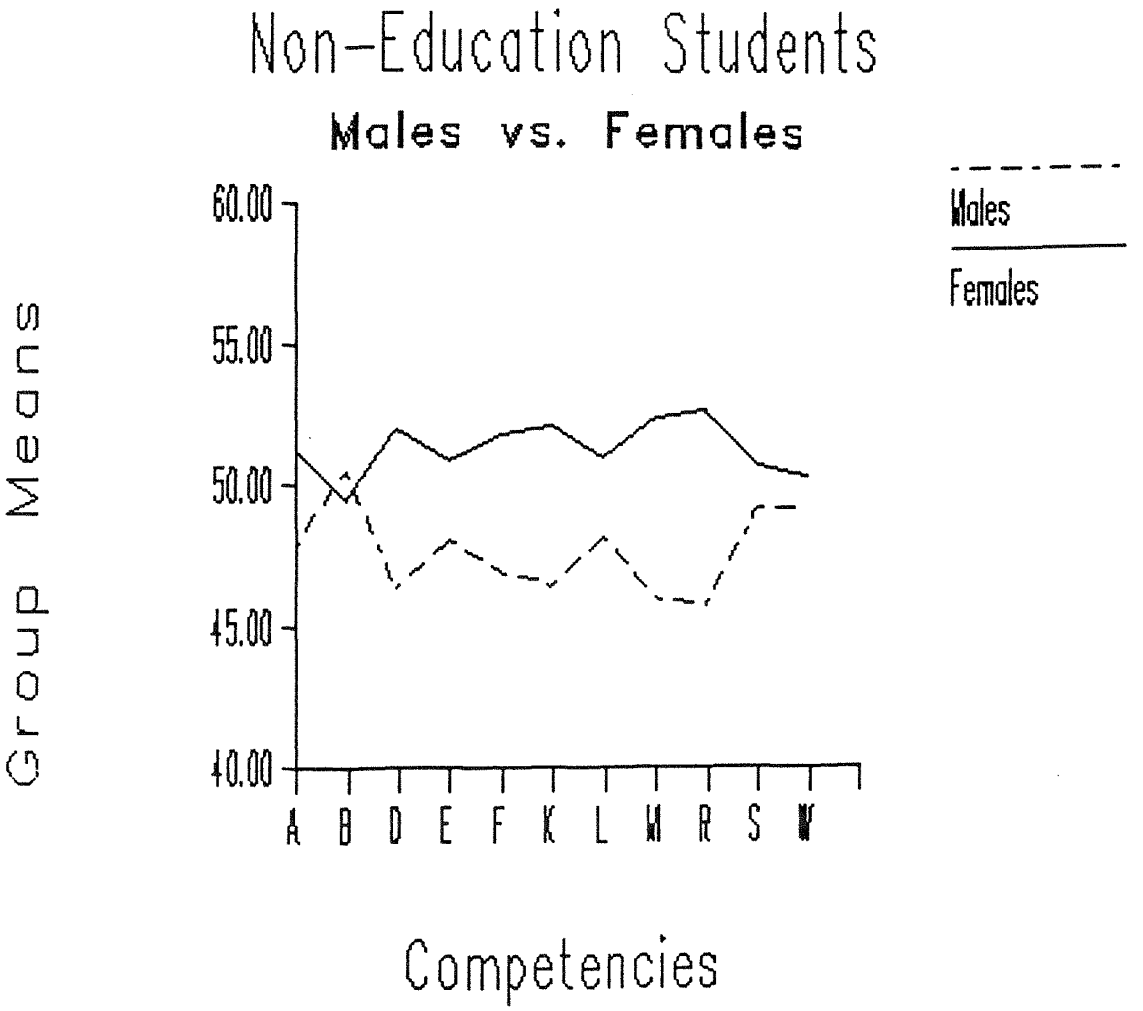

Figure 2: Group Profiles of Non-education Students

A	Academic Learning Time
B	Accountability
D	Individual Differences
E	Evaluation
F	Consistent Rules
K	Affective Climate
L	Learner Self-Concept
M	Meaningfulness
R	Reinforcement
S	Close Supervision
W	Awareness

Institutional Affiliation

The two groups of pre-service teachers were sampled from three different institutions. Within the group of female student teachers (N=43), 38 were from School A. This distribution did not allow for any comparisons regarding the variable of institutional affiliation. A total of 44 female students comprised the group of education students who had not yet student taught with 25 coming from School A, 10 from School B, and the remaining 9 from School C. A significant difference ($F_{2,41} = 5.25$, $p < .01$) in overall performance was found among these three institutions. The group profiles were also significantly different ($F_{20,410} = 2.92$, $p < .05$). Because of this lack of homogeneity, only those students receiving their training from School A were used in further analysis. Group means for the overall performance and each specific competency are presented on Table 10.

Teaching Experience

Of the 28 female elementary teachers included in this study, teaching experience ranged from 2 to 27 years. The mean number of years experience was 14.5 and the median was 15. The relationship between amount of teaching experience and overall test performance was examined with the calculation of a Pearson correlation. A correlation of .04 ($p = .851$) was obtained failing to reveal a significant linear relationship between amount of teaching experience and test performance.

Table 10

Institutional Affiliation and Performance of Pre-service Teachers

Competency	School A (n=25) Mean	SD	School B (n=10) Mean	SD	School C (n=9) Mean	SD
Academic Learning Time	52.5	(10.6)	47.8	(9.9)	45.5	(6.7)
Accountability	46.1	(9.3)	55.5	(9.5)	54.7	(8.0)
Individual Differences	53.1	(6.7)	53.1	(7.0)	38.1	(12.2)
Evaluation	48.4	(9.6)	55.4	(10.0)	48.5	(10.0)
Consistent Rules	48.7	(8.8)	53.2	(9.4)	50.2	(13.7)
Affective Climate	53.4	(6.6)	48.5	(11.4)	42.3	(12.6)
Learner Self-Concept	54.8	(7.0)	44.2	(12.4)	43.0	(6.7)
Meaningfulness	50.5	(9.9)	52.1	(9.5)	46.3	(11.0)
Reinforcement	52.9	(8.0)	46.7	(10.5)	45.5	(12.6)
Close Supervision	51.1	(7.1)	48.8	(9.3)	48.4	(16.7)
Awareness	46.1	(10.1)	55.8	(7.7)	54.5	(7.5)
TOTAL	50.3	(9.2)	54.6	(11.2)	44.0	(8.7)

Final Demographic Composition

The demographic characteristics of the final sample of subjects used in this study are presented on Tables 11 and 12. Table 11 identifies the educational level and undergraduate status of all subjects. It can be seen that the majority of the subjects in this study have yet to receive their undergraduate degree with the exception of the group of practicing teachers in which nearly half have received a masters degree. The range of undergraduate status within the non-education group is more restricted than anticipated with over 95% being third- and fourth-year students. This parallels that of the education students (PreST), thus no comparisons were made among levels of status within the non-education group. The student teachers were comprised of primarily fifth-year students or masters students having no undergraduate status.

Table 12 summarizes the educational training and experience of the four groups of subjects. The distributions presented on Table 12 correspond with the guidelines which were used in constructing each of these groups. These groups were constructed so that there would be little or no overlapping of amount of educational training and highest level of experience across the four groups. Table 12 shows that the non-education students have no classroom teaching experience and have taken very few courses (2 or less) in Education. The remaining three

groups of subjects have completed or nearly completed their coursework in Education and have gained additional experiences in the classroom serving as classroom observers, student teachers, and practicing teachers.

Table 11

Education and Status of Final Sample

	NoED n=78	PreST n=25	ST n=38	Teachers n=28
	n (%)	n (%)	n (%)	n (%)
Education Level				
Working on Bachelors	76 (97.4)	23 (92.0)	37 (97.3)	0 (0)
Completed Bachelors	2 (2.6)	2 (8.0)	0 (0)	9 (32.1)
Working on Masters	0 (0)	0 (0)	1 (2.7)	6 (21.4)
Completed Masters	0 (0)	0 (0)	1 (2.3)	7 (25.0)
Masters plus	0 (0)	0 (0)	0 (0)	6 (21.4)
Missing	0 (0)	0 (0)	0 (0)	0 (0)
Undergraduate Status				
1st Year	0 (0)	0 (0)	0 (0)	*
2nd Year	2 (2.6)	0 (0)	0 (0)	
3rd Year	11 (14.1)	0 (0)	0 (0)	
4th Year	64 (82.1)	23 (92.0)	1 (2.3)	
5th Year	1 (1.3)	0 (0)	37 (97.3)	
Missing data	0 (0)	2 (8.0)	0 (0)	

* NoED - Non-education students
PreST - Education Students (pre- student teaching)
ST - Student teachers
Teachers - Experienced teachers

** Undergraduate Status does not apply in these cases.

Table 12

Training and Experience of Final Sample

	*NoED n=78	PreST n=25	ST n=38	Teachers n=28
	n (%)	n (%)	n (%)	n (%)
Number of Education Courses				
One or fewer	67 (85.9)	0 (0)	0 (0)	0 (0)
Two	11 (14.1)	0 (0)	0 (0)	0 (0)
Three	0 (0)	0 (0)	0 (0)	0 (0)
Four	0 (0)	0 (0)	1 (2.7)	0 (0)
Five or more	0 (0)	25 (100)	37 (97.3)	23 (82.1)
Missing data	0 (0)	0 (0)	0 (0)	5 (17.9)
Highest Level of Classroom Teaching Experience				
None	78 (100)	0 (0)	0 (0)	0 (0)
Observations	0 (0)	1 (4.0)	0 (0)	0 (0)
Mini-teaching (prior to ST)	0 (0)	24 (96.0)	0 (0)	0 (0)
Student Teaching	0 (0)	0 (0)	38 (100)	0 (0)
Practicing Teacher	0 (0)	0 (0)	0 (0)	28 (100)
Missing data	0 (0)	0 (0)	0 (0)	0 (0)

* NoED - Non-education students
PreST - Education Students (pre- student teaching)
ST - Student teachers
Teachers - Experienced teachers

Principal Analysis

Research Questions

Evidence of the construct validity of the Simulation Exercise in Classroom Decision-Making was to be provided in the affirmative answers to each of the three research questions included in this study. The three questions posed in this study were:

1. Do scores on the Simulation Exercise in Classroom Decision-Making reflect differences in the ability to apply teacher effectiveness findings to teaching problems between subjects with professional training and those without such training?
2. Do scores on the Simulation Exercise in Classroom Decision-Making reflect differences in the ability to apply teacher effectiveness findings to teaching problems between subjects with classroom teaching experience and those without experience?
3. Do scores on the Simulation Exercise in Classroom Decision-Making reflect differences in the ability to apply teacher effectiveness findings to teaching problems between experienced classroom teachers and student-teachers?

Each of these three questions was answered of both

overall group performance and patterns across the eleven competencies included in the Simulation Exercise in Classroom Decision-Making. Answers to these questions were obtained by testing six null hypotheses, two for each question: one regarding the total test score over 11 competencies and one regarding the profiles of the 11 scores.

A two factor design with one repeated measure was used to address the three research questions of the current study (Winer, 1971). The results from this analysis are summarized on Table 13 and the means and standard deviations for each of the four groups are shown in Table 14. This design yielded six F-tests. It was found that five of these six F-tests were statistically significant ($p < .05$), two pertaining to overall group performance and three pertaining to specific group profiles. These results indicated an overall test performance difference between those trained in education and those not trained in education and an overall difference between those subjects with classroom teaching experience and those without classroom experience. This analysis failed to reveal an overall difference between experienced teachers and student-teachers. Differences between the group profiles across the eleven competencies were revealed for each of the three specified group contrasts. These profiles will be examined in more detail in the concluding section.

Each of the three group contrasts was analyzed using the two-factor design with one repeated measure outlined previously. Means and standard deviations of the 11 competencies as well as total test performance are reported as they pertain to each of these three contrasts in Tables 15, 16, and 17. The profiles across the 11 competencies were plotted for each of these three group contrasts are displayed in Figures 3, 4, and 5.

Table 13

Results from Two-Factor Design with One-Repeated Measure

Source	df	MS	F
Between Subjects	168		
Groups (A)	3		
Contrast 1 (training)	1	826.42	6.55*
Contrast 2 (classroom exp)	1	825.35	6.54*
Contrast 3 (teaching exp)	1	1.15	.01
Sub w/groups (S/A)	165	126.15	
Within Subjects			
Comp X Group Interaction (A X B)	30	639.13	
Contrast 1 (training)	10	553.43	6.38*
Contrast 2 (classroom exp)	10	899.31	10.37*
Contrast 3 (teaching exp)	10	464.65	5.36*
Comp X Sub/groups (B X S/A)	1650	86.76	
Fitting 11 competency means (Mean = 50)	11		
Total	1859		

* $p < .01$

Table 14

Group Means and Standard Deviations

	NoED (n=78)	PreST (n=25)	ST (n=38)	Teachers (n=28)
	Mean(SD)	Mean(SD)	Mean(SD)	Mean(SD)
Competency				
Academic Learning Time	46.5(10.0)	55.0(8.3)	54.8(7.4)	48.8(10.3)
Accountability	53.3(8.4)	47.3(8.9)	42.9(7.4)	47.1(10.4)
Individual Differences	46.7(10.3)	48.6(8.3)	54.2(9.0)	54.8(7.9)
Evaluation	48.4(9.8)	48.3(9.0)	49.0(9.7)	57.4(9.0)
Consistent Rules	50.2(10.2)	49.2(9.9)	48.1(9.9)	52.7(9.6)
Affective Climate	49.4(11.8)	52.6(4.1)	52.8(4.6)	45.7(12.0)
Learner Self-Concept	49.7(7.5)	53.6(7.1)	51.7(11.6)	45.3(13.8)
Meaningfulness	48.9(9.7)	51.3(10.0)	52.8(9.9)	48.1(10.5)
Reinforcement	47.5(10.9)	54.5(8.7)	49.6(9.6)	53.4(6.5)
Close Supervision	44.4(10.7)	53.7(5.7)	55.5(6.9)	54.8(5.9)
Awareness	51.8(10.0)	46.6(11.2)	47.4(9.3)	51.4(8.5)
TOTAL	48.7(10.3)	52.0(9.5)	50.3(10.5)	51.3(8.9)

Educational Training

Table 15 provides a summary of the results used to address the first research question:

1. Do scores on the Simulation Exercise in Classroom Decision-Making reflect differences in the ability to apply teacher effectiveness findings to teaching problems between subjects with professional training and those without such training?

The group of students who had received training in Education outperformed the non-education group on seven of the eleven competencies as well as the overall test. Both groups scored the same in the competency of Evaluation, while the non-education group outperformed the education group on the remaining three competencies. A series of independent t-tests revealed five significant differences within these group profiles. Those students trained in Education scored significantly higher in the areas of Academic Learning Time, Learner Self-Concept, Reinforcement, and Close Supervision whereas those students who did not receive this training scored significantly higher in the competency area of Accountability. The profiles for these two groups are plotted in Figure 3.

Table 15

Educational Training Contrast

Group Means and Standard Deviations

	NoED (n=78) Mean(SD)	PreST (n=25) Mean(SD)	Diff[a]	t value (df=101)
Competency				
Academic Learning Time	46.5(10.0)	55.0(8.3)	8.5	3.85***
Accountability	53.3(8.4)	47.3(8.9)	-6.0	-4.13***
Individual Differences	46.7(10.3)	48.6(8.3)	1.9	.85
Evaluation	48.3(9.8)	48.3(9.0)	0.0	.05
Consistent Rules	50.2(10.2)	49.2(9.9)	-1.0	-.41
Affective Climate	49.4(11.8)	52.6(4.1)	3.2	1.35
Learner Self-Concept	49.7(7.5)	53.6(7.1)	3.9	2.28*
Meaningfulness	48.9(9.7)	51.3(10.0)	2.4	1.04
Reinforcement	47.5(10.9)	54.5(8.7)	7.0	2.93**
Close Supervision	44.4(10.7)	53.7(5.7)	9.3	4.16***
Awareness	51.8(10.0)	46.6(11.2)	-5.2	-2.18*
TOTAL	48.7(10.3)	52.0(9.5)	1.5	

a Difference between means (PreST - NoED)
* $p < .05$
** $p < .01$
*** $p < .001$

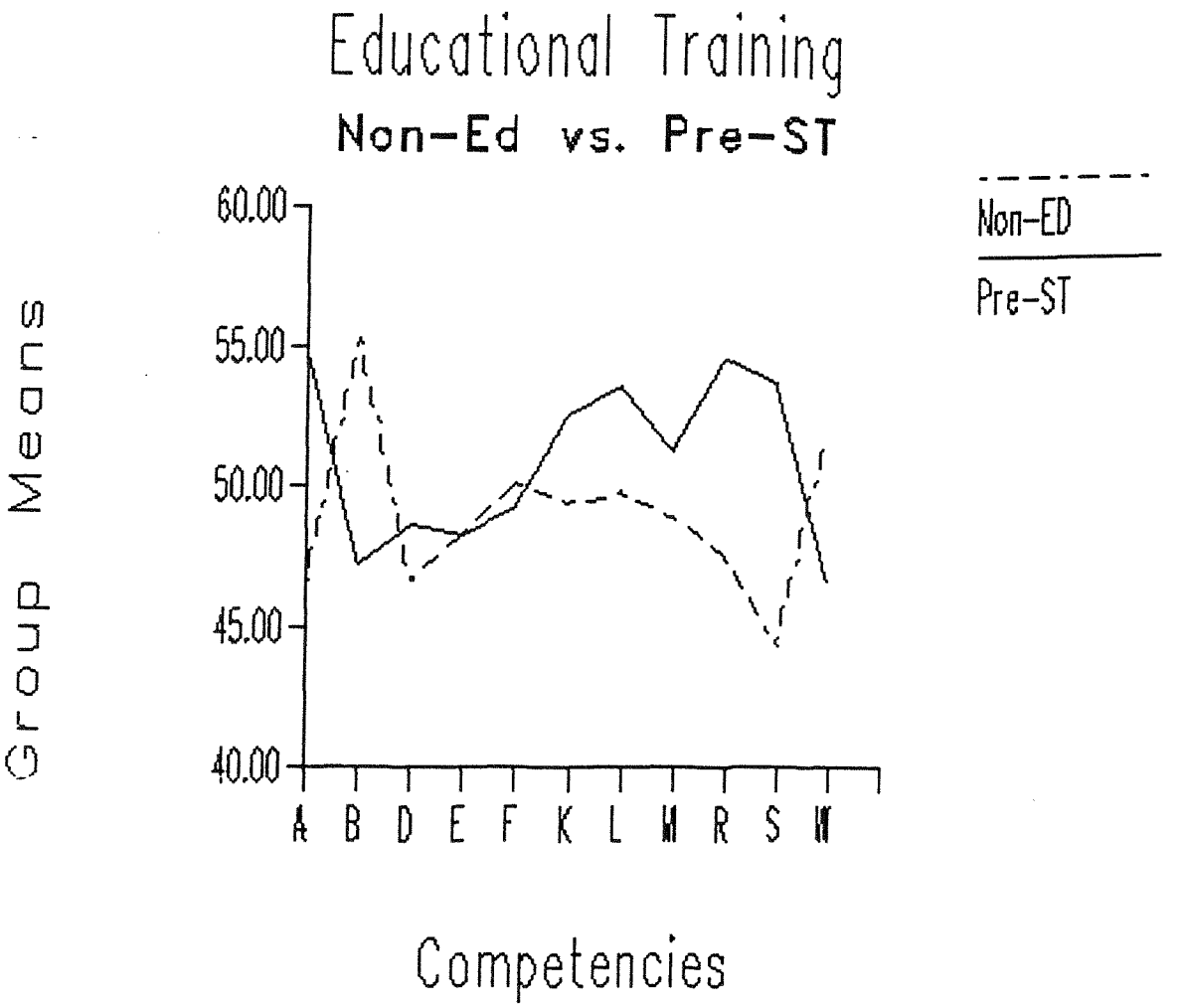

Figure 3: Educational Training Group Profiles
Non-education students vs. Education Students

A Academic Learning Time
B Accountability
D Individual Differences
E Evaluation
F Consistent Rules
K Affective Climate
L Learner Self-Concept
M Meaningfulness
R Reinforcement
S Close Supervision
W Awareness

Classroom Experience

The results from the second group contrast are presented in Table 16. The second research question was concerned with the factor of classroom teaching experience. This question was stated specifically as:

2. Do scores on the Simulation Exercise in Classroom Decision-Making reflect differences in the ability to apply teacher effectiveness findings to teaching problems between subjects with classroom teaching experience and those without experience?

The results from Table 16 indicate that those subjects who have had classroom teaching experience (n=66), either during student teaching or as a practicing teacher, outscored those subjects without this classroom experience (n=103) on the overall test as well as seven specific competency areas. Four of these seven competency areas were significantly higher than that of the inexperienced subjects. These included Academic Learning Time, Individual Differences, Evaluation, and Close Supervision. The subjects without classroom experience scored higher in the remaining four competency areas, significantly higher on the competency of Accountability. These group profiles are displayed in Figure 4.

Table 16

Classroom Experience Contrast

Group Means and Standard Deviations

	Inexp (n=103) Mean(SD)	Exp (n=66) Mean(SD)	Diff[a]	t value (df=167)
Competency				
Academic Learning Time	48.5(10.3)	52.3(9.2)	3.8	2.42**
Accountability	53.4(9.2)	44.7(8.9)	-8.7	-6.05***
Individual Differences	47.2(9.9)	54.4(8.5)	7.2	4.94***
Evaluation	48.4(9.5)	52.6(10.2)	4.2	2.73**
Consistent Rules	50.0(10.5)	50.1(9.9)	0.1	.06
Affective Climate	50.1(10.5)	49.8(9.2)	-0.3	- .24
Learner Self-Concept	50.7(7.6)	49.0(12.9)	-1.7	-1.08
Meaningfulness	49.5(9.8)	50.8(10.3)	1.3	.80
Reinforcement	49.2(10.8)	51.2(8.6)	2.0	1.27
Close Supervision	46.7(10.5)	55.2(6.4)	8.5	5.95***
Awareness	50.6(10.5)	49.1(9.1)	-1.5	- .89
TOTAL	49.5(10.2)	50.7(9.8)	1.2	

a Difference between means (Exp - Inexp)
* $p < .05$
** $p < .01$
*** $p < .001$

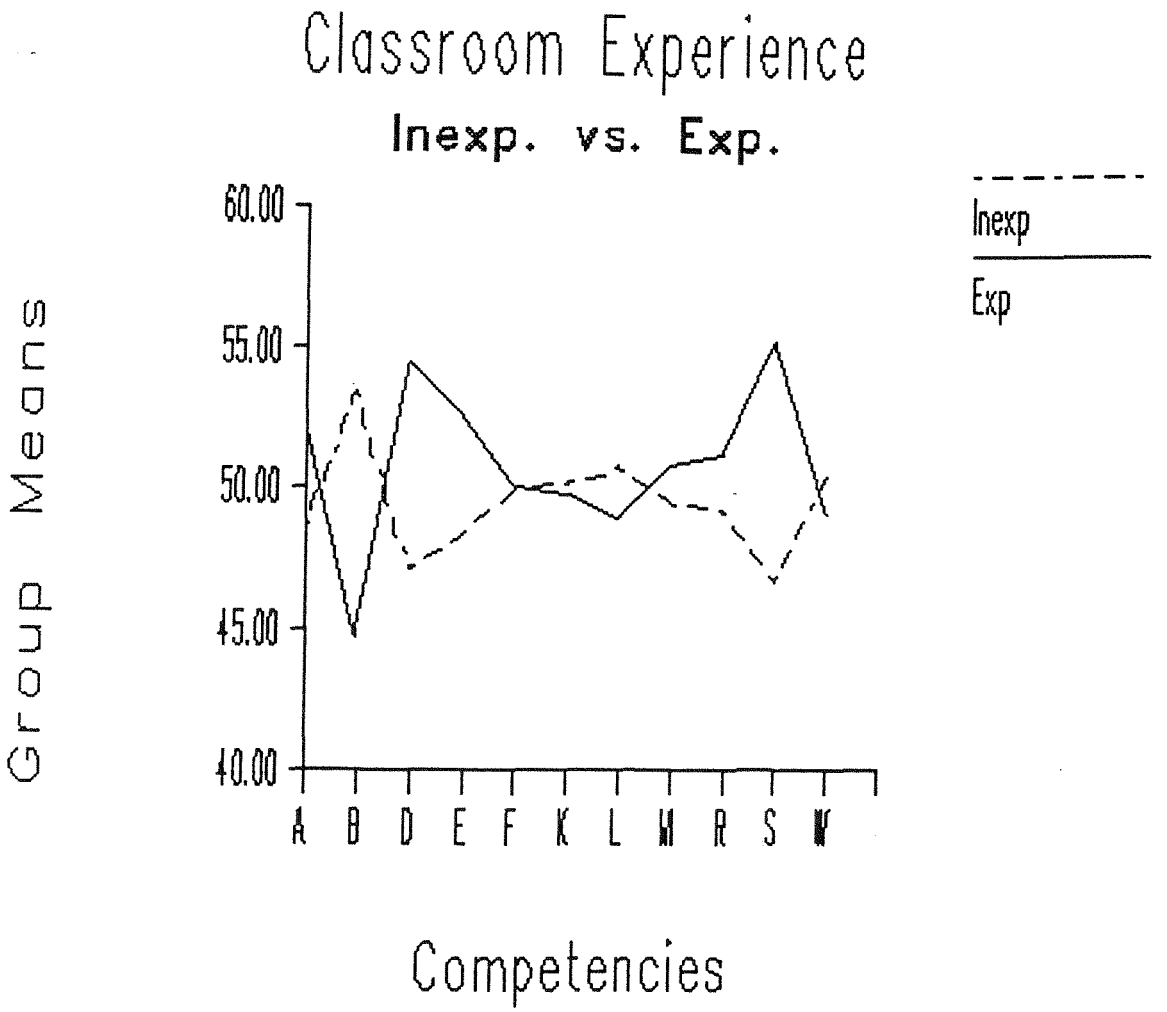

Figure 4: Classroom Experience Group Profiles
Inexperienced vs. Experienced

A	Academic Learning Time
B	Accountability
D	Individual Differences
E	Evaluation
F	Consistent Rules
K	Affective Climate
L	Learner Self-Concept
M	Meaningfulness
R	Reinforcement
S	Close Supervision
W	Awareness

Teaching Experience

The third and final research question regarded the amount of teaching experience and its correspondence to the performance on this simulation in classroom decision-making. The third research question was:

3. Do scores on the Simulation Exercise in Classroom Decision-Making reflect differences in the ability to apply teacher effectiveness findings to teaching problems between experienced classroom teachers and student-teachers?

Table 17 summarizes the information gathered relevant to this research question. An overall difference between teachers with an average of 14.5 years experience and student teachers was not found. However, a difference in the two group profiles across the eleven specific competencies was found. These profiles are plotted in Figure 5. Student teachers scored higher on five competencies, significantly higher in the areas of Academic Learning Time, Affective Climate, and Learner Self-Concept. Teachers scored higher on six of the eleven competencies, significantly higher in the area of Evaluation.

Table 17

Teaching Experience Contrast

Group Means and Standard Deviations

	ST (n=38) Mean(SD)	Teachers (n=28) Mean(SD)	Diff[a]	t value (df=64)
Competency				
Academic Learning Time	54.8(7.4)	48.8(10.3)	-6.0	-2.74**
Accountability	42.9(7.4)	47.1(10.4)	4.2	1.94
Individual Differences	54.2(9.0)	54.8(7.9)	0.6	0.29
Evaluation	49.0(9.7)	57.4(9.0)	8.4	3.55***
Consistent Rules	48.1(9.9)	52.7(9.6)	4.6	1.90
Affective Climate	52.8(4.6)	45.7(12.0)	-7.1	-3.34**
Learner Self-Concept	51.7(11.6)	45.3(13.8)	-6.4	-2.03*
Meaningfulness	52.8(9.9)	48.1(10.5)	-4.7	-1.85
Reinforcement	49.6(9.6)	53.4(6.5)	3.8	1.84
Close Supervision	55.5(6.9)	54.8(5.9)	-0.7	-0.39
Awareness	47.4(9.3)	51.4(8.5)	4.0	1.79
TOTAL	50.3(10.5)	51.3(8.9)	1.0	

a Difference between means (Teachers - ST)
* $p < .05$
** $p < .01$
*** $p < .001$

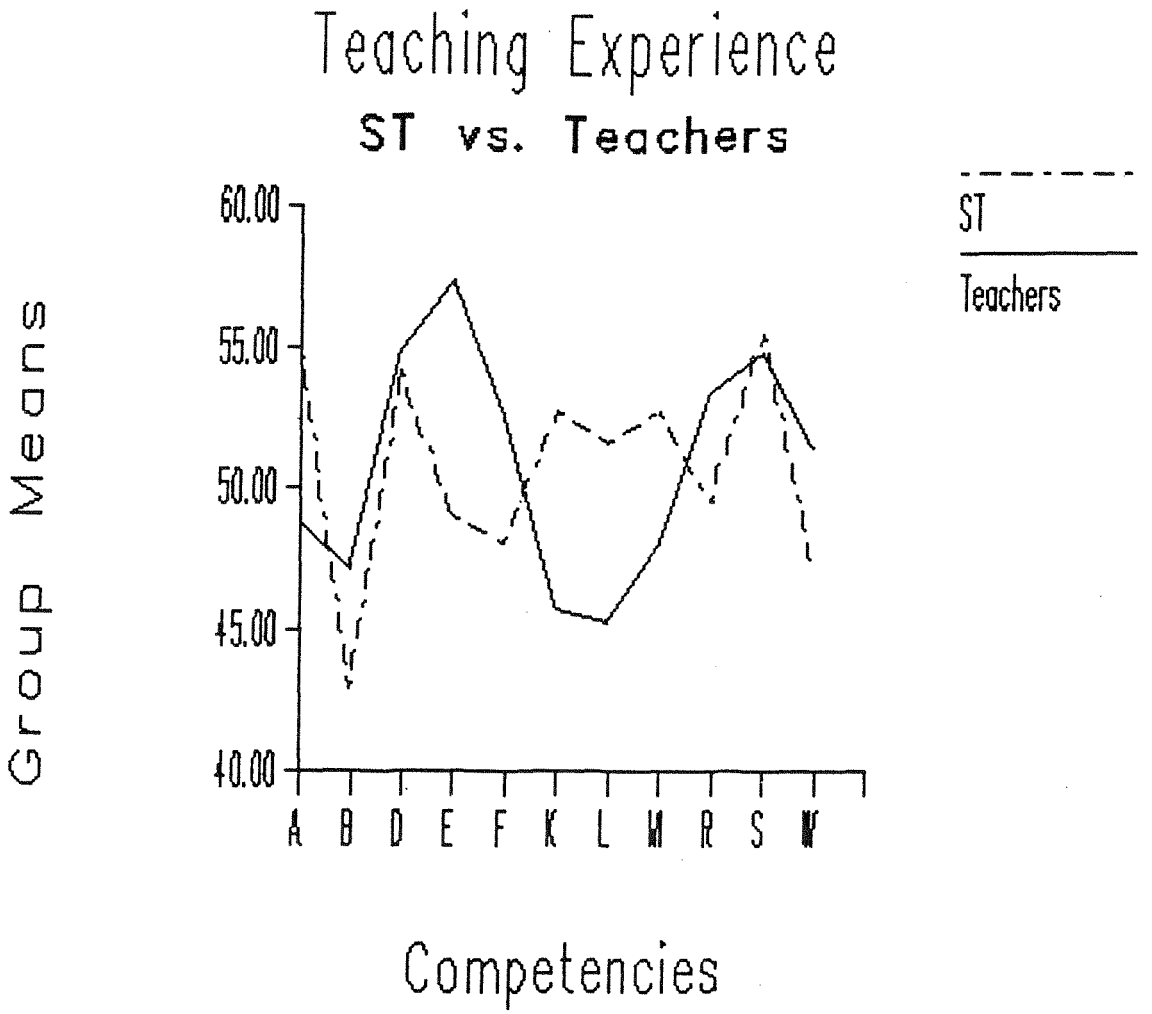

Figure 5: Teaching Experience Group Profiles
Student Teachers vs. Practicing Teachers

A	Academic Learning Time
B	Accountability
D	Individual Differences
E	Evaluation
F	Consistent Rules
K	Affective Climate
L	Learner Self-Concept
M	Meaningfulness
R	Reinforcement
S	Close Supervision
W	Awareness

CHAPTER 5

Summary and Conclusions

This study aimed to verify and extend the findings obtained during the field test of the STITC-H test. This test was constructed as a group test of interactive teaching competencies in an effort to simulate the decision-making process in the classroom more realistically than traditionally used pencil-and-paper tests.

Test Refinement

To fulfill this purpose, the STITC-H test used in the initial pilot study was refined for use in the current study. This refinement process required a series of discriminant and item analyses to examine the internal consistency of each of the thirteen competencies comprising the test. One competency, Questioning Skills, was eliminated during this process. As a result of this process, the test was reduced from 65 to 42 problems, consisting of 188 items and was able to be completed within 45 minutes. The format of the test was also changed from the original synchronized slide/tape presentation to a single videotape.

Sample

The revised test was administered to four groups of subjects. These four groups included 1) university students not majoring in Education, 2) university students majoring in Education, 3) student teachers, and 4) experienced

practicing teachers. The intention was to analyze the data within both elementary and secondary samples. However, the response within the secondary sample, particularly from student-teachers, was not sufficient to construct the four groups described above and thus perform the specified analysis of the data. As a result, this study utilized data exclusively from an elementary sample.

Reliability

Reliability estimates for the internal consistency of each of the twelve competencies were calculated using Cronbach's coefficient alpha. Items written specifically for this study were assigned to a specific competency one at a time and their contribution to that competency was assessed. The items from each competency were also evaluated and reviewed for their consistency with the BTAP indicators (see Appendix A). This content validity examination resulted in the elimination of one competency area, Clarity of Structure. Final alpha coefficients ranged from .33 (Close Supervision) to .60 (Accountability) with a mean of .45.

Preliminary Analysis

The homogeneity of each of the four main groups comprising this study's sample was examined using a two factor design with one repeated measure (Winer, 1971). This design examined overall group performance as well as patterns over the eleven competencies.

It was found at this time that three of the four groups of subjects were comprised of over 80% females and the fourth group, non-education students, was 61% female. A comparison within the non-education group revealed that females scored consistently higher than males. Thus, it was decided to conduct the remainder of the study using only female subjects.

One other subgroup was found to lack homogeneity. The group of education students who had not yet student taught came from three different institutions. Significant differences in both overall performance and profiles across the eleven competencies were found within this subgroup. Due to this finding and the dominance of student teachers from School A, only those subjects who were trained at School A were used in further analysis.

Principal Analysis

The major emphasis in this study was to examine the construct validity of the Simulation Exercise in Classroom Decision-Making. In establishing construct validity, it was theorized that this test was a measure of generic knowledge about interactive teaching competencies. It was proposed that this generic knowledge may be learned through university teacher-training programs, the reading of the research literature on effective teaching, and experiences as classroom teachers. Therefore, it was hypothesized that: 1) university students with training in education would

score higher than students without training in education, 2) those subjects with classroom teaching experience (student teachers and teachers) would score higher than those without classroom teaching experience (non-education and education students) and 3) experienced teachers would score higher than student teachers. Regarding the overall test performance of subjects, two of the three hypotheses were rejected at the .05 level. A significant difference between the overall performance of experienced teachers and student teachers was not found. All three hypotheses were rejected upon the examination of specific group profiles across the eleven competencies. These specific hypotheses will be discussed individually in the following sections. The discussion of these three group profile differences will be outlined using specific behaviors from the simulation exercise.

Educational Training

The first research question was proposed to examine the extent to which the Simulation Exercise in Classroom Decision-Making reflected differences between students who have had training in education and those who have not had training. This question was stated specifically as:

1. Do scores on the Simulation Exercise in Classroom Decision-Making reflect differences in the ability to apply teacher effectiveness findings to teaching problems between subjects

with professional training and those without such training?

An affirmative answer to this question indicated that the students who were trained in education acquired knowledge through their pedagogical coursework that other university students did not. University students trained in education scored significantly higher than those students without such training on the overall test. The education students obtained a mean performance of 52.0 while the non-education students averaged a 48.7.

Students with teacher training scored higher than the students without teacher training on seven of the eleven competencies. Both groups scored the same on the competency of Evaluation. The students trained in education scored significantly higher in the areas of Academic Learning Time, Learner Self-Concept, Reinforcement, and Close Supervision while those students without educational training scored significantly higher on Accountability.

These results suggest that teacher training is especially helpful in providing students with knowledge pertaining to four areas while detrimental in the area of Accountability. Students trained in education identified alternatives that made more efficient use of classroom time, and maintained the focus of the lesson and the attention of the students. These education students also more often responded to situations with a sensitivity to the student's

racial and cultural background, encouraged learners to try harder and do better in class, and avoided responses which would embarrass or make the student feel inferior.

Students trained in education indicated the use of positive reinforcement strategies more often as they pertained to classroom rules, students' abilities in class and during recess, and avoided the use of punitive responses such as slapping and giving failing grades. In addition, students trained in education more often indicated they would circulate throughout the classroom while presenting a lesson or monitoring seatwork.

Those students who were not trained in education scored significantly higher in the competency of Accountability. These students more often indicated that they would hold students responsible for assigned tasks, to be completed on the assigned date, and within the stated guidelines. In contrast to the non-education students, students trained in education were more likely to offer extended deadlines, and ignore certain criteria in students' written work such as neatness and spelling. Non-education students more often indicated that they would hold students accountable for having all their materials such as textbooks and workbooks ready for class. These students were also more likely to have students were frequently unprepared make up work after class.

Classroom Experience

The second research question was proposed to test the hypothesis that those subjects with classroom teaching experience would perform better than those without classroom experience on a test that simulates interactive decision-making. This question was:

2. Do score on the Simulation Exercise in Classroom Decision-Making reflect differences in the ability to apply teacher effectiveness findings to teaching problems between subjects with classroom teaching experience and those without experience?

An affirmative answer to this question provided evidence that the experience gained from teaching in the classroom results in the acquisition of knowledge that contributes to performance on the Simulation Exercise in Classroom Decision-Making. Those subjects who had experience in classroom teaching, either as student teachers or practicing teachers, scored significantly higher than those subjects without classroom teaching experience on the overall test score. Those experienced in classroom teaching obtained a mean of 50.7 whereas the inexperienced group acquired a mean of 49.5.

Specifically, subjects with classroom teaching experience scored higher on seven of the eleven competencies included in the Simulation. Significantly higher

performances were found in the competencies of Academic Learning Time, Individual Differences, Evaluation, and Close Supervision. The group of inexperienced subjects scored higher on the remaining four competencies, significantly higher in the area of Accountability.

These findings suggest that the experience gained while teaching in the classroom contributes to a teacher's knowledge of the research findings and ability to apply them in a simulation exercise. Experience gained in the classroom appears to be especially helpful in four competency areas.

First, subjects with classroom experience responded to classroom situations more consistently with research findings when asked to deal with situations which required them to maintain the focus of the lesson and student involvement in the lesson while dealing with specific individuals who were inattentive. They also made more efficient use of the time spent on procedural matters such as taking attendance and choosing teams for classroom drills. Secondly, those who have gained classroom teaching experience paid more attention to individual differences in providing alternative assignments for those learners with different abilities and background experiences, reviewing important concepts for the class when many students were having difficulty, and spending more time with individual learners having problems with their lessons. A third area

in which classroom experience has been a positive influence is that of Evaluation. Those subjects having classroom experience more often indicated they would check students' understanding during the instructional process by asking questions and requesting students to repeat important points. They also indicated greater use of informal quizzes to gather diagnostic information. Lastly, those with classroom teaching experience were more likely to supervise students closely. They monitored students' activity while engaged in small group tasks, circulated around the room while in the instructional mode and during seatwork, and were less likely to leave the room or correct papers at their desk.

Subjects without classroom experience were significantly higher in the competency of Accountability just as those students without the educational training were relative to education students. These inexperienced subjects more often held students responsible for assigned tasks, the completion of assignments by specific dates, and imposed consequences such as staying after school and reduced grades when assignments were not completed on time or prepared neatly without spelling errors. They also indicated more often that students should have all materials such as textbooks and workbooks ready for class and would need to make up after class if they were continuously unprepared.

Teaching Experience

The third research question was proposed to test the influence of the amount of teaching experience on performance in the Simulation Exercise in Classroom Decision-Making. This question was specifically aimed at testing the hypothesis that experienced teachers would score higher than student teachers and was stated as:

3. Do scores on the Simulation Exercise in Classroom Decision-Making reflect differences in the ability to apply teacher effectiveness findings to teaching problems between experienced classroom teachers and student-teachers?

A significant difference in the overall performance of experienced teachers and student teachers was not found in the current study. However, experienced teachers did achieve a mean of 51.3, which was slightly higher than that of student teachers (50.3). The profiles of the two group's performance across each of the eleven competencies differed significantly.

Experienced teachers scored higher on six of the eleven competencies while student teachers scored higher on five. Experienced teachers scored significantly higher in the area of Evaluation as student teachers scored significantly higher in the areas of Academic Learning Time, Affective Climate and Learner Self-Concept.

Evaluation was identified while investigating the second research question as a competency influenced by classroom teaching experience. Experienced teachers more often indicated checking students' understanding during instruction by asking questions and requesting them to repeat important points, and reported greater use of quizzes for formative evaluation purposes.

Student teachers scored significantly higher than experienced teachers in three competency areas. These areas included 1) Academic Learning Time, which was earlier associated with both educational training and classroom experience, and 2) Learner Self-Concept, also associated with training, and 3) Affective Climate. Student teachers more often responded to situations with alternatives suggesting they maintain the focus of the lesson and student involvement in the lesson while dealing with specific individuals who were inattentive. They also indicated making more efficient use of the time spent on procedural matters such as taking attendance and choosing teams for classroom drills.

Student teachers, like those students who were trained in education, more often responded to situations with a sensitivity to the student's racial and cultural background, encouraged learners to try harder and do better in class, and avoided responses which would embarrass or make the student feel inferior.

Student teachers also responded to simulated classroom situations in a more accepting manner. For example, student teachers less often reported dealing with classroom situation in punitive or threatening ways such as throwing student materials, reprimanding students in front of other students, and putting students "on the spot" or embarrassing them in front of the class.

Discussion

Some consistent patterns among the eleven competencies of interactive teaching were found in the current study. Both educational training and classroom teaching experience were found to be associated with several areas of teacher competence. The knowledge gained from pedagogical training and teaching in the classroom has been identified as especially helpful in seven of the eleven competencies included in this study. The scores in the areas of Academic Learning Time, Individual Differences, Evaluation, Affective Climate, Learner Self-Concept, Meaningfulness, and Close Supervision steadily increased as the groups included in this study acquired more training and experience. The exception to this linear increase in knowledge occurs as teachers gain additional experience in the classroom. Increased experience as practicing teachers only proved to be valuable in one area, Evaluation.

The development of knowledge in these seven competency areas can be mostly attributed to the content of

professional education coursework and the reading of professional journals because the two groups of pre-service teachers were consistently higher in these areas. These two groups, education students and student teachers, are those who have had the most recent exposure to professional courses and have been required to read the professional literature as a result. These students have all been trained within the same teacher preparation program at School A.

The examination of School A's curriculum outlines and course syllabi revealed that students complete coursework in the areas of learning theory, child development, exceptional children, language and reasoning skills instruction, educational psychology, and tests and measurement. The elementary curriculum emphasizes the teaching-learning process focusing on the major functions that teachers perform. For example, classroom organization and management is one area emphasized in the elementary preparation program. Coursework is provided that focuses on maintaining students' interest in the lesson and dealing with other disruptive behaviors while retaining the focus of the instructional lesson. This may be done by circulating around the classroom and monitoring student behavior, not taking critical time away from other students to deal with minor individual disruptions.

Another area of emphasis is that of instruction. It

is within this teaching function that knowledge of specific content and effective instructional strategies are most important. These elementary pre-service teachers have been trained in learning theory, child development, educational psychology and the instruction of both language and reasoning skills. Students gain valuable knowledge pertaining to individual learning styles and developmental stages of children as well as appropriate strategies in dealing with children at various stages. This knowledge has been reflected in their performances in the areas of Affective Climate, Learner Self-Concept, and Meaningfulness. These two groups of pre-service teachers have consistently scored higher than both the non-education students and the experienced teachers in these three competency areas.

The competency areas of Individual Differences and Evaluation were also identified as benefits of training in education. The performance within these two areas increased as groups received more training and more experience. The performance of individuals who had classroom teaching experience was significantly higher than those who had no experience. This finding suggests that dealing with individual differences and use of effective evaluative strategies are competencies best learned from experience working with children in the classroom.

One area in which training and experience was reflected negatively in this simulation was that of

Accountability. Those subjects without training or classroom experience performed the highest on this competency. The experiences gained through an elementary teacher preparation and working with children do not promote holding every student accountable for the completion of assignments within stated guidelines, by specified deadlines, and the imposition of consequences when not completed within these rigid criteria. Perhaps, through the development of an awareness of individual differences and a sensitivity for the child in regards to his or her personal self-concept and the overall affective climate of the classroom, teachers become less rigid and conventional with children.

Summary

The Simulation Exercise in Classroom Decision-Making has been developed as a pencil-and-paper measure of interactive teaching competencies that is inexpensive, quick, and easy to administer to groups of subjects. The mode of presentation has been modified so that the immediacy of decision-making within the actual classroom could be more closely simulated than in traditional pencil-and-paper tests of teacher competence.

Evidence of content and construct validity has been provided in both the pilot test and the current study. The test has been constructed in accordance with reviews of process-product research and yields estimates of consistency

with this research base. This simulation exercise has discriminated among subjects with varying degrees of training and experience in education.

Whether or not this test has predictive validity remains to be seen and that is a major recommendation for further study. If this simulation exercise can be used to predict teacher performance, it could then be used as a screening device within both teacher education programs and in local school divisions. This exercise could also be used to identify strengths and weaknesses of practicing teachers so that the appropriate inservice training could be scheduled and other remedial services provided.

Suggestions for Further Study

Based on the results obtained in this study some recommendations will be made for the refinement of the current study and the development of further research studies. The order in which they appear is not intended to suggest their degree of importance or the priority with which they should be addressed.

1. A future study should be conducted to investigate the extent to which performance on the Simulation Exercise in Classroom Decision-Making is related to teaching behavior in the classroom. The two groups of pre-service teachers included in the current study would serve as the basis for this future predictive

validity study. Many of these students will become teachers in the state of Virginia and will participate in the BTAP Program. An initial estimate of the current instrument's predictive validity could be obtained by correlating current performance with future BTAP observations on specific competencies.

2. The study described above could also be done concurrently. In order to conduct a study of this nature, beginning teachers would be sampled from the state of Virginia who are currently being observed within the BTAP system. This simulation would be administered concurrently and the two performances would be correlated.
3. The scores from this simulation exercise should also be correlated with other tests purporting to measure professional knowledge such as the NTE to further investigate its' concurrent validity.
4. The scores from this simulation should be correlated with measures of general aptitude such as the SAT and the GRE to examine the extent to which these two types of measures draw upon the same general abilities.
5. The sample from the current study should be extended to include students, both those trained

in education and those not trained in education, from more than one institution. The sample of teachers should also be represented by a greater range of experience. The current study was conducted using samples of students from only one institution and teachers with a great deal of experience. The inclusion of teachers having a wider range of classroom teaching experience would allow for more specific investigation of the influence of teaching experience on their performance on this simulation exercise.

6. The current format of the simulation exercise should be further investigated using a scoring system tailored to the needs of a different program. The current study relied primarily on the indicators of effective teaching outlined within the BTAP Program. Other training programs, or other states, may suggest alternative scoring procedures for the types of problems found in this simulation.

7. The specific items found on each of the competency scales should be re-evaluated. Items may be revised, rewritten, or discarded at this time. Specific emphasis should be placed on the items comprising the Accountability competency. Those subjects without pedagogical training or

classroom teaching experience consistently scored higher on this competency. Perhaps the items fail to define the competency as intended in the research findings.

REFERENCES

Adams, R. D. (1982). Teacher Development: A look at changes in teacher perceptions across time. Paper presented at the annual meeting of the American Educational Research Association, New York.

Andrews, B. J. (1980). Can professional competence be measured? In E. H. Lovelane (Ed.), Measuring the hard to measure (pp. 39-52). San Francisco: Jossey-Bass, Inc.

Association of Teacher Educators (1988). Teacher Assessment. Reston, VA: Author.

Ayers, J. B. (1980). A longitudinal study of teachers. Paper presented at the annual meeting of the American Educational Research Association, Boston.

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

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

Berliner, D. (1986). In pursuit of expert pedagogue. Educational Researcher, 17(7), 5-13.

Bloom, B. (1986). Automaticity. Educational Leadership, 43 (5),70-77.

Borich, G. D. (1977). The appraisal of teaching: Concepts and process. Reading, MA: Addison Wesley.

Brookes, D. M. & Hawke, G. (1985). Effective and ineffective session opening teacher activity and task structures. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Brophy, J., & Good, T. L. (1986). Teacher behavior and student achievement. In M. Wittrock (Ed.), Handbook of research on teaching, third edition, New York: Macmillan.

Bunch, J. B. (1990, in press). Professional Advice on Videotaping Instruction. Performance and Instruction Journal.

Bureau of Educational Research (1984). Beginning Teacher Assistance Program Phase II Final Report. Charlottesville, VA: Author.

Calderhead, J. (1981). A psychological approach to research on teachers' classroom decision making. British Educational Research Journal, 7, 51-57.

Chiu, Y. C. (1989). An Investigation of the Relationship between the National Teacher Examinations and a Performance-Based Assessment of Professional Knowledge. Unpublished doctoral dissertation, University of Virginia, Charlottesville.

Clark, D. C., Snow, R. B., Newby, T. J., & Cook, V. A. (1985). Perceived Origins of Teaching Behavior. Journal of Teacher Education, 36, 49-53.

Coker, H., Medley, D. M., & Soar, R. (1980). How valid are expert opinions about effective teaching? Phi Delta Kappan, 62(2), 131-134,149.

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

Crawford, J., & Robinson, C. (1983, March). A review of empirical research in classroom management. Classroom management: Reviews of the teacher education and research literature (pp. 43-62). Princeton, NJ: Educational Testing Service.

Cross, L. H. (1985). Validation of the NTE Tests for Certification Decisions. Educational Measurement: Issues and Practices, 4(3), 7-10.

Darling-Hammond, L., Wise, A., Pease, S..(1983). Teacher Evaluation in the Organizational Context: A Review of the Literature. Review of Educational Research, 53, 285-328.

Dewalt, M. W. (1985) The Effects of Teacher Training on Teacher Competence. Unpublished doctoral dissertation, University of Virginia, Charlottesville.

Educational Testing Service. (1984). A guide to the NTE Core Battery Tests. Princeton, NJ: Author.

Eisenburg, T. E. & Rudner, L. M. (1988). State Testing of Teachers: A Summary of Current Practices. Washington, DC: American Institute for Research.

Ellett, C. D., Capie, W., & Johnson, C. E. (1980). Assessing teacher performance. Educational Leadership. 38, 219-220.

Fagot, B. I. (1978). Reinforcing contingencies for sex-role behaviors: effect of experience with children. Child Development, 49(1), 30-36.

Florida State Department of Education (1984). Florida Teacher Certification Examination. Technical Report: 1981-1982. Tallahassee, FL: Author. (Eric Reproduction Document Service No. ED 248 190)

Florida State Department of Education (1983). Domains of the Florida Performance Measurement System. Tallahassee, FL: Author. (Eric Reproduction Document Service No. ED 283 777)

Fogarty, J. L., Wang, M. C., & Creek, R. (1983). A Descriptive Study of Experienced and Novice Teachers' Interactive Instructional Thoughts and Actions. Journal of Educational Research, 77(1), 22-32.

Frederiksen, N., Saunders, D. R., Wand, B. (1957). The In-Basket Test. Psychological Monographs: General and Applied, 71(9), 1-28.

Frye, H. B. (1972). Development and Initial Validation of the Teaching Situation Reaction Test (Elementary Form), A Measure of Teacher Effectiveness. Dayton, OH: Dayton University, Ohio Research Institute.

Gage, N. L. & Needels, M. C. (1989). Process-Product Research on Teaching: A Review of Criticisms. The Elementary School Journal, 89(3), 253-300.

Gage, N. L. (1978). The scientific basis of the art of teaching. New York: Teachers College Press.

Gallup, A. M. (1986). The 18th Annual Gallup Poll of the Public's Attitudes Toward the Public Schools, Phi Delta Kappan, 68(1), 43-59.

Gallup, A. M. (1984). The Gallup Poll of Teachers' Attitudes Toward the Public Schools. Phi Delta Kappan, 66(2), 97-107.

Gallup, A. M. & Clark, D. L. (1987). The 19th Annual Gallup Poll of the Public's Attitudes Toward the Public Schools, Phi Delta Kappan, 69(1), 17-30.

Gallup Poll: Public Opinion (1986). Wilmington, DE: Scholarly Resources.

Givens, S. (1988). ETS Unveils Plan to Replace National Teacher Examination. AACTE Briefs, 9(9-10), 11,16.

Harris, W. U. (1981). Teacher command of subject matter. In J. Millman (Ed.), Handbook of teacher evaluation (pp. 58-72). Beverly Hills, CA: Sage.

Hays, L. (1988). A Simulation Test of Teacher Competencies. Unpublished doctoral dissertation, University of Virginia, Charlottesville.

Hodgkins, R., & McKenna, B. (1982). Testing and teacher certification: An explosive combination. Educational Measurement: Issues and Practice, 1(2), 10-26.

Hoffman and Roper (1985). How valuable is teacher training to beginning teachers? Paper presented at annual meeting of the American Educational Research Association, Chicago, IL. (ERIC Document Reproduction Service No. ED 258 967).

Holmes Group. (1986). Tomorrow's Teachers. East Lansing, MI: Author.

Hoste, R. (1982). Sources of influence on teaching practice in the evaluation of courses in teacher education. Journal of Education for Teaching, 8(3), 252-261.

Housner, L. & Griffey, D. (1985). Teacher cognition: Differences in planning and interactive decision-making between experienced and inexperienced teachers. Research Quarterly for Exercise and Sport, 56(1), 45-53.

Huberman, M. (1985). What Knowledge is of most worth to teachers? A knowledge use perspective. Teaching and Teacher Education, 1(3), 251-262.

Huberty, C. J. (1975). Discriminant Analyses. Review of Educational Research, 45 (4), 543-598.

Jackson, P. W. (1965). The way teaching is. NEA Journal, 54(8), 10-13, 62.

Kane, M. (1986). The role of reliability in criterion-referenced tests. Journal of Educational Measurement, 23(3), 221-224.

Kinney, L. (1964). Certification in education. Englewood Cliffs, NJ: Prentice-Hall.

Klemp, G. 0. (1979). Identifying, measuring, and integrating competence. In P. S. Pottinger & J. Goldsmith (Eds.), Defining and measuring competence (41-52). San Francisco: Jossey-Bass.

Leinhardt, G., Weidman, C., & Hammond, K. M. (1987). Introduction and integration of classroom routines by expert teachers. Curriculum Inquiry, 17(2), 135-176.

Leinhardt, G. & Greeno, J. G. (1986). The Cognitive Skill of Teaching. Journal of Educational Psychology, 78(2), 75-95.

Levin, B. (1979). Teacher Evaluation - A Review of Research. Educational Leadership, 37 (3), 240-245.

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

Lortie, D. C. (1975). Schoolteacher. Chicago: University of Chicago Press.

Marchant, G. J. & Bowers, N. D. (1988). Teacher Agreement with Research-Based Effective Teaching Behaviors. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

McNergney, R. F., Medley, D. M., Aylesworth, M. S., & Innes, A. H. (1983). Assessing Teachers' Planning Abilities. Journal of Educational Research, 77(2), 108-111.

Medley, D. M., Coker, H., & Soar, R. S. (1984). Measurement-based evaluation of teacher performance: An empirical approach. New York: Longman.

Medley, D. M. (1982). Teacher Competency Testing and the Teacher Educator. Charlottesville, VA: Association of Teacher Educators and the Bureau of Educational Research.

Medley, D. M. (1981). Training teachers to be professional decision-makers. Symposium at the American Educational Research Association Meeting.

Medley, D. M. (1978). Some Obstacles to the Valid Measurement of Teacher Competence. A Paper presented at the Conference on Issues in Competency-Based Education, Georgia State University.

Medley, D. M. (1978). Alternative assessment strategies. Journal of Teacher Education, 29(2), 38-44.

Medley, D. M. (1977). Teacher competency and teacher effectiveness: A review of the process-product research. Washington, DC: American Association of Colleges for Teacher Education.

Medley, D. (1972). Early history of research on teacher behavior. International Review of Education, 18, 430-438.

Medley, D. M., Mitzell, H. E. (1958). A technique for measuring classroom behavior. Journal of Educational Psychology, 49, 86-92.

Millman, J. (Ed.) (1981). Handbook of Teacher Evaluation. Beverly Hills, CA: Sage.

Murray, C. K. (1969). The Construct Validity of the Teaching Situation Reaction Test. Journal of Educational Research, 62(7), 323-328.

Myers, J. L. (1979). Fundamentals of Experimental Design. Boston: Allyn and Bacon, Inc.

National Commission on Excellence in Education (1988). American Education: Making It Work. Washington, DC: US Department of Education.

National Commission on Excellence in Education (1983). A Nation At Risk. Washington, DC: US Department of Education.

Needles, M. (1980). Directed teacher change. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Palladino, J. (1980). The charade of testing teacher competency: Relevant criticism for the New York State Education Commissioner's Task Force on Teacher Education and Certification. (ERIC Document Reproduction Service No. ED 196 901)

Pearce, J. A. (1988). An investigation of teacher performance by levels of teaching experience. Unpublished doctoral dissertation, University of Virginia, Charlottesville.

Pedhauzer, E. J. (1982). Multiple Regression in Behavioral Science. New York: Holt, Rinehart & Winston.

Peterson, P., & Clark, C. M. (1978). Teachers' reports of their cognitive processes during teaching. American Educational Research Journal, 15, 555-565.

Piper, M. K., & Houston, W. W. (1981). The search for teacher competence: CBTE and MCT. Journal of Teacher Education, 31(5), 37-40.

Piper, M. K., & O'Sullivan, P. S. (1981). The National Teacher Examination: Can it predict classroom performance? Phi Delta Kappan, 62. 401.

Poggio, J. P., Glasnapp, D. R., Miller, M. D., Tollefson, N., & Burry, J. A. (1986). Strategies for Validating Teacher Certification Tests. Educational Measurement: Issues and Practice, 5, 18-25.

Pottinger, P. (1979). Competence assessment: Comments on current practices. In P. S. Pottinger & J. Goldsmith (Eds.), Defining and measuring competence (25-39). San Francisco: Jossey-Bass.

Quirk, T., Witten, B., & Weinberg, S. (1973). Review of Studies of the Concurrent and Predictive Validity of the National Teacher Examinations. Review of Educational Research, 43 (1), 89-114.

Rabinowitz, W. and Travers, R. (1953). Problems of Defining and Assessing Teacher Effectiveness. Educational Theory, 3 (3), 214-219.

Richardson, M. S. (1985). Perceptions of principals and teachers on effective management of student behavior. Spectrum, 3(3), 25-30.

Riley, J. F. (1982). The development of an instrument to measure creative teaching abilities. Paper presented at annual meeting of the Southwest Educational Research Association, Austin, TX. (ERIC Document Reproduction Service No. ED 220 521)

Ropo, E. (1987). Teachers' Conceptions of Teaching and Teaching Behavior: Some Differences between Expert and Novice Teachers. Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.

Rosenholtz, S. J. & Smylie, M. A. (1984). Teacher Compensation and Career Ladders. Elementary School Journal, 85(2), 149- 166.

Rudner, L. M. (1987). Testing to Improve the Quality of Our School Teachers. Washington, DC: National Governors Association.

Sandefur, J.T. (1985). Competency assessment of Teachers. Action in Teacher Education, 7(1-2), 1-6.

Saunders, S. L., Herbert, J., & Chronister, A. (1989). Using Information to Make Planning Decisions in a Computer Simulation. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Scriven, M. (1981). Summative teacher evaluation. In J. Millman (Ed.). Handbook of teacher evaluation (244-271). Beverly Hills, CA: Sage.

Shavelson, R., Webb, N. & Burstein, L. (1986). Measurement of Teaching. In Gage, N.L. (3rd Ed.), Handbook of Research on Teaching, Chicago, IL.: Rand McNally.

Shavelson, R. J. (1983). Review of Research on Teachers' Pedagogical Judgments, Plans, and Decisions. The Elementary School Journal, 83(4), 392-423.

Smith, B. O., Peterson, D. & Micceri, T. (1987). Evaluation and Professional Improvement Aspects of the Florida Performance Measurement System. Educational Leadership, 44 (7), 16-19.

Smith, D. C. (Ed.). (1983). Essential knowledge for beginning educators. Washington, DC: American Association of Colleges for Teacher Education and ERIC Clearinghouse on Teacher Education.

Southern Regional Educational Board (1984). A Comparison of teacher certification test scores and performance evaluations for graduates in teacher education and arts and sciences in three states. Atlanta, GA. Author.

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

SPSSx Users' Guide (2nd ed.) (1986). Chicago: SPSS.

Thompson, B. (1981). Teachers' Preferences for Various Teaching Methods. NASSP Bulletin, 65(446), 96-100.

Vold, David J. (1985). The Roots of Teacher Testing in America. Educational Measurement: Issues and Practices, 4(3), 5-7.

Washington State Board of Education (1987). State Board of Education Initiatives: Professional Education in the 1980's. Olympia, WA: Author.

Weber, W. A. & Roff, L. A. (1983). A review of the teacher education literature on classroom management. Classroom management: Reviews of the teacher education and research literature (pp. 7-42). Princeton, NJ: Educational Testing Service.

Wilson, A. J. (1985). Knowledge for Teachers: The Origins of the Natioanl Teacher Examinations Program. Paper presented at the annual meeting of the American Educational Research Association. Chicago, IL. ERIC Reproduction Document No. ED 262 049.

Winer, B. J. (1971). Statistical Principles in Experimental Design. New York: McGraw-Hill Book Company.

Wise, A. E., Darling-Hammond, L., McLaughlin, M. W. & Berstein, H. T. (1985). Teacher Evaluation: A Study of Effective Practices. The Elementary School Journal, 86 (1), 61-121.

Appendix A

BTAP Competency Definitions and Indicators

Revised: April 1988

VIRGINIA BEGINNING TEACHER ASSISTANCE PROGRAM
BEGINNING TEACHER COMPETENCIES

INTRODUCTION

The basic purpose of the Beginning Teacher Assistance Program is to assess beginning teachers relative to the degree to which they can demonstrate the possession of selected classroom competencies and to provide assistance to those teachers who fail to demonstrate the possession of these competencies. This document contains a list of the competencies a beginning teacher will be expected to demonstrate in the classroom before receiving the Virginia Collegiate Professional Teaching Certificate.

This program defines a COMPETENT TEACHER as one who possesses certain knowledge and has the ability to translate this knowledge into appropriate classroom teaching behaviors. Given this definition, being a competent teacher is necessary but not sufficient to being an effective teacher.

Teacher competency is necessary because it is presumed that a teacher who cannot demonstrate these competencies does not know enough about teaching and learning to become a good teacher. On the other hand, teacher competency is not sufficient because it is possible that an individual who can demonstrate the possession of these competencies may not necessarily be an effective teacher. The notion of effective teaching is much broader than the demonstrated possession of a limited number of teaching competencies. Competent teaching as defined by BTAP does not include some factors which would be included in the concept of satisfactory teaching performance, e.g., the ability to work well with the other teachers in the school or the ability to communicate effectively with parents. A certificate of competence to teach has not in the past been considered to be a guarantee of satisfactory performance as a teacher nor should it be so regarded in this case.

In addition, the idea of effective teaching implies typical performance while competency assessment according to BTAP is based on demonstrated capability in a specified number of structured situations. An individual may demonstrate the capability of behaving in appropriate ways during the BTAP assessment but, for various reasons, may not typically behave in that manner. BTAP PROPOSES TO ASSESS AN INDIVIDUAL'S CAPABILITIES BUT CANNOT GUARANTEE THAT THE INDIVIDUAL WILL ROUTINELY BEHAVE IN A MANNER CONSISTENT WITH THESE CAPABILITIES.

A. ACADEMIC LEARNING TIME

The competent teacher knows that learning is directly related to the amount of time learners are actively engaged in planned learning activities. The one relationship most clearly established by process-product research is the relationship between academic engaged time and learners' achievement gains.

The beginning teacher should demonstrate knowledge of this competency by:

- planning for the efficient use of class time
- minimizing the amount of time spent on procedural matters
- using non-punitive techniques for maintaining learner involvement in assigned tasks
- maintaining continuous focus on the lesson topic

B. ACCOUNTABILITY

The competent teacher knows the importance of holding learners responsible for completing assigned tasks. It is important for the teacher to make sure that every student actually undergoes the learning experiences planned for him or her. Holding the learner personally responsible for completing assigned learning tasks is also important because it gives learners practice in assuming and discharging personal responsibilities.

The beginning teacher should demonstrate knowledge of this competency by:

- planning just what tasks each learner is supposed to complete
- making clear to the learners what they are expected to accomplish
- clearly establishing consequences of not completing an assigned task
- checking to see whether learners work on their tasks

C. CLARITY OF STRUCTURE

The competent teacher knows that learning is facilitated if the lesson is presented in a clear systematic sequence consistent with the objectives of instruction. Learning is a conscious activity of the learner which proceeds (according to research in human learning) most efficiently when the learner is aware of the relationship of each part of the activity to the other parts and to the whole.

The beginning teacher should demonstrate knowledge of this competency by:

- preparing outlines, reviews, and summaries, beforehand
- beginning the lesson or unit with a statement of purpose, outline, etc.
- making interrelations among parts of the lesson clear to learners
- ending the lesson or unit with a summary or review

D. INDIVIDUAL DIFFERENCES

The competent teacher knows that learners progress at different speeds, learn in different ways, and respond to different kinds of motivation. Few generalizations about learning are better established than this one. Research indicates that teaching strategies should be adapted to these differences if all learners are to achieve at their full potential.

The beginning teacher should demonstrate knowledge of this competency by:

- planning ways of dealing with individual differences in learners' abilities, cultural background, handicaps
- defining different objectives for different learners
- providing alternate ways for different learners to achieve common objectives
- providing for learners with special problems (such as hearing or visual impairment, severe learning differences)
- providing for learners with unusual talents or abilities
- arranging the classroom for easy access for physically handicapped learners

E. EVALUATION

The competent teacher knows that learner progress is facilitated by instructional objectives which are known to the learners and which coincide with the objectives of evaluation# Important information about learner status can be obtained informally while teaching and can be used in making tactical teaching decisions. Additionally, formal assessment of the learner's progress is important in making instructional decisions, grading, and promotional decisions. Competence in matching given instructional objectives with informal and formal evaluation contributes to the soundness of the teacher's decisions during the course of instruction.

The beginning teacher should demonstrate knowledge of this competency by:

- planning evaluation (formal and informal) whenever he or she plans instruction
- designing formal evaluation procedures that are both relevant and fair
- asking questions, observing learners' work, and checking learners' understanding regularly during instruction to evaluate progress
- informing learners about how their performance will be evaluated

F. CONSISTENT RULES

The competent teacher knows that rules for classroom behavior must be clear and consistent and that learners must understand and accept the rules and the consequences of violating them. When rules are unclear or applied inconsistently, classroom management is difficult; when rules are clear and consistently applied, the classroom seems almost to manage itself.

The beginning teacher demonstrates knowledge of this competency if:

- it is seldom necessary to restate rules of conduct
- disruptive pupil behaviors are infrequent
- when disruptive behaviors occur, the teacher deals with them in a non-punitive manner

K. AFFECTIVE CLIMATE

The competent teacher knows that learning occurs more readily in a classroom environment which is non-punitive and accepting. Research indicates that achievement gains are related positively to an affective environment which is non-punitive, i.e., relatively free from hostility and threats.

The beginning teacher should demonstrate knowledge of this competency by:

- avoiding hostility and punitiveness
- acting relaxed, good-humored, and accepting learner behavior

L. LEARNER SELF-CONCEPT

The competent teacher knows that a learner's achievement may be enhanced by improving his self-concept, and that his self-concept is enhanced if the teacher's expectations are high and if the teacher shows appreciation of the learner's personal worth#

The beginning teacher should demonstrate knowledge of this competency by:

- planning lessons that relate to learner's background and interests
- encouraging learners to do better
- praising correct performance of difficult tasks or correct answers to a difficult question
- helping learners develop appreciation of their own cultural heritage
- helping learners develop feelings of personal worth
- showing courtesy to and concern for learners

M. MEANINGFULNESS

The competent teacher knows that learning is facilitated when content is related to learners' interests, common experiences or to information with which they are familiar. Although the importance of meaningfulness in learning has been formally established by research in human learning done in psychological laboratories, practicing teachers have recognized its importance for many years.

The beginning teacher should demonstrate knowledge of this competency by:

- planning ways of relating instruction to interests and previous knowledge of learners
- pointing out relationships between lesson or unit content and things learners already know
- pointing out relationships between lesson or unit content and outside or "real world" interests of learners
- asking questions of learners that require them to identify relationships between what they are learning and something they already know
- planning activities that require learners to identify relationships between what they are learning and something that is important to them outside the classroom
- relating instruction to the cultural backgrounds of learners

P. PLANNING

The competent teacher knows the importance of deliberate and varied planning activities. Instructional planning should reflect the teacher's knowledge that: 1) learning activities should match the instructional objectives; 2) learning is facilitated when ideas are communicated in more ways than one and when two or more sensory modes are employed; 3) the current literature on the teaching profession should be consulted; and 4) learners' scores on standardized tests contain important and useful information about the class as a group and about individual learners.

Planning activities should reflect the teacher's knowledge of instructional objectives and activities, multiple instructional modes, the current professional literature, and the interpretation of test data to facilitate instruction and learner progress.

The beginning teacher demonstrates knowledge of this competency by:

- using test data in defining objectives or choosing learner activities, content, materials, or media
- using relevant professional literature in defining objectives or choosing learning activities, content, materials or media
- defining objectives that move learners toward long-term goals
- defining objectives on the basis of differing needs of groups and individuals
- defining objectives in measurable terms
- using information about test reliability, validity, and test norms
- using objectives as a basis for planning learning activities
- planning different activities for learners with different abilities, interests, and cultural backgrounds
- planning alternative ways for learners to achieve the same objectives
- planning for use of different media

Q. QUESTIONING SKILLS

The competent teacher knows how to phrase different kinds of questions, that different types of questions are most effective in promoting different types of learning, and which type to use for which purpose. Asking questions is a major professional teaching tool and the skillful use of questions has been extensively studied and researched.

The beginning teacher demonstrates knowledge of questioning skills by:

Questioning skills during drill and practice sessions

- asking questions that test learners' ability to recall or apply material learned previously
- avoiding difficult questions, higher order questions, and questions that call for original answers

Questioning skills during a recitation

- asking questions that test students' knowledge or comprehension of a topic being studied
- giving feedback on student answers and voluntary comments
- answering students' questions

Questioning skills during a discussion

- asking open-ended questions that call for student opinions, beliefs, etc.
- giving positive feedback to students answers and voluntary comments
- accepting student answers and voluntary comments
- avoiding questions that test students' knowledge or comprehension

R. REINFORCEMENT

The competent teacher demonstrates awareness that the skillful use of reinforcement is an effective means of encouraging and discouraging particular behaviors. Establishment of the importance of reinforcement in modifying human behavior and clarification of principles governing its use has been one of the principal achievements of research in learning.

The beginning teacher should demonstrate knowledge of this competency by:

- giving positive rather than negative feedback
- not using punishment to motivate learners
- calling attention to desirable behaviors
- using positive feedback to cue learners to behavioral expectations

S. CLOSE SUPERVISION

The competent teacher knows that more is learned during individual, small and whole group activities if the learners are monitored. Research indicates that learning is facilitated during activities when they are monitored rather closely, presumably because this increases the amount of learner engagement. Close supervision also provides opportunities for the teacher to assist and encourage learners.

This competency is demonstrated only when learners are working independently or in small groups. The beginning teacher should demonstrate knowledge of this competency by:

- monitoring activity of all learners
- helping learners who have difficulties

W. AWARENESS

The competent teacher knows that effective classroom management depends on knowing what is occurring in the classroom and that the learners perceive the teacher knows what is going on# The teacher who is aware of what is going on in the classroom is likely to increase learner participation in learning activities and reduce disruptions.

This competency is demonstrated when the teacher is working with the class as a whole. The beginning teacher should demonstrate knowledge of this competency by:

- maintaining constant awareness of level of interest and attention of learners
- making learners aware of teacher awareness

Appendix B

A Simulation Exercise in Classroom Decision Making

Instructions and Selected Problems

A Simulation Exercise in Classroom Decision Making

Directions:

This is an exercise in making decisions in various teaching situations. This exercise includes a series of problems depicting situations that may occur in the actual classroom. In the classroom a teacher must often make decisions quickly, so in this simulation you will be given only a few seconds to make each decision.

You will see printed on the screen and hear read descriptions of classroom situations. After seeing and hearing each situation, you will hear described several possible courses of action. These options will not appear on the screen. Listen to each option. Mark the bubble in the first column if you think you might do that. Mark the bubble in the second column if you do not think you would follow that course of action.

Think about what you would do as a teacher in a normal classroom. Do not try to think of unusual or rare circumstances or exceptions to normal practice. This is not a test in which each problem has a "right" or "best" answer. You may find several alternatives appropriate in some cases and few or none in others.

Remember to mark the first column if you might follow that course of action. Mark the second column if you probably would not follow that course of action. Mark your answer quickly because you will only have five seconds to make each decision.

Let's try a sample problem now.

SAMPLE PROBLEM

You have just begun a class discussion when the principal unexpectedly appears at your door, tells you to continue with whatever you were doing, and seats himself in the back of the room. You might:

A. Try to do as he says and continue with the discussion as if he were not there.

B. Invite him to leave.

C. Ask him to join in the class discussion.

D. End the discussion and begin a lesson that includes the teaching skills that are part of the principal's evaluation.

If there wasn't enough time to think over possible alternatives of action on the previous sample problem, then this exercise is working because that's the way teaching is. Teaching involves mainly decisions based on insufficient information that must be made too quickly to allow adequate thought - just like those in this exercise.

We are now ready to begin this

simulation exercise

Remember to mark the first column if you might follow that course of action. Mark the second column if you probably would not follow that course of action. Mark your answer quickly because you will have only five seconds to make each decision.

Here is the first situation.

A. Elementary students in your school are expected to line up are not supposed to talk as they walk through the halls. You have a first-grade class. At the beginning of the school year, you might:

1. Explain the rules to your class before the first time the class walks in the hall.

2. Walk at the head of the line.

3. Stop the first time anyone talks in line and say, " Remember that the school rule says we must be quiet in the hall."

4. If nobody is talking, say, "Good, you are walking quietly."

5. Say, "The first student who talks or gets out of line will have to stay after school."

I. You are teaching in a private school in a predominantly Christian community. It is December and your class's turn to put up the hall bulletin board. Most of the students want to do a Christmas mural. You have one Jewish girl, Becky, in the class. You might:

38. Have some of the other students help Becky do a second mural about Hanukkah.

39. Send Becky to the library during this activity to do special research.

40. Have Becky share her experiences of Hanukkah with the class in a discussion.

41. Ask Becky to paint a picture by herself to be placed next to the Christmas mural.

42. Have Becky work on the Christmas mural with the rest of the children.

FF. Speeches on current events topics are due on Friday. You might say:

143. "Be ready on time if you want to get full credit for your speech."

144. "I'd like as many of you as possible to be ready for Friday. The rest of you can give your speeches on Monday."

145. "I am going to take ten points off your grade if you are not ready when I call on you."

146. "If you are not ready when I call on you, you must stay after school and finish."

LL. You are teaching addition facts to eighth graders with learning disabilities in math. You might:

169. Have the students play a game using flash cards.

170. Conduct an oral math drill using addition problems.

171. Teach students to use a classroom bowling set and to keep score.

172. Have the students select items that they could buy from the school store with five dollars.

Appendix C

Simulation in Classroom Decision Making

Background Questionnaire

SIMULATION OF CLASSROOM DECISION-MAKING

Background Information Form

Purpose - The purpose of this survey is to gather information supplementary to that obtained from the simulation exercise. The responses from this survey as well as the simulation will be tabulated and examined in summary form so as to ensure that individual responses remain confidential.

Directions - Please read every item carefully and respond as accurately as possible by bubbling in the appropriate letter on the answer sheet. Should you find that an item does not apply to you, leave the item blank. Please be as specific as possible when responding to a particular item.

When you have completed this survey, please return it to the instructor as you leave the room.

* Please be sure that your social security number is recorded on both this questionnaire and bubbled on your answer sheet.

Begin recording your responses to the following items on the answer sheet at number 189.

PART A Social Security Number: _____

189. Do you wish to receive a copy of the simulation results?

A. Yes
B. No

If yes, please supply your current mailing address below:

Name: _____

Address: _____

_____

190. Age:

A. 18 or under
B. 19-21
C. 22-30
D. 31-40
E. over 40

191. Sex:

A. Male
B. Female

192. Highest Level of Education:

A. Have not yet received undergraduates degree
B. Bachelors Degree (____ BA, ____ BS)
C. Bachelors plus (# of graduate credits: _____)
D. Masters Degree
E. Masters plus (# of graduate credits: _____)

193. If you are currently in an undergraduate degree program, please indicate your present status below.

A. 1st Year
B. 2nd Year
C. 3rd Year
D. 4th Year
E. 5th Year

194. Are you currently enrolled in an undergraduate <u>teacher preparation</u> program?

A. Yes
B. No

Undergraduate Major:

_____

195. If you are currently in an undergraduate degree program, please indicate the number of education courses you have previously enrolled in.

A. 1 or less
B. 2
C. 3
D. 4
E 5 or more

196. Classroom Teaching Experience:

(please indicate the highest level only)

A. No Classroom Teaching Experience
B. Field Observations Only
C. Mini-Teaching (prior to Student Teaching)
D. Student Teaching
E. Classroom Teaching Experience

197 - 200. Please indicate the type of program you are currently enrolled in:

197. Elementary Education A. Yes B. No

198. Secondary Education A. Yes B. No

Area of Concentration: _____

199. Special Education A. Yes B. No

(please indicate grade levels: _____

200. Program other than Education A. Yes B. No

Please specify: _____

PART B

* NOTE: Complete the following questions only if you have had classroom teaching experience

Please feel free to indicate additional comments about a particular item or items in the space provided below the rating scale.

201. Please indicate the extent to which you feel the situations in this simulation resembled normal classroom experiences.

A. never
B. seldom
C. often
D. always

Comments:

202. To what extent do you feel the limited response time in this simulation corresponded to that of the actual classroom?

A. never
B. seldom
C. often
D. always

Comments:

203. Do you feel this simulation reflects the actual classroom decision-making process compared to the typical pencil-and-paper test?

A. no, it is much worse than pencil-and-paper tests
B. no, it is worse than pencil-and-paper tests
C. it is about the same as pencil-and-paper tests
D. yes, it is somewhat better than pencil-and-paper tests
E. yes, it is much better than pencil-and-paper tests

Comments:

* Please answer questions 204 through 210 only if you are currently employed as an practicing teacher

204. If you have been an practicing teacher, how many years have you been teaching?

_____

Grade Levels Taught:

205. Kindergarten A. Yes B. No
206. 1-3 A. Yes B. No
207. 4-6 A. Yes B. No
208. 7-8 A. Yes B. No
209. 9-12 A. Yes B. No

Please indicate your area of certification:

210. Elementary Education A. Yes B. No
211. Secondary Education A. Yes B. No

Subject Area: _____

212. Special Education A. Yes B. No

Grade levels: _____

213. Other: A. Yes B. No

Please specify: _____

What do you feel are the major strengths of this simulation exercise? Please be specific.

What do you feel are the major weaknesses of this simulation exercise? Please be specific.

Thank you for your participation in this simulation exercise.

Consent Form

I agree to participate in a Simulation Exercise of Classroom Decision-Making. This exercise will take approximately one hour to complete. I understand that the results from this exercise will be summarized and presented in aggregate form so that my individual performance will remain confidential.

_____ _____

Signature Date