

DETERMINING THE RELIABILITY AND VALIDITY  
OF AN INSTRUMENT TO MEASURE BEGINNING TEACHER  
KNOWLEDGE OF READING INSTRUCTION

by

LESLIE TARBET

B. A., Northwestern University, 1961

---

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

College of Education

KANSAS STATE UNIVERSITY  
Manhattan, Kansas

1984

Approved By:

Nancy J. Smith  
Major Professor

LD  
2668  
.T4  
1984  
T37  
C. 2

A11202 666065

COPYRIGHT

DETERMINING THE RELIABILITY AND VALIDITY  
OF AN INSTRUMENT TO MEASURE BEGINNING TEACHER  
KNOWLEDGE OF READING INSTRUCTION

LESLIE TARBET

1984

## TABLE OF CONTENTS

	Page
LIST OF TABLES . . . . .	iii
ACKNOWLEDGEMENTS . . . . .	iv
CHAPTER	
I. INTRODUCTION . . . . .	1
Statement of the Purpose of the Study. . . . .	2
Limitations. . . . .	4
Definition of Terms. . . . .	4
Summary. . . . .	5
II. REVIEW OF THE LITERATURE . . . . .	7
Growth of Teacher Evaluation . . . . .	7
Recent Proposals Calling for Change. . . . .	34
Summary. . . . .	37
III. METHODOLOGY. . . . .	39
Procedure for Developing the Test. . . . .	39
Pilot Study. . . . .	41
Method of Analyzing Data . . . . .	42
IV. RESULTS. . . . .	44
Overview . . . . .	44
Pilot Test I . . . . .	45
Pilot Test II. . . . .	52
Summary. . . . .	55
V. CONCLUSIONS. . . . .	56
Test Reliability Results . . . . .	56
Item Analysis. . . . .	57
Discussion . . . . .	59

	Page
REFERENCES . . . . .	61
APPENDIXES . . . . .	64
Appendix A, Pilot Test I . . . . .	64
Appendix B, Inter-rater Item Classification. . .	81
Appendix C, Pilot Test II. . . . .	86
Appendix D, Raw Data . . . . .	104

LIST OF TABLES

Table	Page
1. Comparison of Pilot Study Test Data . . . . .	46
2. Difficulty Index Comparisons . . . . .	48
3. Discrimination Index Comparisons . . . . .	50
4. Inter-rater Agreement on Knowledge Areas . . . . .	51
5. Comparison of Areas of Knowledge, Pilot Test II Data . . . . .	53
6. Difficulty Index Comparison of Areas of Knowledge from KTREG II. . . . .	54
7. Discrimination Index Comparison of Areas of Knowledge from KTREG II . . . . .	54

## ACKNOWLEDGEMENTS

The writer wishes to thank the following:

Dr. Nancy Smith for guidance throughout this study and for the many hours of assistance.

Dr. Clyde Colwell and Dr. Robert Newhouse for their time and suggestions throughout the study.

Phi Delta Kappa for their support through a monetary grant.

My family and friends for their ongoing support.

CHAPTER I  
INTRODUCTION

The current concern for excellence in education challenges educators to re-examine the content of reading courses and students' knowledge of reading instruction. This is necessary to ascertain whether students are grasping the necessary information to meet today's standard for the teaching of reading.

The widely publicized "A Nation At Risk: The Imperative for Educational Reform," published in 1983 by the National Commission of Excellence in Education, revealed the decline in basic skills competence to the general public. This has created a public demand for re-evaluation and improvement in reading and other basic skills. This quest for excellence especially applies to those who are studying to become elementary teachers of reading. Parents and educators alike have the right to expect that beginning teachers are knowledgeable about the subjects they will be required to teach.

Many universities, colleges, and departments of education have responded to this call for reform with modifications of teacher education programs. During the last five years, according to the Chronicle of Higher Education (Sept. 7, 1983), 85 percent of schools, colleges and departments of education (SCDE's) have begun to improve the quality of their teacher education programs and about 75 percent have raised their entrance requirements. An AACTE survey of SCDE's published

in the July/August, 1983 Briefs concurs with this finding. However, to begin to meet this challenge, at the specific teaching knowledge level, reading professionals within schools of education will need to examine closely their course content and pre-service elementary students' knowledge to establish criteria which will ensure that graduates are at a safe level of beginning professional practice. This will mean evaluating their own effectiveness in preparing future teachers to teach reading in order to meet societal demands and expectations.

#### Statement Of The Purpose Of The Study

Several tests measuring knowledge of teaching reading have been developed in the past. In 1971, J. T. Sandefur developed a model for evaluating teacher education graduates, part of which encompassed teacher evaluation in reading. His model was comprised of obtaining and analyzing data from four sources: career-line data, direct classroom observations, student, peer and supervisor ratings, and standardized methods. Each of these took time and money to complete. This complexity led others to try to approach this problem from a simpler perspective.

Educational Testing Service (ETS), in their role as test developer for Phase II of The Beginning Teacher Evaluation Study, conducted in the State of California in 1973-1974, devised a Knowledge of Teaching Reading instrument to measure how knowledgeable beginning teachers were in this specific content area.

Sterl Artley and Veralee Hardin of the University of



Missouri published an "Inventory of Teacher Knowledge of Reading" in 1971 and revised it in 1975. Neither of these two forms of the test were widely accepted or used, due in part, by the poor reviews received in Buro's Mental Measurement Yearbook.

During 1977-1979, the University of Houston developed and completed two pilot studies using their Student Teacher Summative Performance Instrument (PEI) in order to ascertain their teacher education program effectiveness and validation. The instrument was found to be able to distinguish high competent students from low competent students. A change in the administration at the University caused the long-range use of this instrument to be terminated.

Through legislation passed in 1978, the State of Florida's teacher education system established several new components in its educational program. Among these new components is a call that all teacher candidates seeking certification must be able to demonstrate mastery of essential generic and specialized teaching competencies on a written examination. This examination covers knowledge of teaching of reading. Oklahoma (effective in February, 1982) also requires competency examinations in subject areas before graduation.

These attempts do indicate a conscious desire on the part of reading educators and institutions to investigate the preparedness of pre-service elementary teachers to teach reading in the elementary grades. Much has been written about the need for measuring knowledge of teaching reading but little has been done about it.

The major purpose of this study was to construct a Knowledge of Teaching Reading in Elementary Grades test. The results of giving the test in a Pilot Study were then investigated to measure the reliability and validity of the test.

#### Limitations

The study is limited in that it will be a cognitive measure of knowledge of teaching reading. It is realized that there are other aspects necessary to be an effective teacher of reading and that there are other methods of evaluation.

#### Definition of Terms

Content Validity--the extent to which the content of the test represents a balanced and adequate sampling of the traits or learning outcomes about which conclusions are to be drawn.

Difficulty Index--sometimes called difficulty range, indicates the proportion of students who pass a given test item. It is also expressed as "p value."

Discrimination Index--gives the difference in correct responses between the high and low groups as a proportion of either group, i. e., if all of the high group and none of the low group are correct, the index equals 1.00 or a perfect discrimination for the item.

Face Validity--a professional appraisal of what appears to be valid for the content the test attempts to measure.

Knowledge of Teaching Reading in the Elementary Grades test--the test created in this study. It is also referred to as KTREG test.

Pre-service elementary students--those students who have completed all the necessary coursework prior to exiting an institution's program for teaching reading in the elementary grades.

Reading competencies--the knowledge of teaching reading in the elementary grades.

Reliability--a measure of the consistency of a test.

Validity--an indication of whether a test measures what it says it measures.

### Summary

The major purpose of this study was to construct a Knowledge of Teaching Reading in Elementary Grades test and to determine its reliability and validity.

Chapter 2 will provide a background of research related to the study. Teacher evaluation and recent developments on the education scene will be reviewed to show the need for this test instrument.

Chapter 3 describes the methodology used to design the test and gather the data.

Chapter 4 describes the findings from two administrations of the test.

Chapter 5 summarizes the findings and states conclusions drawn from the examination of the data. It also makes some

suggestions in which future revisions of the test can be improved as well as future applications for the test.

CHAPTER II  
REVIEW OF THE LITERATURE

Growth Of Teacher Evaluation

According to J. T. Sandefur (1970), "evaluating teacher effectiveness has been the most difficult of all problems the education community has had to face." (p. 3) Many educators and authorities have given and written their opinions on what constitutes effective teaching, yet few solid advances have been made in the development of test instruments to uniformly measure teaching behavior. Sandefur said that the diverse opinions of authorities as to what constitutes effective teaching has stunted the creation and development of evaluation measures.

Over the years many new theories of teaching and learning have emerged, accompanied by methodologic support, to be added to what already had accumulated. Teachers had an almost unlimited variety of teaching behaviors to choose from as models. Consequently this wealth of theory and methodology caused a lack of uniform, definitive terminology to describe teaching. Any observer of teaching behavior was free to pick and choose among this explosion of terminology to write a "value judgement." The lack of experimental research done in teacher education prior to 1960, either to determine the characteristics of good teaching or to measure them also constrained the defining of effective teaching.

However, in the early 1960's, due to significant amounts

of federal money made available to teacher education as well as to researchers, there emerged a great deal of research about teacher education.

In summing up the information gathered from this era, Sandefur (1970) relied on a survey of the research in teacher education conducted during the period 1955-1971, as reported by Peck and Tucker. He says they identified six themes which seemed to emerge from the body of research which were significant. They were:

1. A "systems" approach to teacher education, often called "instructional design," substantially improves its effectiveness. A good deal of research clustered around three special cases of the general model: training teachers in interaction analysis, micro-teaching, and behavior modification.
2. Teacher educators should practice what they preach. That is, when teachers are treated in the same way they are supposed to treat their pupils, they are more likely to adopt the desired style of teaching behavior.
3. Direct involvement in the role to be learned, or such close approximations as sensitivity-training laboratories or classroom simulation laboratories, produce the desired teaching behavior more effectively than remote or abstract experiences such as lectures on instructional theory.

4. Using any or all of the techniques just mentioned, it is possible to induce a more self-initiated, self-directed effective pattern of learning, not only in teachers but, through them, in their pupils.
5. Traditional ways of educating teachers have some intended effects, but they also have some quite undesirable effects...
6. One long-needed methodological advance is beginning to appear in the research: the use of pupil gain measures as the ultimate criterion of the effectiveness of any given process of teacher education. (p.3-4)

Edward C. Pomeroy wrote in the Forward to Haefele's (1971) "Systems and Modeling: Self-Renewal of Teacher Education," that

our society is undergoing an intense period of introspection during which the fabric of our institutions is being tested to determine its responsiveness to demand for innovativeness, relevance, and change. Particularly, teacher education has engaged in rigorous self-studies in response to pressures from students and faculty as well as from the public. (p. iv)

In order that teachers make a difference in the learning outcomes of students, it is incumbent upon the teacher training institutions to exercise quality controls which will ensure the effectiveness of teacher education.

Arthur W. Heilman (1965) in his "Effects of An Intensive In-Service Program on Teachers' Classroom Behavior and Pupils'

Reading Achievement," makes use of Mary Austin's remark from her book, The Torch Lighters, Tomorrow's Teachers Of Reading, that "potential teachers of reading receive meager training in existing teacher training institutions. The totally inadequate preparation for such an extremely difficult and complicated teaching role has necessitated a need for effective in-service training program." (p. 1)

Heilman goes on to say that the common format of in-service training programs has proven to be totally inadequate to meet the needs of teachers and school systems. The "short term reading conferences" and "highly standardized field courses in reading" have had little impact on the serious problem of reading instruction in American schools. Morrison and Austin (1977) returned to the same population used in The Torch Lighters, Tomorrow's Teachers Of Reading to explore their progress in meeting recommendations made by Austin in her book. They found that a majority of the recommendations were in effect. One of the most important was the recommendation that the equivalent of a three hour course in reading be required of prospective teachers of reading. They also found some significant changes had taken place in teacher preparatory programs such as extending the scope of the reading programs to include more courses, more specialization opportunities, and broader content coverage. Another important discovery they found was that less than half of the respondents indicated any recommendations for the future programs they would propose to initiate at their respective institutions.



Related to this problem, says Heilman (1965), is the fact that

the teaching of reading is little affected by the hundreds of examples of fragmented research presently available on teaching beginning reading. There has been in the past and there is at the present a notable lack of synthesizing, translating, and implementing available knowledge from research into classroom practices and procedures. (p. 3)

He also points out that the training of elementary teachers, in the majority of cases, has not conditioned them to read or analyze research. Morrison and Austin (1977) found that about 50 percent of the institutions they canvassed had some emphasis on acquainting prospective teachers with techniques, interpretation, and evaluation of current and past research. They received comments such as, "Varies from course to course as to research emphasis." or "Done only on a graduate level." or "Let's not try to make the undergraduate 'how to' course into a graduate research course. Other than a few professional journals, I hesitate to cover research."

Teaching is a highly complex activity. Improvement of the teaching of reading calls for insights into present strengths and weaknesses. It also calls for knowledge of alternative ways of approaching and achieving specific goals and behaviors.

Under the sponsorship of the American Association of Colleges for Teacher Education (AACTE), J. T. Sandefur (1970) proposed a model for the evaluation of teacher education graduates. It included three major generalizations, each with

a series of sub-generalizations drawn from and presented as a synthesis of research findings related to teaching characteristics of good teachers. They served as the basis of his evaluation model.

1. Good teaching utilizes maximal involvement of the student in direct experiential situations.
  - a. Good teachers attempt to foster problem-oriented, self-directed, actively inquiring patterns of learning behavior in their students.
  - b. Good teachers elicit pupil-initiated talk and allow more pupil-initiated exploration and trial solutions.
  - c. When teachers try to elicit independent thinking from their students, they get it.
  - d. Good teachers involve students in decision-making processes in active, self-directing ways.
  - e. Teachers who are interested in student involvement are less prone to dominate the classroom through lecture and other teacher activities.
2. Good teaching encourages maximal freedom for the student.
  - a. Good teachers use significantly more praise and encouragement for the student.

- b. They accept, use, and clarify student ideas more often.
  - c. They give fewer directions, less criticism, less justification of the teacher's authority, and less negative feedback.
  - d. They use a relaxed conversational teaching style.
  - e. They use more divergent questions, do more probing, and are less procedural.
  - f. They are more inclined to recognize the "affective climate" of the classroom and are responsive to student feelings.
  - g. Teachers with low dogmatism scores are more likely to use indirect methods than those with more closed-minded attitudes.
3. Good teachers tend to exhibit identifiable personal traits broadly characterized by warmth, a democratic attitude, affective awareness, and a personal concern for students.
- a. Good teachers exhibit characteristics of fairness and democratic behavior.
  - b. They are responsive, understanding and kindly.
  - c. They are stimulating and original in their teaching.
  - d. They are responsible and systematic.
  - e. They are poised and confident, and emotionally self-controlled.

- f. They are adaptable and optimistic.
- g. They are well-versed in subject matter and give evidence of a broad cultural background. (p. 28-29)

Sandefur's evaluation model was comprised of four major categories of data sources: 1) career line information, 2) direct classroom observation, 3) pupil, peer, and supervisory evaluation, and 4) standardized measurements. From the career line information it was suggested that teacher training institutions that were going to evaluate their graduates should somehow systematically collect data on wastage from teaching, promotions, advanced degrees earned, writing, research, project activity and the mobility of teachers.

From the direct classroom observation systems, it was recommended that one of two systems be used; either a 14-category modification of the Flanders system and the Hough variation of interaction analysis, or the Classroom Observation Record, a product of the Teachers' Characteristics Study. Both instruments are valid and are highly objective in assessment.

From the third source of data, it was suggested that the Student Evaluation of Teaching developed by Veldman and Peck be used by students. Peers and supervisors could use a teacher evaluation form developed by Kansas State Teachers College.

The fourth area of data was that of standardized measures. It was suggested that institutions wishing to measure pupil-gains choose from available tests. These would include tests of intelligence, achievement, personality and attitude.

Sandefur recommended that an institution attempting to implement a comprehensive evaluation system for its teacher education graduates based upon his model take into account the expense and time needed for a comprehensive evaluation program. Expense might or might not be a factor. The time involved would certainly be a prime factor for consideration.

In 1973 the California Commission for Teacher Preparation and Licensing developed and conducted the Beginning Teacher Evaluation Study (BTES) in order to determine what factors should be considered in this process. The second phase of this study, the hypotheses-generating and instrument-developing phase, was conducted by the Educational Testing Service (ETS).

ETS was given two tasks to complete. The first was to develop an assessment system to measure both teacher and pupil behaviors as well as other factors which might be related to those behaviors. The second task was to generate hypotheses about the inter-relationships between teacher and pupil behaviors and related factors.

The study was conducted in 43 schools with a total of 95 teachers participating. Because of the complexity of Phase II, a variety of techniques was used to measure teacher and pupil behaviors. Among the techniques used to measure teacher behavior in reading was a Knowledge of Teaching Reading Test. In the BTES, the knowledge of teaching reading was looked upon as a somewhat more complex concept than subject knowledge and involved not only the techniques of classroom management

(what Sandefur synthesized from research) but also the methods and materials which can be used to teach a given subject (again concurring with Sandefur). Also, this concept included the ability to know what methods and materials should be applied to different students and when each is appropriate or inappropriate; a throwback to the Progressives' desire for individualized education for every child.

Along with the Knowledge of Teaching Reading Test, the teachers also took a test of decoding skills and a measure to assess their higher level of comprehension skills. The assumption made was that a teacher who had trouble with determining letter-sound correspondence, for example, might be less effective in teaching that skill than a teacher with more skill in these areas. A possible cause of this could be that the teacher is less able to communicate the essential knowledge to the pupil or, because the teacher spends less instructional time on areas in which he/she feels less competent or less comfortable, thereby providing the pupils less opportunity to learn. Teachers also took a test of application of their knowledge of teaching reading to certain hypothetical situations. Phase II found that the most

likely model for the relationship between teacher aptitudes, knowledge, and performance is that there exists some minimal aptitude level which is necessary but not sufficient for the acquisition of knowledge, both about the subject matter area and about teaching skills, and that this knowledge

will produce adequate pupil acquisition of information and adequate teaching performance. The aptitudes specific to certain subjects are probably moderated by the teacher's communication techniques which are also probably partially dependent on other aptitudes, such as verbal facility. (p. 52)

ETS decided that there were practical as well as theoretical reasons why teachers aptitudes were considered an important part of the study. Since the ultimate outcome of the BTES was to suggest possible modifications in teacher licensing standards, it seemed logical and desirable to have variables that could be readily quantified for use as selection devices. BTES found that this approach followed a model used in industrial psychology whereby an analysis is made of the task or job being studied and the relationship of the task elements to known aptitudes and abilities. This approach allows the prospective employer both to select individuals on the basis of existing skills and also to select for training those individuals who have the greatest likelihood of being successful.

One of the most important results of the BTES was that the data showed that teachers do make a significant contribution to pupils' learning. In an analysis of the data, the factor most strongly related to a pupil's final scores were the initial scores, which was to be expected. In general, pupils who read better at the beginning of the year were still the better readers at the end of the year. The second most important

factor contributing to change was what the teachers did; the kinds of instruction and teaching that went on in their classrooms. Another important result was that no single skill or teaching practice was found to be effective at both of the grade levels used in this study or in teaching reading. The study also found that effective teaching patterns will differ by subject and by grade level. The study results supported the hypothesis that it is patterns of practices rather than single teaching practices or teaching skills which account for effectiveness.

The teachers in the study used a variety of ways in which to organize the pupils for instruction. The critical feature in organizing instruction seemed to be the degree to which the procedure facilitated managing the instructional process and maintaining a reasonable high level of pupil involvement. Appropriate material was not found to be sufficient. Patterns of classroom organization and teaching performance which provide more direct instruction to pupils by teachers improved instruction and pupil learning. The data suggested that an effective teacher practice was one in which the classroom organizational structures or the teaching performance provided for direct instruction. This was found to improve learning. Also, the study found that the teachers' knowledge of teaching methodology and of reading appeared to be related to their teaching performances.

The data also indicated that considerable time spent in organizing for instruction rather than in instructing is



an ineffective use of instructional time. It was recognized that a teacher must spend time in organizing and managing instruction, but when more time was spent in management, then there is less time for instruction. The BTES findings confirm what many educators have been writing about the teaching of reading; that it is a highly complex activity drawing on many areas of knowledge needed by teachers of reading.

In 1971, two professors of reading, A. Sterl Artley and Veralee B. Hardin from the University of Missouri, devised, developed, and published an "Inventory of Teacher Knowledge of Reading." Their Inventory was to be used for elementary school teachers and college students in methods courses. It was an evaluative devise for measuring reading knowledge but another of its stated purposes was to improve the teacher education program. The test was reviewed by Daniel T. Fishco in Reading World, May, 1973. Fishco found it to be a questionable tool for measuring teachers' knowledge of teaching reading because the items were either narrow or nebulous in meaning leading to doubtful validity.

There are other tests used to measure competency in knowledge of teaching reading available. The American College Testing Program (ACT) Proficiency Examination in Reading Instruction in the Elementary School (1973-76) for college accreditation of nontraditional study or for advanced placement or assessment of educational achievement was developed and administered in New York State. This was part of their College Proficiency Examination Program. Another test offered by ACT is a proficiency exam in Corrective Instruction in Reading.

Another test used for college seniors or adults as a means of assessing teaching reading competency is the National Teacher Examinations: Introduction to the Teaching of Reading. This test was developed in 1972-76. Jack C. Merwin, a reviewer of this test for The Eighth Mental Measurement Yearbook compiled by Buros, states that the technical data of interest to most users - data showing the extent to which scores can be used to identify potential success in teaching - is largely unavailable. He also stresses as does reviewer James V. Mitchell, that the National Teacher Examinations are "grossly misused" if they are used in any way to predict effective classroom teaching as conventionally measured.

Several years ago (1977-1979) the University of Houston undertook an evaluation of their teacher education program. In order to assess their program they developed and piloted an evaluation instrument entitled "Student Teacher Summative Performance Evaluation Instrument" or PEI. The categories covered by the instrument were: 1) planning and preparing for instruction, 2) performing instructional functions, 3) using student data, 4) relating interpersonally, 5) professionalism, and 6) related experiences. Many of these category names are the same types of behaviors mentioned by Sandefur and by the BTES. The one that differs somewhat is the last category, related experiences. This category covered other meaningful experiences thought to be related to good teaching behaviors that had been experienced by the teacher.

The test was given just prior to student teaching in

order to check teaching competence, to diagnose potential problems in teaching, and to provide useful information to student teacher supervisors. Concurrent studies of the PEI were made to ascertain if the scores obtained could validly and reliably predict high competent students from low competent students, as well as investigate reliability among the raters. According to a study done by Piper and O'Sullivan in 1981, it was concluded that the PEI could predict high and low competent student teachers. A change in the administration of the College of Education in 1979 discontinued the long-range research plans and use of the PEI.

The faculty, however, adapted in 1982, several competencies from this evaluation study for the Department of Education at Houston. These are: 1) identifies and teaches reading skills necessary for student understanding of given content areas, 2) assesses and alters the difficulty level of instructional materials, 3) diagnoses the reading abilities of students, 4) incorporates strategies for improving vocabulary and comprehension, and 5) incorporates strategies for improving study skills appropriate to given content areas. The wording of several of these competencies is different from prior studies but the teaching behaviors are similar. One noteworthy inclusion is the competency of diagnosing the reading abilities of students. This aspect of teaching reading is basic to any instructional program devised for students. Also worthy of note is the attention given now not only to skills, but to comprehension and also to study skills.

Mentioned earlier in the BTES, teachers' verbal ability had been identified as the most significant school variable in explaining student differences in achievement. Daniel S. Arnold (1977) states,

from the Coleman report, perhaps the most extensive attempt ever made to assess the nation's entire educational system, as well as from studies by Bowles (1969), Bowles and Levin (1968), Hanushek (1968), and Guthrie et al. (1969), this variable was identified. Guthrie (1970) suggests that we need to view teachers' verbal ability, however, as a proxy measure for a number of related skills and qualities, such as finding means to motivate students, adapting materials to students' ability levels, and communicating in ways which make the subject matter more understandable, rather than considering it a single factor. (p. 5)

From this same report, Arnold states that attempts to relate these factors to success in teaching often produces a lack of validity conclusion. He states,

perhaps instead of examining single criteria for their validity, we need to examine in greater depth the potential of a cluster of criteria. If, as has been suggested in a study of teacher education in Ohio (Ryan, Kleine, and Lrasno, 1972), criteria should be consistent with the goals of teacher preparation programs and should include a measure

of intellectual competence, ability to succeed in academic coursework, facility in dealing with children, and commitment to the teaching profession. (p. 5)

Pre-service teachers, according to Arnold, are most likely to be successful if they are above average in academic ability and preparation in their teaching field. Studies confirming this relationship mentioned by Arnold include

W. A. Skinner's (1947) report in which he reported 23 separate researches into the relationship between teaching performance and intelligence quotients of teachers. Although small, the correlations were positive. Hellfritzsch (1945), reporting on a factor analytic study, found general knowledge and mental ability combined to provide a positive predictor of performance. LaDuck (1945) subsequently noted that the intelligence of teachers correlated significantly with student gain scores. Using both tests of intelligence and tests of teacher knowledge in the content field, Rostker (1945) found similar correlations. Jones (1956) more recently discovered that 'good' and 'poor' teachers differed with respect to grade-point averages in professional courses and grade-point averages in courses in their majors. Teachers are more likely to be successful if they demonstrate flexibility and emotional stability

as character traits and if they manifest favorable attitudes toward children. (p. 20)

As a result of the many studies and surveys that have been done on positive interaction with children as a criteria for effective teaching, many teacher training institutions now require early experiences with children or adolescents as a condition for admission to the professional program. Morrison and Austin (1977) set up a model college program in which they suggest students at the sophomore year be involved in a variety of experiences with children in a number of different settings. Frequent contacts with children in classrooms would help these students adjust with greater confidence and competence in their initial teaching assignments. The State of Florida, for example, passed legislation in 1978 requiring its institutions to require early field experiences with children.

Teacher preparation institutions must accept a very large portion of the responsibility for product characteristics upon releasing teachers as certifiable and employable. Many faculties at colleges and universities object to all formal evaluation although they are continuously assessed informally by students, fellow professors, and the public. Some faculty members believe that what goes on in the classroom is too complex to evaluate. And given the weakness of current assessment methodologies, this argument may have considerable validity. On the other hand, evaluation may force individuals to face up to professional deficiencies. Public outcry is

demanding that educators in schools and colleges document the effectiveness of their programs as a basis for justifying continued or expanded support. The press for accountability leads quite naturally to a concern for programs that make a difference in the students. Whether or not teacher training institutions choose to adopt Competency Based Education as the whole program for its teacher preparation efforts, the essential elements of CBE seem to be a minimal requirement for goal setting. Among many professional educators who have studied the practice of teacher competency examinations, J. T. Sandefur (1983) concluded that very little testing of teachers was done prior to 1977. However, as of the Spring of 1983, twenty-eight states test or plan to test teacher candidates prior to the issuance of a certificate. The purpose of this exit examination is to ensure that those who enter the teaching profession have the basic knowledge in those areas considered fundamental in effective teaching. Such tests cover basic skills, professional knowledge (of teaching), and academic proficiency (subject-area specialization).

Arnold (1977) cites five necessary elements in establishing goals and determining assessment procedures devised by Elan (1975). They are:

1. Competencies to be demonstrated by the student should be derived from teacher roles, stated so as to make evaluation possible.
2. Criteria utilized to evaluate these competencies must be explicit in outlining areas of mastery.

and must be known by student and evaluator alike.

3. Assessment of the competencies should be objective and based upon student performance.
4. Demonstrated competency, rather than time or course completion determines the student's rate of progress.
5. The purpose of the instructional program is to facilitate development and evaluation of the student's achievement of the specified competencies. (p. 33-34)

In teacher education programs, the treatment usually takes the form of courses or other special experiences provided for the students. Since most teacher education programs have large enrollments, courses are offered in several sections or settings. Differences among instructional staff and environments create varying experiences for the students. Thus the competencies deemed necessary for completion of the program should be decided upon by the total faculty who comprise the teaching staff of the program.

In a program reviewed by Schalock (1976), the Oregon College of Education found that prospective teachers need not have the same subject matter base nor engage in the same practice teaching experiences to become effective teachers.

A great deal of latitude is provided students to negotiate the particular sets of knowledges (sic)



and skills they think are appropriate to prepare them for the context within which they wish to teach. Some non-negotiable knowledges and skills are required within the program (these would be the competencies the total staff agrees are necessary), but these constitute no more than a third or so of those that students going through the program are expected to obtain. (p. 25)

Such knowledge and skills are obtained through a wide variety of learning experiences. Schalock continues by stating, the linkage of knowledges and skills to be mastered to competencies to be demonstrated and the continuous testing of students as to the effectiveness of that linkage provides a level of structure and guidance to the curriculum refinement process that is immensely helpful. (p. 26)

The program was implemented at the Oregon College of Education on an experimental basis with 50 students in 1972. It was adapted the following year by the faculty within the elementary division of the college as the single elementary teacher preparation program to be offered by the college. In 1973-74, the college received the AACTE award for outstanding teacher education program in the nation.

Other noteworthy attempts of successful reforms are activities going on in Florida and Oklahoma. Florida passed legislation in 1978 and in subsequent years to establish several new components to its teacher education system. Early

field experiences and improved admissions, counseling and retention policies must be provided by the state's institutions. Florida also requires a score at the 40th percentile in basic skills on a nationally standardized college entrance examination as a minimum requirement for students entering a teacher education program. Teacher candidates must show demonstrable mastery of essential generic and specialized teaching competencies on a written examination before certification. The examination specifically covers competency in reading. At least 80 percent of the graduates of a teacher education program must pass the Florida Teacher Certification Examination or the college will lose its approved program status. The internship/beginning teacher program, effective July, 1982, is another important component of the Florida system. Successful completion of a one-year beginning teacher program is a requirement for the initial regular teaching certificate. This program provides, among other things, a verification of satisfactory performance of state-identified generic teaching competencies. Formative (for spotting specific weaknesses) and summative (for identifying immediate needs) evaluation instruments are being developed.

Oklahoma legislation, effective in February, 1982, called for 1) raising the standards for admission to colleges of education, 2) requiring competency examinations in subject areas prior to graduation, 3) mandated an entry-year internship prior to certification, 4) monitoring the beginning teacher's performance by a team representing the profession,

and 5) providing for the continuing education of teachers and teacher educators. Oklahoma also requires competency in oral and written English, a minimum GPA for entrance into a teacher education program and successfully passing an examination appropriate to their teaching specialty or level.

"Guidelines For The Professional Preparation of Reading Teachers," a pamphlet developed by the Professional Standards and Ethics Committee of the International Reading Association in 1978, recommended that all teachers, as well as reading personnel, should be prepared in reading education as it relates to their particular roles, their specific attitudes, concepts and skills. In particular they cite as demonstrable, cognitive capabilities for primary and intermediate grades the following:

1. Language Foundations For Reading

A. The English language as a communication system

1. Understands the concept of language as a symbol for transmitting ideas, information, and feelings.
2. Understands phonemic structure, morphemic structure, semantics, syntactic and intonation patterns.
3. Understands the historical development and the patterns of change in the English language.
4. Understands American English dialects and usage styles.

5. Understands major theories concerned with the processes of language use, especially the process of reading, and the nature of language development.

B. Language Development

1. Understands the principles and stages of physical, socioemotional, and intellectual development.
2. Understands the importance of oral language development as a foundation for beginning reading instruction.
3. Understands the relationship between continuous language development and readiness for reading achievement at any level.

2. Comprehension

A. Literal and interpretive comprehension

1. Develops a functional understanding of comprehension processes.

B. Critical comprehension

1. Develops a functional understanding of the nature of critical comprehension.

3. Word Analysis

- A. Demonstrates an understanding of the inter-relatedness of word analysis skills and comprehension skills and of the limitations of word analysis in isolation.
- b. Understands that a variety of word analysis strategies is required to meet the needs of

individual learners.

- C. Understands the difference in procedures for teaching word analysis among the various approaches to reading instruction.

#### 4. Enjoyment of Reading

- A. Is familiar with appropriate materials for the enjoyment of reading of students at the level taught.
- B. Is familiar with materials designed to appeal to the reading-handicapped learner.

#### 5. Diagnostic Teaching

##### A. Diagnostic evaluation

- 1. Understands the various methods for estimating achievement level and their limitations.
- 2. Understands the characteristics, strengths, and weaknesses of different types of measurement devices.

##### B. Organizing school and classroom for diagnostic teaching

- 1. Understands the extent of student variability in language aptitude, experience, and skills attainment.

##### C. Adapting instruction to needs of students with varied linguistic backgrounds

- 1. Understands the difference between oral reading errors and those miscues which

are a result of speech variations related to special linguistic backgrounds.

- D. Instruction of students with special reading needs
  - 1. Understands the nature and causes of reading/learning disability.
- 6. Program Planning and Improvement
  - A. Interaction with parents/community
    - 1. Knows the influences that social, economic, and cultural patterns have on language development.
  - B. Instructional planning: curriculum and approaches
    - 1. Has broad familiarity with developmental, supplemental, nonprint, and library materials useful in teaching reading.
    - 2. Understands several instructional approaches including group instruction in a developmental reading series, individualized self-selected reading, individualized systems-management approach, language experience approach. (p. 2-4)

Furthermore, the American Association of Colleges for Teacher Education, in their series of position papers on Educating A Profession, published in 1983, state that beginning teacher competencies, as expected by parents and educators alike, should include:

- 1. Care about children and are committed to their education.

2. Are broadly educated in the liberal arts and capable of engaging the young in the knowledge and intellectual processes fundamental to a common general education.
3. Are knowledgeable about the subjects that they teach.
4. Possess a thorough grasp of the knowledge base undergirding teaching practice, a repertoire of instructional strategies, and the skills to apply these to the education of individual students.
5. Are able to transcend their own personal experiences in the classroom as a student, and subsequently as a teacher, in order to make instructional decisions based on professional knowledge.
6. Are able to understand and use new technologies.
7. Are able to communicate effectively with children, parents, policy-and decision-makers, and other persons.
8. Understand and use methods of inquiry and research findings in making professional decisions.
9. Honor the dignity and rights of every individual learner consonant with the values of our democratic and pluralistic society.
10. Are prepared to implement the concept that

schools should provide open and equal educational opportunities for all.

11. Collectively represent the cultural differences associated with the society at large. (p. 3)

These qualities expected of teachers must be developed through programs offered at teacher training institutions. The public and others in the profession are insisting that these teacher training institutions establish criteria which ensure that graduates are at a competent level to begin professional teaching. Teacher training institutions must, according to the AACTE (1983),

- employ an assessment system which provides relevant, timely and recurrent measures of the prospective teachers' progress toward becoming effective teachers. The assessment system should be appropriate to each institution's organization of its programs...teacher education institutions have an obligation to establish entry, retention and exit criteria which measure the knowledge, skills, and attitudes of successful teachers. (p. 2-3)

#### Recent Proposals Calling For Change

Many reform proposals, addressing the major issue of how to provide for quality in teacher education, include as one of their specific areas in need of reform, the preparation of teachers; including program entrance and exit requirements. Some of the reform proposals that Sikula and



Roth (1984) cite follow.

1. In 1983, the National Commission of Excellence in Education published "A Nation At Risk: The Imperative for Educational Reform" Some of the recommendations listed for institutions pertinent to teachers of reading are 1) adopt more rigorous and measurable standards and higher expectations for academic performance, 2) raise admission requirements, 3) require demonstratable aptitude for teaching and demonstratable competence in an academic discipline.
2. "The Need For Quality" issued in 1981 by the Southern Regional Educational Board (a consortium of 14 Southern states) recommended that states should develop tighter standards for admission into teacher preparation programs, conduct performance evaluations of all beginning teachers, and require tests for certification with specific passing scores.
3. In 1983, the Council of Chief State School Officers (CCSSO) Ad Hoc Committee on Teacher Certification, Preparation, and Accreditation submitted a report to the Council. Among the recommendations it made were: 1) teacher training standards should be strengthened through state program approval, 2) there is a need for preparation programs to be balanced - current programs are overgeneralized for elementary teachers and overspecialized for secondary teachers, 3) extend teacher training programs with the increased time being devoted to learning pedagogy and self-evaluation, 4) competency-based programs should be appraised to improve programs.

4. In 1982 the NEA issued "Excellence In Our Schools, Teacher Education: An Action Plan." This plan called for more rigorous criteria for admission and graduation from teacher education programs, programs should provide more opportunities for practice, programs should be approved on the basis of standards derived from the practice of teaching, the state should review program content to ensure that beginning teachers have the capability to begin practice within that state. (Kansas is one of ten states working with NEA to implement these standards. To date, efforts are only in the initial stages.)

5. The American Federation of Teachers (AFT) submitted a proposal in 1981 before the House Sub-committee on Post-secondary Education calling for 1) improving the competency of the teaching force, 2) improving teacher education programs, 3) improving teacher certification and licensure, 4) developing entrance standards for teacher preparation programs and enforcing them, 5) using a written examination that tests the level of literacy, knowledge of subject matter and pedagogy, to qualify for entrance into the profession, and 6) need for a one-to-two-year internship prior to permanent certification.

6. The Education Commission of the States (ECS), a nonprofit organization comprised of governors, state legislators and educators, issued a 1982 report, "Raising Teacher Quality Levels," which recommended tougher admission standards for teacher education, a screen or test before certification, and an internship prior to certification.

7. Another prominent name in education, Educational Testing

Service (ETS), an organization active in teacher testing through the National Teachers Examination, published "Teacher Competence" in a 1982 issue of Focus. Among other conclusions reached, they concluded that the crisis of competence was rooted in the way teachers are trained. They suggested that five-year programs and other quality innovations may attract brighter students and a more rigorous teacher preparation program would draw quality students.

### Summary

The review of the literature concerning evaluation of teacher education and the need for evaluation of specific teaching areas, such as the teaching of reading, suggests that teachers of reading need to demonstrate knowledge in eight broad categories. These are:

1. Vocabulary and word attack skills
2. Reading comprehension
3. Diagnostic teaching for grouping/individualization
4. Organizing for instruction
5. Instructional planning
6. Evaluation of the reading process
7. Exceptional children/linguistically different learner
8. Discipline, interaction with parents and community

These knowledges can be demonstrated cognitively and can be analyzed for strengths and weaknesses in students' knowledge and in institutions' programs.

The defining process of behaviors necessary for effective

teaching has been honed over the years and it was this process that makes possible the creation of an assessment instrument to measure pre-service students' knowledge of teaching reading in the elementary grades. The time for testing of beginning teachers in basic skills and their teaching specialties has arrived. The demand for excellence and accountability has shown a need for institutions to carefully screen their students before exit from their teacher preparation programs. Prior tests, inventories and evaluation instruments have pioneered in this area but have not fully met today's need. Some of the instruments were created for specific instances, some are for private use by the creating agency, while others do not fully meet the criteria for testing knowledge of teaching reading in elementary grades. A need for a test measuring pre-service elementary teachers knowledge of teaching reading in elementary grades has been shown to exist.

Chapter 3 will describe the creation of the test instrument and method of analyzing data from a Pilot Study.

### CHAPTER III METHODOLOGY

The major purpose of this study was to construct a Knowledge of Teaching Reading in Elementary Grades test and then to investigate the results of giving the test in a Pilot Study in order to measure the reliability and validity of the test.

The research conducted in this study is Ex Post Facto research. A post-test only will be given to intact groups. Since this study concerns an existing program's effectiveness, the research design will be to measure reading knowledge after completion of the selected institution's stated coursework for pre-service elementary education teachers through the use of the Knowledge of Teaching Reading in the Elementary Grades test (also called KTREG). At this time Kansas State University and other existing institutions' treatment may vary from each other but all institutions present coursework acceptable for teacher certification in the State of Kansas. Therefore, although treatment will not be uniformly the same for all students, their over all general treatment meets the criteria for taking the KTREG test.

#### Procedure For Developing The Test

Sets of items were developed by several professors who teach reading methods courses as well as by the researcher. These items covered the eight broad categories suggested in the review of the literature. These in turn were simplified

into three categories of knowledge; Skills, Comprehension, and Theory/Research.

Items were constructed from course and field work materials presented to elementary education teachers of reading at some time during their teacher training program. Test items were constructed as multiple choice items with four answer choices. For Pilot Study purposes, approximately even amounts of questions pertaining to the three categories of knowledge were included on the test. The Theory/Research category received a double amount since there is a great amount of overlapping in this area. (Appendix A contains the first version of the Knowledge of Teaching Reading in Elementary Grades test.)

To establish face validity, a working copy of the test was submitted to a panel of reading professors for scrutiny. Their comments and suggestions were used to reword, add, or delete stems and foils. In addition, the test items were independently classified to the three categories of competencies by a panel of reading experts (different from the panel who looked at the test for face validity) to further establish the validity of the test items. (See Appendix B) An inter-rater analysis of the item classifications was used to determine agreement among raters' classified test items.

Internal reliability was measured using the Kuder-Richardson formula. This formula provides an estimate of the average correlation which might be obtained if all possible split-half combinations of a group of items were correlated

with one another. The split-test method provides an estimate of the extent to which a test is internally consistent, and permits a reliability coefficient to be obtained from a single administration of the test.

### Pilot Study

To ensure that the research conclusions to be drawn would be statistically valid, Cohen's (1977) suggested initial strategy of not less than 15-25 subjects per group minimum was used to form the sample populations for the Pilot Study. The first subjects group used to pilot test the KTREG test was comprised of 30 pre-service elementary teachers who had completed a methods course in elementary school reading. This sample group was an intact group and formed one of the Fall, 1983 reading methods classes at Kansas State University. The subjects took the KTREG test as part of their final examination. Answers to test items were marked on IBM cards to facilitate scoring. Subjects were told they were taking part in a study being conducted by a graduate student for a Master's Thesis. If they did not wish to participate they were given this option. If they wished data results they could so indicate to the professor in charge of the methods course and arrangements would be made to get a copy of the results to them.

The data from the first pilot test was analyzed for internal reliability using the Kuder-Richarson formula. The difficulty and discrimination indexes for each test item were analyzed to make stem and/or foil changes or to delete

the item entirely. Suggestions made by the reading professors who scrutinized the test for face validity were also incorporated into stem and/or foil changes. Some new items were written to replace deleted items and to add length to the test and equalize the three areas of knowledge. These items (99) were then submitted to a panel of reading experts to be classified to three categories of knowledge. These 99 items, plus five additional items written specifically for the comprehension category of knowledge (for a total of 104 items) comprise the second pilot study test. (See Appendix C)

The second pilot test was administered to an intact group of 24 pre-service elementary teachers who had completed a reading methods course during the Spring Semester, 1984. Answers were again marked on IBM cards to facilitate scoring. The test was administered by the professor in charge of the methods course as part of the final examination. Students were given the same explanation and participation choice as the first pilot test group. Data was analyzed for internal reliability using the Kuder-Richardson formula. The difficulty and discrimination indexes for each test item were analyzed to suggest further revisions in order to make a more valid and reliable test.

#### Method of Analyzing Data

The data gathered from the panel of reading experts, who classified items to categories of knowledge, was analyzed by comparing percentage agreements to show inter-rater agreement; a measure of test validity. The data from both pilot tests,



in the form of marked IBM cards, was processed by the Kansas State University Computer Center using the Grader-Roster Program developed by the Computer Center. This program allowed for a print-out to be made of each student's identification number, total scores, and individual responses to each test item. In addition, the program allowed for a computer print-out of the number of correct responses for each item, the percent of students who scored it correctly (difficulty index), the item discrimination index and student responses ( a display of the number of students selecting each option or alternative for each item). Finally, this program printed out data on the frequency of the score, percentile rank, Z-score, T-score, Kuder-Richardson coefficient, mean, median, and 1st and 3rd quartiles, standard deviation, and the standard error.

The Pilot Study test data was analyzed for reliability using the Kuder-Richardson formula and for validity through an analysis of the difficulty index and the discrimination index for test items as well as the analysis of the data gathered from the panel of reading experts on item classification. The second pilot test was also analyzed by individual reading knowledge category for reliability and validity using the same measures that were used for the Pilot Study as a whole.

## CHAPTER IV

### RESULTS

#### Overview

The Knowledge of Teaching Reading in Elementary Grades test is basically a criterion-referenced test; that is the items comprising the test focus on three discrete categories of knowledges the review of the literature suggested as necessary for beginning teachers to have mastered before they are released to intern or beginning teacher programs and certification. Popham (1981) describes a criterion-referenced test as a measure used to ascertain an individual's status with respect to a defined behavioral domain. The three categories of knowledge - skills, comprehension, and theory/research - are major, defined behavioral domains within the larger domain, teaching reading in elementary grades. They lend themselves readily to a measure of cognitive ability. Scoring of a criterion-referenced test can readily show an examinee's weaknesses and strengths within the whole test and in the discrete categories.

The purpose of the Knowledge of Teaching Reading in Elementary Grades test is to identify those examinees who demonstrate they possess the requisite knowledge for beginning teachers of reading. Only those examinees who show mastery of the requisite knowledge should be allowed to enter into the profession. Conversely, those students who show, by virtue of other assessment devices, high potential, should profit from the Knowledge of Teaching Reading in Elementary Grades test and

remediate deficit areas. This could be done during the intern year with re-examination at a fixed future time prior to certification.

On another level, a criterion-referenced test can be used to appraise the worth of an instructional program. Popham (1981) lists three positive factors for program evaluation: 1) a criterion-referenced test can be specific enough to match an instructional program's curricular emphasis, 2) it can provide information for instructional improvement, and 3) it is more "instructionally sensitive."

## Results

### Pilot Test I

The first Knowledge of Teaching Reading in Elementary Grades pilot test was comprised of 76 items and was administered to 30 pre-service elementary teachers from Kansas State University at the end of the Fall Semester, 1983 as part of their final exam in a reading methods course.

The Kuder-Richardson formula was used as a measure of internal reliability. The reliability coefficient was .73. Popham (1981) states that criterion-referenced tests, by virtue of their content construction, will have lower than the .80 to .90 range of a norm-referenced test.

Table 1 shows that the test scores had a range of 23 points from 45 to 68 out of a possible 76 with a standard error of measurement of 2.82. The mean score was 58.40. The standard deviation was 5.39. Test scores tended to cluster less than

one standard deviation from the mean with the exception of the top four scores and the bottom four scores.

Table 1

Comparison of Pilot Study Test Data

Test Data	KTREG I	KTREG II
Reliability		
Coefficient	.73	.77
Total Examinees	30	24
Number of Items	76	104
Score Ranges	45-68	63-89
Total Group Scores-		
Mean	58.40	77.46
Total Group Scores-		
Median	59.00	79.50
Total Group Scores-		
Standard Deviation	5.39	7.96
Standard Error	2.82	3.85

An item analysis of the difficulty index and the discrimination index was undertaken as a measure of the validity of the test. (Appendix D) Item analysis techniques help to improve tests and provide thorough data on class achievement on the concept measured by the item. Item analysis is also used to discover defective and ineffective items so they can be

reworded or replaced for future use. Scannell and Tracy (1975) suggest, as do many other test and measurement specialists, that the highest scoring 27 percent of the examinees and the lowest 27 percent of the examinees be used in the item analysis process.

Scannell and Tracy also differentiate between the meaning of item difficulty for a criterion-referenced or mastery test and that for a discriminatory test. A distinction between those who pass and those who fail is important. On a criterion-referenced test the knowledges to be measured are generally very basic and professors would expect a rather high proportion of a class to have mastered them. For criterion-referenced tests they feel that the ideal difficulty would probably range from .80 to 1.00. For the criterion-referenced test, they stress, items may be easy or difficult. They go on to state,

There is no statistical basis for an ideal difficulty level since maximizing variation among examinees is not relevant to the purpose of the test...the proportion of examinees who pass the item indicated the degree to which the concept measured by the item has been acquired and whether remedial teaching is needed.

(p. 221)

The desired difficulty level can be set by a professor, an institution, or a state board of education.

Disney (1971) does not differentiate among the types of tests in his discussion of the difficulty index. He suggests that "keepers" are those items with a p value between .20 and .80; although items with values outside those ranges are

sometimes retained on the basis of uniqueness of content. Popham (1981) concurs but goes one step further in saying that the actual ease or difficulty of an item is tied to the instructional program surrounding it. He poses the question, does a p value of .95 really indicate that the item is intrinsically easy or does it mean that the instructor did a first class job of teaching?

The difficulty index data for Pilot Test I (Table 2) shows that of the 76 items, 27 of the items were in the p value range of .20 to .80; 5 were below .20 and 44 were above .80.

Table 2

Difficulty Index Comparisons


---

Difficulty Index	KTREG I		KTREG II	
	#	%	#	%
above .80	44*	59	51**	49
.20 to .80	27	35	50	48
below .20	<u>5</u>	<u>6</u>	<u>3</u>	<u>3</u>
Total	76	100	104	100

---

\* of these, 1 item was at .83.3 and 5 items were at 86.7

\*\* of these, 10 items were at .83.3

The discrimination index also provides a basis for evaluating items and viewing validity at the item level. Disney (1971) suggests that the ideal discrimination value for an item is +1.00; however, in practice, values of 0.40 and higher are usually quite acceptable. He makes note that there may be compelling arguments for retaining items which should otherwise be discarded due to their low value. Popham (1981), in discussing item discrimination, refers to Robert L. Ebel's (Essentials of Educational Measurement, 1979) experienced-based guideline for indicating the quality of norm-referenced test items;

.19 and below	poor items to be rejected or improved by revision
.20 - .29	marginal items, usually needing and being subject to improvement
.30 - .39	reasonably good but possibly subject to improvement
.40 and above	very good items(p. 298)

These parameters took several decades to construct according to Disney(1971). Criterion-referenced test discrimination indexes, according to Popham (1981), still show a certain amount of indeterminacy. Kansas State University's Grader-Roster Program suggests that the minimum acceptable discrimination index level should be around .20 for all but extremely easy or difficult items.

Scannell and Tracy (1975) state that,

In general, an item which tends to be passed by students

who have a good command of the objective measured by the total test and which tends to be failed by students who do poorly on the total test may be considered to contribute in a positive way to the ability of the test to discriminate accurately among different levels of student performance, and then to the reliability of the test scores. (p. 223-224)

They feel that criterion-referenced tests should have lower discrimination indexes than a discriminatory or norm-referenced test.

The discrimination index data for Pilot Test I in Table 3 shows that of the 76 items, 30 items were .20 or higher, 18 items were at .30 or higher, and 6 items at .40 or higher. There were 46 items that fell below the minimum discrimination level as rated against a norm-referenced discrimination index.

Table 3

Discrimination Index Comparisons

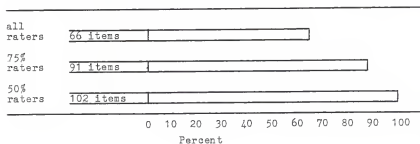
Discrimination Index	KTREG I		KTREG II	
	#	%	#	%
negative discrimination at 0.00	5	7	11	10.5
0.00 to .19	24	32	21	20.5
.20 to .29	17*	22	32**	30
.30 to .39	12	15.5	0	0
.40 and above	12	15.5	25	24.5
Total	<u>6</u>	<u>8</u>	<u>15</u>	<u>14.5</u>
	76	100	104	100
* 17 items at .12 level	** 32 items at .17 level			



The content validity of a criterion-referenced test is very important. As stated in the chapter dealing with Methodology, sets of items were written by a panel of reading professors and the author of this study. The items were submitted to a panel of reading professors who ascertained that the items covered material dealing with the teaching of reading in elementary grades. Suggestions made by them were incorporated into the test items. After Pilot Test I was given, an item analysis of the difficulty and discrimination index levels indicated which items were in need of revision. This was undertaken in an effort to improve the validity of the test.

Concurrently, a classification sheet with directions, the first KTRREG test as well as a set of 26 new questions and an accompanying cover letter (Appendix B) were sent to a different panel of reading experts so that an independent analysis could be made of the items classifications they felt each of the items belonged. The data received from the panel of reading professors who classified the test items to the categories of knowledges is shown in Table 4.

Table 4

Inter-rater Agreement on Knowledge Areas

All raters agreed on 65 percent of the item classifications (66 out of 102 items). Three-fourths of the raters agreed on 90 percent of the item classifications (91 out of 102 items).

### Pilot Test II

The second Knowledge of Reading in Elementary Grades test was comprised of 104 items (made up of revised items from Pilot Test I and new items) and was administered to 24 pre-service elementary teachers from Kansas State University at the end of the Spring Semester, 1984 as part of their final exam in a reading methods course.

The Kuder-Richardson formula was used to measure internal reliability. The reliability coefficient was .77, a gain of .04 from the first pilot test. The test scores ranged 26 points from 63 to 89 out of 104 items (See Table 1) with a standard error of 3.85. The mean score was 77.46. The standard deviation was 7.96. Test scores tended to cluster less than one standard deviation from the mean with the exception of the top four scores and the bottom four scores.

The difficulty index data from Pilot Test II (Table 2) shows that of 104 items, 50 of the items were in the p value range of .20 to .80; 3 were below .20, and 51 were above .80. The discrimination index data for Pilot Test II (Table 3) shows that of the 104 items, 40 items were .20 or higher; 40 items were at .30 or higher, and 15 items at .40 or higher. There were 64 items that fell below the minimum discrimination level as rated against a norm-referenced discrimination index. Of these 64 items, 32 were at the .17 level.

The three categories of knowledge of the second Pilot Test were subjected to the same evaluation process as the two Pilot Tests as wholes. Difficulty index and discrimination index data on the three categories of knowledge are show in Tables 5, 6, and 7.

Table 5

Comparison of Areas of Knowledge, Pilot Test II Data

Test Data	Skills	Comprehension	Theory/Research
Reliability			
Coefficient	.25	.62	.62
Total Examinees	24	24	24
Number of Items	23	24	57
Score Ranges	13-20	14-23	35-50
Total Group Scores-			
Mean	16.58	19.00	41.87
Total Group Scores-			
Median	16.50	19.50	42.25
Total Group Scores-			
Standard Deviation	2.12	2.70	4.76
Standard Error	1.85	1.67	2.93

Table 5 shows that of the three areas of knowledge, Skills had the lowest reliability. This area of knowledge was comprised of discrete subskills that are not in themselves homogeneous; yet the subskills as a total are basic to being competent as a teacher of reading.

Table 6 (Difficulty Index Comparison of Areas of Knowledge from KTREG II) and Table 7 (Discrimination Index Comparison of Areas of Knowledge from KTREG II) show the results of the item analysis for the three areas of knowledge on the second Pilot Test.

Table 6

Difficulty Index Comparison of Areas of Knowledge from KTREG II

Difficulty Index	Skills		Comprehension		Theory/Research	
	#	%	#	%	#	%
above .80	11	48	16	65	23	42
.20 to .80	11	48	7	31	32	57
below .20	<u>1</u>	<u>4</u>	<u>1</u>	<u>4</u>	<u>1</u>	<u>1</u>
Total	23	100	24	100	57	100

Table 7

Discrimination Index Comparison of Areas of Knowledge from KTREG II

Discrimination Index	Skills		Comprehension		Theory/Research	
	#	%	#	%	#	%
negative discrimination	1	4	0	0	6	10.5
at 0.00	6	26.5	3	13	14	24.5
0.00 to .19	5	21.5	10	41.5	10	18
.20 to .29	0	0	0	0	0	0
.30 to .39	6	26.5	5	20.5	15	26
.40 and above	<u>5</u>	<u>21.5</u>	<u>6</u>	<u>25</u>	<u>12</u>	<u>21</u>
Total	23	100	24	100	57	100

## Summary

Test data was reported separately for each of the Pilot Tests on reliability and validity. Data of the three areas of knowledge from KTREG II were also reported. Inter-rater agreement on item classifications was shown to be high.

Chapter 5, Conclusions and Implications, will discuss these results. Suggestions for test improvement will be presented. Future use of the test will be discussed.

CHAPTER V  
CONCLUSIONS

This chapter is devoted to a summary of test reliability results, item analyses, and a discussion of further implications. The results of the study are reviewed and suggestions regarding further revision and testing are made.

The Knowledge of Teaching Reading in Elementary Grades test was created, evaluated, and then tested in a Pilot Study involving two testing situations. Results showed improved reliability and validity of the test between the two testing sessions.

Test Reliability Results

By examining Tables 2 (Difficult Index Comparisons) and 3 (Discrimination Index Comparisons) for validity and Table 1 (Test Data) for reliability, it can be seen that the result of this study is a reliable instrument for measuring beginning teacher's knowledge of teaching reading in the elementary grades. A comparison of the reliability coefficients between the two pilot tests showed an improvement between the two tests of .73 to .77. This can possibly be attributed to several factors: 1) lengthening of the test from 76 items to 104 items, 2) strengthening of the difficulty and discrimination indexes which led to 3) a larger standard deviation from the mean shown on the second pilot test. Compared to a similar test, "Knowledge of Teaching Reading" created by ETS and used in The Beginning Teacher Evaluation Study in California (1973-74,

pgs. 58-59), the Knowledge of Teaching Reading in Elementary Grades test appears to have equal or better test reliability.

### Item Analysis

Dizney (1971), Popham (1981), and Scannell and Tracy (1975) have reported that the content of the item should have the ultimate say in whether or not an item should be deleted due to poor p value and/or discrimination. Because the Knowledge of Teaching Reading in Elementary Grades test evaluates whether or not examinees have mastered the requisite knowledge concerning the teaching of reading, certain items, classified easy or hard according to parameters of the difficulty index or, not discriminating enough according to discrimination parameters, have been retained because of the content of the items. For example, in the second pilot test, test items 10 and 65 (see Appendix C) both have a difficulty index of 91.7 and a discrimination index of 0.00; yet the content of the test items is fundamental to the teaching of reading and therefore, it is necessary to know if examinees have grasped this material.

Table 2, showing Difficulty Index Comparisons, indicates that for KTREG I 44 percent of the items were acceptable (.20 to .86.7) while for KTREG II 58 percent (.20 to .83.3) were acceptable. This represents a gain of 14 percent. The remaining items are subject to qualification on the basis of their content.

Table 3, Discrimination Index Comparisons, indicates that

the acceptable item rate for KTREG I was approximately 40 percent (.20 to .40 or higher) while for KTREG II, it was approximately 70 percent (.17 to .40 or higher) or a gain of 30 percent. The remaining items are subject to qualification on the basis of content.

Since the individual item analyses for the three areas of knowledge were only incorporated into the second pilot test, a comparison of this data between tests could not be made. However, it is apparent from the data on the second pilot test that the Skills area of knowledge as well as the Theory/Research area of knowledge show greater amounts of items showing no discrimination; 6 and 14 respectively compared to only 3 items for Comprehension. This suggests that either the material in the item was well learned by all examinees (the professor did a good job of teaching) or the foils were not doing the job sufficiently.

Difficulty index data for the three areas of knowledge shows that Comprehension had the greatest percentage of "easy" items (classified against a norm-referenced test scale): 65 percent compared to 48 percent for Skills and 42 percent for Theory/Research. However, of the 51 items from the total test that were above .80 on the difficulty index, 10 of these items had a p value of .83.3 which is close enough to call acceptable. Of the 40 items that had a p value of above 83.3, 11 items had good discrimination indexes of .30 or higher. The remaining 29 items out of 104 items on the test are in need of additional work to stems and/or foils.



### Discussion

Discrete guidelines have been drawn for norm-referenced test item analysis whereas criterion-referenced test item analysis has not been defined in such a distinctive manner. Taking this into account, however, future work on this test should look into improving both the difficulty and the discrimination of the test items. One area that would help this would be an analysis of the distractors or foils provided for each item.

Another consideration for improving the test would be to upgrade the content validity of the test by continually appraising the items so that a group or pool of homogeneous items is formed that is consistent with the description of the behavior domain being tested. One issue not easily approached in doing this is the continuing dialogue between reading professors as to what constitutes the best wording for any given item. A future study could look into this problem of semantics and item construction.

Another avenue worthy of consideration would be to create an inventory beginning teachers could use to rate themselves on how they perceive their strengths and weaknesses in the area of teaching reading. This could be analyzed along with the Knowledge of Teaching Reading in Elementary Grades test data to give examinees as well as their supervisors/reading professors a clearer picture of themselves as elementary teachers of reading.

Future work on this test should provide a useful tool

in making decisions regarding the knowledge of pre-service elementary teachers of reading. Further studies should be undertaken using this test to ascertain how effectively pre-service elementary teacher are acquiring basic knowledge of the teaching of reading.

## References

- American Association of Colleges For Teacher Education. (1983). Educating A Profession: Profile Of A Beginning Teacher. Washington, D. C.
- Arnold, Daniel S., et al. (1977). Quality Control In Teacher Education: Some Policy Issues. Washington, D. C., American Association Of Colleges For Teacher Education.
- Carey, Lou M. (undated). State Level Teacher Performance Evaluation Policies. Syracuse: National Council of States on Inservice Education.
- Cohen, J. (1977). Statistical Power Analysis For The Behavioral Sciences. New York: Academic Press.
- Disney, Henry. (1971). Classroom Evaluation For Teachers. Dubuque: Wm. C. Brown Co.
- Ekstrom, Ruth B. (1975). Pupil and Teacher Tests, Final Report, Beginning Teacher Evaluation Study Phase II, 1973-74. Princeton: Educational Testing Service.
- Fults, Anna Carol, et al. (1972). Readings In Evaluation. Danville: The Interstate Printers & Publishers, Inc.
- Haberman, Martin, & Stinnett, T. M. (1973). Teacher Education And The New Profession Of Teaching. Berkeley: McCutchan Publishing Corp.
- Haefele, Donald. (1971). Systems And Modeling: Self-Renewal of Teacher Education. Washington, D. C.: American Association of Colleges For Teacher Education.

- Heilman, Arthur W. (1965). Effects Of An Intensive In-Service Program On Teachers' Classroom Behavior And Pupil's Reading Achievement. University Park: The Pennsylvania State University.
- International Reading Association. (1978). Guidelines For The Professional Preparation Of Reading Teachers. Newark.
- McDonald, Frederick J. (1975). Beginning Teacher Evaluation Study: Summary Report. Princeton: Educational Testing Service.
- Morrison, Coleman, & Austin, Mary C. (1977). The Torch Lighters Revisited. Newark: International Reading Association.
- National Association Of Secondary School Principals. (1974). Competency Tests And Graduation Requirements. Reston.
- Popham, W. James. (1981). Modern Educational Measurement. Englewood Cliffs: Prentice-Hall, Inc.
- Roane, Elinora. (1974). A Study Of Reading Professors' Perceptions Of Competencies Needed By Pre-Service Elementary Majors To Teach Reading To Culturally Different Black Children. Dissertation, Kansas State University.
- Sandefur, J. T. (1970). An Illustrated Model For The Evaluation Of Teacher Education Graduates. Washington, D. C.: American Association of Colleges For Teacher Education.
- Sandefur, J. T. (1983, June). Teacher Competency Testing: The Public's Mandate. Paper presented to the 55th Annual Conference of the National Association of State Directors of Teacher Education and Certification. San Francisco, CA.

- Scannell, Dale P. & Tracy, D. B. (1975). Testing And Measurement In The Classroom. Boston: Houghton Mifflin Co.
- Schallock, H. D., Kersh, B. Y., & Garrison, J. H. (1976). From Commitment To Practice: The Oregon College of Education Elementary Teacher Education Program. Washington D. C.: American Association of Colleges For Teacher Education.
- Sikula, John P., & Roth, Robert A. (1984). Teacher Preparation and Certification: The Call for Reform. Bloomington: Phi Delta Kappa Educational Foundation.
- Toburen, Don L. (1981). Attitudes Of Kansas Competency-Based Teacher Education Graduates Toward Their Teacher Preparation Programs Over A Prolonged Time Period. Dissertation, Kansas State University.

APPENDIX A

Pilot Test I

KNOWLEDGE OF TEACHING READING IN ELEMENTARY GRADES

1. Reading literature aloud to children fulfills the following reading readiness requirement
  - a. lets children know what they will learn next
  - b. develops experiential background
  - c. teaches phonetic reading skills
  - d. organizes cognitive skills in advance
2. The effect of a child's background of experience on the ability to recognize words is
  - a. visually very minor
  - b. likely to be substantial
  - c. greater for boys than for girls
  - d. greater for girls than for boys
3. Why is a child's oral language development highly important?
  - a. it is a basic ingredient for success in reading
  - b. it is symbolic and requires discipline to learn
  - c. it is a system of sounds common to our culture
  - d. it is conventional
4. How is readiness best accomplished?
  - a. by giving children meaningful experiences of all kinds
  - b. by having children complete a readiness workbook
  - c. by teaching specific skills through a sequence of exercises
  - d. by having children learn the alphabet
5. Which is not a desirable way for parents to facilitate reading readiness?
  - a. read to children frequently
  - b. converse with children to enlarge their vocabularies
  - c. provide a variety of experiences
  - d. make the child learn reading skills at home
6. When should systematic and sequential reading instruction begin?
  - a. at the end of kindergarten
  - b. after pupils have finished a reading readiness program
  - c. at different times for different children
  - d. no later than the first half of the first grade
7. Vowels are more consistent in their sounds than are
  - a. consonants
  - b. consonant blends
  - c. consonant digraphs
  - d. none of the above

8. In the word car the vowel sound is an example of a(n)  
a. long sound  
b. short sound  
c. "r" controlled sound  
d. hard sound
9. The vowel letters in the word neat are representative of  
a  
a. consonant digraph  
b. vowel digraph  
c. diphthong  
d. vowel blend
10. "Ou" as in house is a  
a. blend  
b. diphthong  
c. digraph  
d. phoneme
11. Ch as in machine is a  
a. blend  
b. diphthong  
c. digraph  
d. phonogram
12. The first two letters in the word dwarf are representative of a  
a. consonant digraph  
b. consonant diphthong  
c. diphthong  
d. consonant blend or cluster
13. Which group of nonsense words below would be representative of the hard "g" sound?  
a. gane, gobbb, guvot  
b. oge, gint, gish  
c. gripit, goot, gyt  
d. gyimb, gunt, goveret
14. What does every syllable contain?  
a. a consonant sound  
b. a vowel sound  
c. a digraph  
d. a diphthong
15. An important suggestion to remember in the teaching of phonics is  
a. give plenty of attention to letters out of context  
b. emphasize meanings of words  
c. give all children equal training  
d. teach phonics in a separate period



16. If children are to use phonics effectively,
- they should be able to recognize the word if they hear it
  - they should have mastered the basic phonics generalizations
  - they should be able to recognize the structure of the word
  - they should have training in phonics
17. Ms. Brown conscientiously teaches every phonics skills lesson in the basal reader series. This practice is
- a good model for others
  - based on student's needs
  - not necessary
  - supported by research
18. The word John's is
- the possessive form of John
  - the contracted form of John is
  - dependent upon the context in which the word is found
  - dependent upon whether you are referring to more than one John
19. How many phonemes does the word sit have?
- 1
  - 2
  - 3
  - 4
20. How many morphemes does the word unfairly contain?
- 1
  - 2
  - 3
  - 4
21. Which characteristic is not part of the first phase of a Directed Reading Lesson?
- reading silently
  - creating interest
  - introducing new vocabulary
  - developing concept background
22. Developing interest and background of experience in a Directed Reading Lesson should
- receive about the same amount of time in beginning reading as it does in intermediate grades
  - be limited to not more than one minute for most selections at any reading level
  - depend on the teacher's interest in the selection at any reading level
  - be determined by the needs of the students in relation to that selection

23. What are good types of words to teach as sight words?
- irregularly spelled words
  - regularly spelled words
  - words which are unknown to children
  - words of three or more syllables
24. Most readability formulas measure the difficulty of semantic and syntactic variables by tabulating
- the sentence length and number of syllables in a given selection
  - the number of words comprised of 10 or more letters in a given selection
  - the length and number of sentences in a given selection
  - the ratio of numbers of words to numbers of sentences
25. One application of the schema theory for teachers is
- accept the limitations of a student's background of experiences
  - provide instruction involving concept development relevant to reading selections
  - provide instruction in comprehension only in materials students are familiar with
  - teach the concepts of assimilation and accommodation
26. Teaching students to paraphrase is
- a practice which will likely improve their comprehension
  - a practice which most instructional materials include
  - too difficult until the intermediate grades
  - not likely to influence their comprehension of text
27. When checking a pupil's comprehension of material, the teacher should be most concerned about
- factual detail in the selection
  - asking questions consistent with the stated purpose for reading
  - the pupil's ability to evaluate the selection critically
  - how long it took the child to read the selection
28. Which is a true statement about reading comprehension?
- it takes less time to teach reading comprehension skills than word recognition skills
  - national test results seem to indicate that comprehension instruction is presently effective
  - all of the phases of the Directed Reading Lesson can contribute to comprehension improvement
  - basal reading materials provide an adequate source for development of comprehension

29. To evaluate pupils' general reading comprehension of a particular type of material, the elementary teacher should
- include several types of comprehension questions
  - use objective questions as often as possible
  - give a worksheet to all children in the class in order to assess each individual's comprehension
  - depend heavily upon standardized reading tests
30. Asking students to explain how they arrived at correct and incorrect answers to comprehension questions
- takes too much time
  - is distracting and confusing to students
  - enables the teacher to diagnose comprehension ability
  - is recommended in basal teacher manuals
31. Which is a true statement about preparing comprehension questions for a selection?
- ask questions that are all of the same type
  - make the questions as tricky as possible
  - ask more literal questions than interpretive questions
  - ask questions which require a variety of comprehension skills
32. Which of the following should not be the basis of comprehension questions to ask the child after reading a passage?
- factual items
  - inferential questions
  - vocabulary items
  - yes/no questions
33. Joe does not know the word bark. If the word appeared in the sentence, "The dog's bark was very loud," the most effective way of unlocking the meaning of the word would be through the use of
- phonic analysis
  - structural analysis
  - contextual analysis
  - the sight word method
34. Reading to find details is
- always essential for comprehension
  - usually neglected in reading instruction
  - often overemphasized in reading instruction
  - generally should be avoided

35. A cloze test is often
- a simple and effective technique for fitting textbooks to students
  - best used to determine the frustration levels of students
  - a poor technique to use with library books
  - too time consuming and complex for teachers to use in the classroom
36. To develop a Criterion Referenced Test one must have
- clearly defined objectives
  - norms
  - reliability and validity of at least .58
  - children who can read
37. A standardized reading test score of 7.6 means
- that the child with that score reads as well as most children in the sixth month of the seventh grade
  - that the child with that score reads as well as most children who are seven years and six months old
  - that the child with that score should be instructed from a seventh grade reading text
  - that the child with that score should be instructed from a book written for the average seven-year old child
38. In assessing children's reading abilities, their levels of reading achievement should be considered in relation to what?
- the norm for their grade levels
  - their chronological ages
  - their expectancy levels
  - the level of the basal used for instruction
39. The most accurate means of determining a child's instructional level is through the use of
- a cloze test
  - an informal reading inventory
  - an achievement test
  - a word list
40. The level for use in selecting basal readers or textbooks for instruction in reading is the
- instructional level
  - independent level
  - capacity level
  - frustration level
41. In order to effectively assist a youngster in the selection of a library book, a teacher should be most concerned with the child's
- understanding reading level
  - independent reading level
  - instructional reading level
  - frustration reading level

42. A comparison of the "fourth grade" geography book and a fourth grade basal reading will likely show
- the geography book to be written on an easier level
  - the reader on a lower level than the geography book
  - that the two are written on about the same level
  - that the two begin on the same level but the geography book ends on a more difficult level
43. The Guided Reading Procedure is most appropriately used with
- basal readers
  - content area textbooks
  - library books
  - language experience stories
44. Ms. Brown assigned reading materials based on standardized test scores. This practice is considered
- a good one
  - a poor one
  - as good as any other
  - the most effective strategy for placement
45. Which is a true statement about round-robin reading?
- it is an excellent teaching technique
  - it keeps better readers interested
  - it does not keep most students on task
  - it assures the attention of all children during the exercise
46. When a child does not know a word during oral reading the teacher should
- have the child sound it out
  - tell the child the word
  - tell the child to guess using context clues
  - let other children tell the word
47. The greatest disadvantage in using the Language Experience Approach is that
- the materials are expensive and not as interesting
  - it requires more planning to use a systematic approach to teaching skills
  - the vocabulary is too difficult for the children
  - teachers do not have adequate knowledge of reading skills to be effective
48. Language Experience stories can serve as a basis for teaching
- about differences in the way students talk
  - comprehension skills
  - word recognition skills
  - comprehension skills and word recognition skills

49. The approach that best builds on the relationship between oral language, background of experience, and reading is
- basal
  - programmed
  - language experience
  - interest
50. The factor that makes the most difference in successful reading instruction is
- teachers
  - grouping patterns
  - materials
  - methods of instruction
51. Teaching context clues
- has little effect on vocabulary development
  - has an important impact on vocabulary development
  - has an important role in the Directed Reading Activity
  - has an important role in the Guided Reading Procedure
52. Several parents of Mr. Brown's students have volunteered to help him in the classroom. He should
- politely discourage them because teaching reading is the teacher's job
  - immediately give them a teacher's manual and a reading group
  - encourage them in such activities as reading stories
  - limit their participation to grading papers
53. Teaching test taking skills is
- viewed unfavorably by ethical teachers
  - recommended to achieve accurate test results
  - a poor use of instructional time
  - known to produce inflated test results
54. The Directed Reading Activity seems the most consistent with which theory of reading
- bottom-up theory
  - top-down theory
  - interactive theory
  - none of these
55. The best suggestion to make to parents who ask what should be done at home to help promote interest in reading is for them to
- provide drill on word recognition skills
  - pressure their children to use their library cards
  - require equal reading time and television viewing time
  - provide a model by reading aloud to their children

56. A practice whose effectiveness has been demonstrated in promoting an interest in voluntary reading is
- the assignment of written book reports
  - sustained silent reading
  - oral reading during the Directed Reading Activity
  - bulletin board displays of book covers
57. Ms. Cohen is concerned about her students' oral reading skills and seeks advice from other teachers about good teaching practices. Which of the following is the best advice?
- have each child read aloud every day
  - teach the children about eye-voice span
  - have children read silently before reading aloud
  - teach the children to read with expression
58. Providing appropriate reading instruction for exceptional children in the regular classroom is
- impossible because they lack the prerequisites for reading
  - part of the teacher's responsibility of matching instruction to all learners
  - more effective if exceptional learners work independently in the classroom
  - more effective if the emphasis is on word recognition skills
59. The best conclusion to be drawn from the research on methods of teaching reading is that
- the most successful method is the one preferred by each individual teacher
  - individualized instruction produces the greatest achievement
  - there is no best method for teaching reading
  - the method recommended by the teaching materials works best
60. Teachers should instruct students that the rate at which they read a selection should be determined by
- the amount of time available for completing the assignment
  - their fastest reading rate and their interest
  - the rate of reading they use when reading in the basal reader
  - their purpose for reading and the difficulty of the material

61. In selecting books to read aloud to students, teachers should
- be concerned with the independent reading level of the students
  - never read books with characters stereotyped by sex, race, or ethnic group
  - read good literature but explain the presence of stereotyped characters
  - primarily read titles nominated for the William Allen White Award
62. An effective way to manage reading instruction in a classroom consisting of children from culturally different backgrounds is to
- form homogeneous cultural groups
  - use cooperative learning activities
  - use extensive whole group instruction
  - use totally individualized instruction
63. One instructional implication of a bottom-up model of reading is an emphasis on
- comprehension questions
  - skills instruction
  - language experience activities
  - sight words
64. A Directed Reading Activity (DRA) is usually
- a whole-class study skills technique
  - a process that can give a reading or content lesson more structure
  - an individualized four-step reading strategy
  - a small group, student-centered process to review previously read material
65. Diagnostic teaching
- is dependent upon a child's chronological age
  - is instruction based on a child's strengths and weaknesses
  - requires highly specialized training
  - should be used only by reading teachers in the self-contained classroom
66. Diagnostic information can be
- gathered only from standardized tests
  - gathered by the teacher through observation
  - gathered by the teacher only from teacher-made tests
  - recorded either on tape or video for accuracy



67. All but one of the following are examples of structural analysis. Which is that one?
- diphthongs
  - inflected forms
  - derived forms
  - compound words
68. Which of the following words does not lend itself easily to structural analysis?
- running
  - parties
  - slowly
  - gone
69. When the analysis of compound words has been over-emphasized, the pupil may
- spend too much time looking for little words in big words
  - simply divide all multi-syllabic words in the middle
  - depend too heavily on phonic analysis for the recognition of all big words
  - avoid the usage of context clues
70. Words that end in a, es, ed, ing, er, and est are called
- derived forms
  - phonemes
  - definitives
  - inflected forms
71. Unhappy, made from the prefix un and the base word happy, is
- a derived form
  - a phoneme
  - a definitive
  - an inflected form
72. If a child is insensitive to sound differences as they occur in the spoken words fan and can, he/she apparently needs assistance with which of the following?
- visual discrimination
  - auditory discrimination
  - visual acuity
  - auditory acuity
73. When checking a pupil's comprehension of material that has been read, the teacher should be most concerned about
- factual detail in the selection
  - asking questions consistent with the stated purpose for reading the selection
  - the pupil's ability to evaluate the selection critically
  - how long it took the child to read the selection

74. It is preferable to give children a choice of main ideas for selection rather than requiring the child to formulate one
- in primary grades
  - in all elementary grades
  - in upper elementary grades
  - in special reading classes
75. Survival skills for remedial readers should include
- initial consonants, inflectional endings, and contractions
  - sight words, contractions, and phonics principles
  - phonics principles, structural analyses, and context clues
  - sight words, initial consonants and context clues
76. A student may need instruction in structural analysis skills if he/she lacks mastery of
- recognition of consonant sounds and clusters, vowel sounds, vowel digraphs, and auditory blending
  - recognition of base words, prefixes, suffixes, syllabication, accents, compound words, and contractions
  - obtaining the meaning and pronunciation of an unknown word by examining the context in which it is located
  - word-recognition skills

## NEW QUESTIONS

1. Reading comprehension does not include
  - a. the ability to read rapidly
  - b. the ability to understand charts and graphs
  - c. the ability to use encyclopedias
  - d. the ability to survey a variety of materials
2. Children acquire reading skills
  - a. in much the same way that they acquire oral language
  - b. in a structured sequence of tasks
  - c. in an unstructured sequence of tasks
  - d. in much the same way that they acquire comprehension strategies
3. In order to have successful reading, students need
  - a. phonics skills and word attack skills
  - b. phonic instruction and comprehension skills
  - c. word attack skills and a sight vocabulary
  - d. a sight vocabulary and comprehension skills
4. Word attack and phonic skills should
  - a. begin each reading instruction lesson
  - b. be concentrated in the primary grades
  - c. be followed with written exercises
  - d. be followed with reading in context
5. Students who know words in context but not in isolation are
  - a. telling you they have small sight vocabularies
  - b. in need of structural analysis skills
  - c. in need of semantic and syntactic skills
  - d. telling you that context aids their reading
6. Students participate in comprehension
  - a. through the active process of predicting
  - b. through the passive process of silent reading
  - c. through reading familiar materials
  - d. through reading unfamiliar materials
7. In comprehension development two situations exist for students to answer questions. They are
  - a. locate and orally read
  - b. recall and locate
  - c. recall and summarize
  - d. locate and summarize

8. Student responses are increased and are more complete when
  - a. teachers decrease wait time between question and response
  - b. other students are allowed to give clues
  - c. teachers allow more wait time between question and response
  - d. other students are not allowed to give clues
  
9. Reading skills are most effectively developed and refined by
  - a. focusing on readers' weaknesses
  - b. giving individual instruction
  - c. giving whole group instruction
  - d. focusing on readers' strengths
  
10. One important fundamental principle to remember when selecting materials to teach comprehension strategies is
  - a. to be sure that the context is sufficiently meaningful for students to relate what they already know to the text information
  - b. to be sure students are forced to solely attend to words and sound symbol information
  - c. to be sure that the comprehension strategies being taught have not been modeled or guided by the teacher
  - d. to be sure that the context is new and different from what the students already know
  
11. Content area reading
  - a. uses the same methods of presentation that are used during reading instruction
  - b. uses a different method of presentation than is used during reading instruction
  - c. is not taught by elementary school teachers
  - d. can not be presented through the language experience method
  
12. Most low-achieving readers
  - a. face problems that are largely in their control
  - b. face problems that are largely out of their control
  - c. face a greater demand on them than average or above average readers
  - d. face receiving instruction from many rather than just one teacher

13. Teachers often face the pressure to satisfy others that all readers are reading on grade level. This expectation is
  - a. realistic
  - b. idealistic
  - c. unreasonable
  - d. reasonable
  
14. Studying word families as a means of increasing vocabulary knowledge is
  - a. appropriate for primary students
  - b. appropriate for upper elementary students
  - c. appropriate for all grades
  - d. not appropriate for all grades
  
15. When students encounter difficulty in learning to read, the reaction of
  - a. teachers and parents is crucial
  - b. parents is crucial
  - c. peers and parents is crucial
  - d. peers and teachers is crucial
  
16. Reading difficulties can be caused by
  - a. emotional disturbances or disabilities
  - b. physical disabilities
  - c. intellectual and emotional disabilities
  - d. intellectual, physical and emotional disabilities
  
17. If reading goals seem worth reaching, either because they are immediately interesting or because they serve some need, students
  - a. will direct their energies towards those ends without prodding
  - b. will often overlook important details in their efforts to complete the task
  - c. will become competitive with peers and reduce their time on task
  - d. will direct their attention towards those goals with teacher prodding
  
18. Elementary teachers who obtain more alert and enthusiastic participation are observed to be
  - a. strict and well-organized
  - b. considerate and well-organized
  - c. impersonal and goal orienting
  - d. indulgent and orderly

19. Periodic evaluation of each student's reading is primarily important
  - a. for teachers' grading reports
  - b. as a means of evaluating teaching effectiveness
  - c. as feedback to the learner
  - d. for determining regrouping of students
  
20. The most common cause for difficulty in learning is inadequate fundamental skills in
  - a. written expression
  - b. reading
  - c. arithmetic
  - d. ability to plan and organize
  
21. One goal of the elementary school is the development of the fundamental skills of reading, quantitative thinking, and communication. Most of the teacher's evidence of this
  - a. comes from seatwork, workbooks and written work
  - b. comes from the oral responses of students
  - c. comes from quizzes and tests
  - d. comes from observation of daily performance
  
22. Older elementary students may find it beneficial to follow a specific study technique in the reading of content area materials. The most well-known of these is the
  - a. D-R-T-A
  - b. IEP
  - c. SQ3R
  - d. SRA
  
23. Reluctant readers are
  - a. those students who never volunteer to read orally
  - b. those students who have the necessary skills but don't use them
  - c. those students who never have the correct place when it is their turn to read orally
  - d. those students who can not answer comprehension questions
  
24. Which one is not a basic concept essential to good educational practices?
  - a. the family is not an important institution in education
  - b. all children are capable of being educated
  - c. teaching and learning should be a planned systematic process
  - d. everyone needs to learn how to read and write

25. One way not to grow in your profession as a teacher is to
- be willing to improve skills and increase knowledge
  - be willing to try new things
  - be willing to maintain an open-door policy in teaching
  - be willing to let seasoned teachers create your daily teaching schedule
26. Comprehension questions requiring pupil reactions, insights, judgements, or reflections before answering are
- highly effective
  - highly reflective
  - low in effectiveness
  - low in content

APPENDIX B

Inter-rater Item Classification



April 9, 1984

Dear

Thank you for participating in the classification of the test items into content areas. By doing this, you will help me with the validity of placing test items into particular content areas. I will tally all respondents' classifications. Hopefully, they will all agree. It promises to be interesting.

Enclosed you will find a Classification Sheet that has the content area definitions to help you with your item classifications.

Also enclosed is a copy of the test. It is in two sections. Just list the item number under the correct heading and that will differentiate between the two sections.

I appreciate your help. If you would complete this classification right away and mail it back to me in the enclosed stamped, addressed envelope, I will be most grateful.

Sincerely,

Enclosures: Classification Sheet  
Test (two sections)  
Stamped, addressed envelope

## CLASSIFICATION SHEET

Content Area Definitions

1. Reading Skills which includes the word recognition skills-getting access to word meaning at all stages of reading acquisition-understanding the meaning and function of words as well as their pronunciation. Phonics- consonant blends, consonant digraphs, vowel digraphs, vowel diphthongs, syllables and grapheme bases; Graphic Cues- such as spelling pattern cues, morphemes and configuration; Syntactic Cues- such as word classform (nouns, verbs), syntactic patterns; and Semantic Clues- such as pictures, morphemes, preceding contexts and following contexts.
2. Reading Comprehension which is using strategies to acquire meaning from text. Strategies used could be word grouping, cloze exercises, prediction, construction or reconstruction, using world knowledge, divergent thinking, convergent thinking, literal responses, inferential responses, paraphrasing, opinion responses, evaluation, concept elaboration, concept organization, and study reading.
3. Theory/Research which includes traditional and recent views of the reading process and reading instructional practices, reading programs, specific methods of teaching, classroom organization, ways of assessing and interpreting the reading performance of groups and individuals.

-----	
Knowledge of Teaching Reading in Elementary Grades	
	New Questions
Reading Skills	
Reading Comprehension	
Theory/ Research	

Directions: Place the number of each test item in the appropriate content area and under the appropriate heading.

Pilot Test I  
Inter-rater Item Classifications

Skills				Comprehension				Theory/Research			
1	2	3	4	1	2	3	4	1	2	3	4
	2	2		1	1				1		1
								2	2		
								3	3	3	3
					4			4	4		4
								5	5	5	5
								6	6	6	6
7	7	7	7								
8	8	8	8								
9	9	9	9								
10	10	10	10								
11	11	11	11								
12	12	12	12								
13	13	13	13								
14	14	14	14								
15	15	15	15								
16	16	16	16								
17	17	17	17							17	
18	18	18	18								
19	19	19	19								
20	20	20	20								
								21	21	21	21
								22	22	22	22
23	23	23	23					24	24	24	24
				25	25			25	25		
				26	26	26			26		
				27	27	27			27		
				28	28	28			28		
				29	29	29			29		
				30	30	30			30		
				31	31	31			31		
				32	32	32			32		
				33	33	33					
33	33	33	33	34	34	34					
								34			
								35	35	35	35
								36	36	36	36
								37	37	37	37
								38	38	38	38
								39	39	39	39
								40	40	40	40
								41	41	41	41
								42	42	42	42
								43	43	43	43
								44	44	44	44

Pilot Test I  
Inter-rater Item Classifications

Skills				Comprehension				Theory/Research			
1	2	3	4	1	2	3	4	1	2	3	4
								45	45	45	45
								46	46	46	46
								47	47	47	47
								48	48	48	48
								49	49	49	49
								50	50	50	50
								51	51	51	51
								52	52	52	52
								53	53	53	53
								54	54	54	54
								55	55	55	55
								56	56	56	56
								57	57	57	57
								58	58	58	58
								59	59	59	59
								60	60	60	60
								61	61	61	61
								62	62	62	62
								63	63	63	63
								64	64	64	64
								65	65	65	65
								66	66	66	66
67	67	67	67								
68	68	68	68								
69	69	69	69								
70	70	70	70								
71	71	71	71								
72	72	72	72								
								73	73	73	73
								74	74	74	74
								75	75	75	75
								76	76	76	76

26 New Questions  
Inter-rater Item Classifications

Skills				Comprehension				Theory/Research			
1	2	3	4	1	2	3	4	1	2	3	4
				1	1	1			1		
					2				2	2	
		3			3				3	3	
4		4			4				4		
		5							5	5	
				6	6	6	6				
				7	7	7	7				
					8	8			8	8	
		9							9	9	9
				10	10	10	10				
				11	11					11	11
									12	12	12
									13	13	13
14									14	14	*
									15	15	15
									16	16	16
					17				17	17	
									18	18	18
									19	19	19
									20	20	20
									21	21	21
									22	22	22
					22				23	23	23
									24	24	24
									25	25	25
				26	26	26	26				

\* Rater number 4 did not classify item 14

APPENDIX C

Pilot Test II

KNOWLEDGE OF TEACHING READING IN ELEMENTARY GRADES

1. Which one of the following best helps develop the experiential background necessary for children's readiness for reading?
  - a. take children on fieldtrips
  - b. read literature aloud
  - c. show films and filmstrips
  - d. provide educational toys and games
  
2. The effect of a child's experiential background on the ability to recognize words is
  - a. minor for boys and girls
  - b. substantial for boys and girls
  - c. greater for boys than for girls
  - d. greater for girls than for boys
  
3. The oral language development of a child is important because
  - a. it is a basic ingredient for reading success
  - b. it is a basic ingredient for verbal success
  - c. it is emphasized in the communication skills
  - d. it is emphasized at the primary level
  
4. Readiness is best accomplished
  - a. by giving children meaningful experiences of all kinds
  - b. by having children complete a readiness workbook
  - c. by giving children specific skills through a sequence of exercises
  - d. by having children learn the alphabet
  
5. Which is not a desirable way for parents to facilitate reading readiness?
  - a. read to child frequently
  - b. converse with child to enlarge vocabulary
  - c. provide variety of experience
  - d. help them read library books
  
6. When should systematic and sequential reading instruction begin?
  - a. at the end of kindergarten if the child is ready
  - b. after pupils have finished a reading readiness program
  - c. when the child shows he/she is ready
  - d. no later than the first half of the first grade

7. Vowels are more consistent in their sounds than are
  - a. consonants
  - b. consonant blends
  - c. consonant digraphs
  - d. syllables
8. In the word car the vowel sound is an example of a(n)
  - a. long sound
  - b. short sound
  - c. "r" controlled sound
  - d. hard sound
9. The vowel sound you hear in the word meat is a
  - a. long sound
  - b. vowel digraph
  - c. diphthong
  - d. vowel blend
10. The vowel sound you hear in house is a
  - a. blend
  - b. diphthong
  - c. digraph
  - d. phoneme
11. The medial consonant sound in machine is a
  - a. blend
  - b. diphthong
  - c. digraph
  - d. phonogram
12. The first two letters in the word dwarf are representative of a
  - a. consonant digraph
  - b. consonant diphthong
  - c. diphthong
  - d. consonant blend or cluster
13. Which group of nonsense words below would be representative of the hard g?
  - a. gane, gobbb, guvot
  - b. oge, gint, gisb
  - c. gripit, goot, gyt
  - d. gyymb, gunt, goveret
14. What do the syllables in machine, china, and ache have in common?
  - a. a consonant sound
  - b. a vowel sound
  - c. a digraph
  - d. a diphthong



15. An important suggestion to remember in the teaching of phonics is
  - a. give plenty of attention to letters out of context
  - b. emphasize meaning of words
  - c. give all children equal training
  - d. teach phonics in a separate period
16. If children are to use phonics effectively
  - a. they should be able to recognize the word if they hear it
  - b. they should have mastered all the basic phonic generalizations
  - c. they should be able to recognize the structure of the word
  - d. they should have training in phonetics
17. The practice of teaching every phonics skills lesson in a basal reader series is
  - a. a good model
  - b. based on student's needs
  - c. not necessary
  - d. supported by research
18. All but one of the following are examples of structural analysis. Which one is that one?
  - a. diphthongs
  - b. inflected forms
  - c. derived forms
  - d. compound words
19. Which of the following words does not lend itself easily to structural analysis?
  - a. builders
  - b. building
  - c. rebuild
  - d. built
20. When the analysis of compound words has been over-emphasized, the pupil may
  - a. spend too much time looking for little words within the big word
  - b. simply divide all multi-syllabic words in the middle
  - c. depend too heavily on phonic analysis for the recognition of all big words
  - d. avoid the usage of context clues

21. The word John's is
- the possessive form of John
  - the contracted form of John is
  - dependent upon the context in which the word is found
  - dependent upon whether you are referring to more than one John
22. Words that end in s, es, ed, ing, er, and est are called
- derived forms
  - phonemes
  - definitives
  - inflected forms
23. Unhappy, made from the prefix un and the base word happy is
- a derived form
  - a phoneme
  - a definitive
  - an inflected form
24. How many phonemes does the word bite have?
- 1
  - 2
  - 3
  - 4
25. How many morphemes does the word attainments contain?
- 1
  - 2
  - 3
  - 4
26. Which one of the following begins the first phase of a Directed Reading Lesson?
- reading silently
  - creating interest
  - introducing new vocabulary
  - developing concept background
27. Developing interest and background of experience in a Directed Reading Lesson should
- be determined by students' abilities in relation to length of selection
  - be determined by the length of the selection and students' needs
  - be determined by the students' abilities in relation to that selection
  - be determined by the needs of the students in relation to that selection

28. What are good types of words to teach as sight words?
- irregularly spelled words
  - regularly spelled words
  - words which are unknown to children
  - words that have more than one pronunciation
29. Most readability formulas measure the difficulty of semantic and syntactic variables by tabulating
- the sentence length and number of syllables in a given selection
  - the number of words comprised of 10 or more letters in a given selection
  - the length and number of sentences in a given selection
  - the ratio of numbers of words to numbers of sentences
30. One application of the schema theory for teachers is
- accept the limitations of a student's background of experiences
  - provide instruction involving concept development relevant to reading selections
  - provide instruction in comprehension only in materials students are familiar with
  - teach the concepts of assimilation and accommodation
31. Teaching students to paraphrase is
- a practice which will likely improve their comprehension
  - a practice which will likely improve their reading rate
  - a practice which will likely impair their comprehension
  - a practice which is not likely to influence their comprehension
32. The teacher's major concern when checking a pupil's comprehension of material is
- asking questions relative to the factual details in the selection
  - asking questions consistent with the stated purpose for reading the selection
  - the pupil's ability to evaluate the selection critically
  - the pupil's ability to state the main idea

33. Which is a true statement about reading comprehension?
- it takes less time to teach reading comprehension skills than word recognition skills
  - national test results seem to indicate that comprehension instruction is presently effective
  - all of the phases of the Directed Reading Lesson can contribute to comprehension improvement
  - basal reading materials provide an adequate source for development of comprehension
34. To evaluate pupils' general reading comprehension of a particular type of material, the teacher should
- include several types of comprehension questions on the evaluation
  - include objective questions as often as possible on the evaluation
  - give a worksheet to all children in the class in order to assess each individual's comprehension
  - include cloze and maze type sections on the evaluation
35. Asking students to explain how they arrived at correct or incorrect answers to comprehension questions
- is less important than oral reading skills
  - doesn't help the teacher diagnose comprehension ability
  - helps the teacher diagnose comprehension ability
  - helps the teacher know who read the story closely
36. Which is a true statement about preparing comprehension questions for a selection?
- ask questions that are all of the same type
  - ask questions that discriminate between students
  - ask questions that offer alternative answers
  - ask questions requiring a variety of comprehension skills
37. Which of the following should not be the basis of comprehension questions to ask the child after reading a passage.
- questions about facts
  - inferential questions
  - questions about vocabulary
  - yes/no questions

38. The most effective way of unlocking the meaning of the word bark in the sentence, "The dog's bark was very loud," would be through the use of
- phoneme analysis
  - structural analysis
  - contextual analysis
  - sight word method
39. It is preferable to give children a choice of main ideas for a selection rather than requiring the child to formulate one in their own words
- in primary grades
  - in all elementary grades
  - in upper elementary grades
  - in first grade only
40. Reading to find details is
- always essential for comprehension
  - usually neglected in reading comprehension
  - often overemphasized in reading instruction
  - generally to be avoided
41. To develop a Criterion Referenced Test one must have
- clearly defined objectives
  - clearly defined norms
  - reliability and validity of at least .58
  - reliability and validity of at least .85
42. A standardized reading test score of 7.6 means
- that the child with that score read as well as most children in the sixth month of the seventh grade
  - that the child with that score read as well as most children who are seven years and six month old
  - that the child with that score should be instructed from a seventh grade reading text
  - that the child with that score should be instructed from a book written for the average seven-year old
43. In assessing children's reading abilities, their levels of reading achievement should be considered in relation to
- the norm for their grade levels
  - their chronological ages
  - their expectancy levels
  - the level of the basal used for instruction

44. The most accurate means of determining a child's instructional level is through the use of
  - a. a cloze test
  - b. an informal reading inventory
  - c. an achievement test
  - d. a word list
45. In selecting basal readers or textbooks, the student's level to use as a guide for his reading group is
  - a. instructional level
  - b. independent level
  - c. capacity level
  - d. frustration level
46. In order to effectively assist a youngster in the selection of a library book, a teacher should be most concerned with the child's
  - a. interest
  - b. reading level
  - c. concentration length
  - d. silent reading habits
47. Comparing the readability of a fourth grade basal to a fourth grade geography book would most likely show
  - a. the geography book written on an easier level
  - b. the basal reader on a lower level than the geography book
  - c. the basal reader and the geography book written on the same level
  - d. both books begin on the same level but the geography book ends on a more difficult level
48. The Guided Reading Procedure is most appropriately used with
  - a. basal readers
  - b. content area textbooks
  - c. library books
  - d. language experience stories
49. The practice of assigning the same level of reading materials as standardized test scores is considered
  - a. acceptable
  - b. poor
  - c. as good as any other
  - d. the most effective strategy for placement
50. Round-robin reading
  - a. is an excellent teaching technique
  - b. keeps better readers interested
  - c. does not keep most students on task
  - d. assures the attention of all children

51. When a child does not know a word during oral reading the teacher should
  - a. have the child sound it out
  - b. tell the child the word
  - c. tell the child to guess using context clues
  - d. let other children tell the word
  
52. The greatest disadvantage to employing the Language Experience Approach is that
  - a. the materials are not as interesting as basal texts
  - b. it requires more planning to use a systematic approach to teaching skills
  - c. the children feel that they are in the low reading group
  - d. teachers do not have an adequate knowledge of reading skills to be effective
  
53. Language Experience stories can serve as a basis for teaching
  - a. comprehension skills and sight vocabulary
  - b. word recognition and comprehension skills
  - c. sight vocabulary, comprehension and word recognition skills
  - d. sight vocabulary and word recognition skills
  
54. The method that best builds on the relationship between oral language, background of experience, and reading is
  - a. basal texts
  - b. programmed texts
  - c. LEA
  - d. interest approach
  
55. The factor that makes the most difference in successful reading instruction is
  - a. teachers
  - b. grouping patterns
  - c. materials
  - d. methods of instruction
  
56. Teaching context clues
  - a. has little effect on vocabulary development
  - b. has an important impact on vocabulary development
  - c. has an important role in the Directed Reading Activity
  - d. has an important role in the Guided Reading Procedure

57. In Mr. Brown's classroom several students' parents have volunteered to help him in the classroom during reading. He should let them
- work with small group activities
  - correct workbooks and grade papers
  - read stories or listen to students read orally
  - know teaching really is the teacher's job and politely discourage them
58. Survival skills for remedial readers should include
- initial consonants, inflectional endings, and contractions
  - sight words, contractions, and phonics principles
  - phonics principles, structural analyses, and context clues
  - sight words, initial consonants, and context clues
59. Teaching test taking skills is
- known to produce deflated test results
  - recommended to achieve accurate test results
  - recommended to achieve accurate objectives
  - known to produce inflated test results
60. The Directed Reading Activity seems most consistent with which theory of reading
- bottom-up theory
  - top-down theory
  - interactive theory
  - inferential theory
61. The best suggestion to make to parents who ask what should be done at home to help promote interest in reading is for them to
- help them do their reading homework
  - require more reading time than television viewing time
  - require equal reading time and television viewing time
  - provide a model by reading aloud to their children
62. A good practice to promote an interest in voluntary reading is
- the assignment of written book reports
  - sustained silent reading
  - oral reading during the Directed Reading Activity
  - bulletin board displays of book covers



63. The best advice to receive about good teaching practices for oral reading skills is
- have each child read aloud every day
  - teach the children about eye-voice span
  - have children read silently before reading aloud
  - teach the children to read with expression
64. Providing appropriate reading instruction for exceptional children in the regular classroom is
- more effective if exceptional learners work independently in the classroom
  - part of the teacher's responsibility for matching instruction to all learners
  - not part of the teacher's responsibility for matching instruction to all learners
  - more effective if exceptional learners do not work independently in the classroom
65. The best conclusion to be drawn from the research on methods of teaching reading is that
- each teacher might prefer to use a different method
  - individualized instruction produces the greatest achievement
  - there is no best method for teaching reading
  - the method recommended by the basal teachers' manuals works the best
66. Students should be made to understand that the rate at which they read a selection should be determined by
- the amount of time available for completing the assignment
  - the rate of reading they use when reading in a basal reader
  - the rate at which they read when they are interested in the content
  - the purpose for reading and difficulty of material
67. In selecting books to read aloud to students, teachers should
- be concerned with the independent reading level of the students
  - never read books with characters stereotyped by sex, race, or ethnic group
  - read good literature but explain the presence of stereotyped characters
  - primarily read titles nominated for the William Allen White award

68. An effective way to manage reading instruction in a classroom consisting of children from culturally different backgrounds is to
- form homogeneous cultural groups
  - use cooperative learning activities
  - use extensive whole group instruction
  - use totally individualized instruction
69. The instructional implication of a bottom-up model of reading is an emphasis on
- comprehension questions
  - skills instruction
  - language experience activities
  - sight words
70. A Directed Reading Activity (DRA) is usually
- a whole-class study skills technique
  - a process that can give a reading or content lesson more structure
  - an individualized four-step reading strategy
  - a small group, student-centered process to review previously read material
71. Diagnostic teaching
- is instruction based upon a child's strengths
  - is instruction based on a child's strengths and weaknesses
  - is instruction based on a child's weaknesses
  - is instruction to be used only by reading specialists
72. Diagnostic information
- can be gathered only from standardized tests
  - can be gathered by the teacher through observation
  - should be gathered from teacher-made tests
  - should be recorded either on tape or video for accuracy
73. A student may need instruction in structural analysis skills if he/she lacks mastery of
- recognition of consonant sounds and clusters, vowel sounds, vowel digraphs, and auditory blendings
  - recognition of base words, prefixes, suffixes, syllabication, accents, compound words, and contractions
  - obtaining the meaning and pronunciation of an unknown word by examining the context in which it is located
  - word recognition skills

74. Reading comprehension does not include
- the ability to read rapidly
  - the ability to understand charts and graphs
  - the ability to use encyclopedias
  - the ability to survey a variety of materials
75. Children acquire reading skills
- in much the same way that they acquire oral language
  - in a structured sequence of tasks
  - in an unstructured sequence of tasks
  - in much the same way that they acquire comprehension strategies
76. In order to have successful reading, students need
- phonics skills and word attack skills
  - phonic instruction and comprehension skills
  - word attack skills and a sight vocabulary
  - a sight vocabulary and comprehension skills
77. Word attack and phonic skills should
- begin each reading instruction lesson
  - be concentrated in the primary grades
  - be followed with written exercises
  - be followed with reading in context
78. Students who know words in context but not in isolation are
- telling you they have small sight vocabularies
  - in need of structural analysis skills
  - in need of semantic and syntactic skills
  - telling you that context aids their reading
79. Students participate in comprehension
- through the active process of predicting
  - through the passive process of silent reading
  - through reading familiar materials
  - through reading unfamiliar materials
80. In comprehension development two situations exist for students to answer questions. They are
- locate and orally read
  - recall and locate
  - recall and summarize
  - locate and summarize

81. Student responses are increased and are more complete when
- teachers decrease wait time between question and response
  - other students are allowed to give clues
  - teachers allow more wait time between question and response
  - other students are not allowed to give clues
82. Reading skills are most effectively developed and refined by
- focusing on readers' weaknesses
  - giving individual instruction
  - giving whole group instruction
  - focusing on readers' strengths
83. One important fundamental principle to remember when selecting materials to teach comprehension strategies is
- to be sure that the context is sufficiently meaningful for students to relate what they already know to the text
  - to be sure students are forced to solely attend to words and sound symbol information
  - to be sure that the comprehension strategies being taught have not been modeled or guided by the teacher
  - to be sure that the context is new and different from what the students already know
84. Content area reading
- uses the same methods of presentation that are used during reading instruction
  - uses a different method of presentation than is used during reading instruction
  - is not taught by elementary school teachers
  - can not be presented through the language experience method
85. Most low-achieving readers
- face problems that are largely in their control
  - face problems that are largely out of their control
  - face a greater demand on them than average or above average readers
  - face receiving instruction from many rather than just one teacher

86. Teachers often face the pressure to satisfy others that all readers are reading on grade level. This expectation is
- realistic
  - idealistic
  - unreasonable
  - reasonable
87. Studying word families as a means of increasing vocabulary knowledge is
- appropriate for primary students
  - appropriate for upper elementary students
  - appropriate for all grades
  - not appropriate for all grades
88. When students encounter difficulty in learning to read, the reaction of
- teachers and parents is crucial
  - parents is crucial
  - peers and parents is crucial
  - peers and teachers is crucial
89. Reading difficulties can be caused by
- emotional disturbances or disabilities
  - physical disabilities
  - intellectual and emotional disabilities
  - intellectual, physical and emotional disabilities
90. If reading goals seem worth reaching, either because they are immediately interesting or because they serve some need, students
- will direct their energies towards those ends without prodding
  - will often overlook important details in their efforts to complete the task
  - will become competitive with peers and reduce their time on task
  - will direct their attention towards those goals with teacher prodding
91. Elementary teachers who obtain more alert and enthusiastic participation are observed to be
- strict and well-organized
  - considerate and well-organized
  - impersonal and goal orienting
  - indulgent and orderly

92. Periodic evaluation of each student's reading is primarily important
- for teachers' grading reports
  - as a means of evaluating teaching effectiveness
  - as feedback to the learner
  - for determining regrouping of students
93. The most common cause for difficulty in learning is inadequate fundamental skills in
- written expression
  - reading
  - arithmetic
  - ability to plan and organize
94. One goal of the elementary school is the development of the fundamental skills of reading, quantitative thinking, and communication. Most of the teacher's evidence of this
- comes from seatwork, workbooks and written work
  - comes from the oral responses of students
  - comes from quizzes and tests
  - comes from observation of daily performance
95. Older elementary students may find it beneficial to follow a specific study technique in the reading of content area materials. The most well-known of these is the
- DRTA
  - IEP
  - SQ3R
  - SRA
96. Reluctant readers are
- those students who never volunteer to read orally
  - those students who have the necessary skills but don't use them
  - those students who never have the correct place when it is their turn to read orally
  - those students who can not answer comprehension questions
97. Which one is not a basic concept essential to good educational practices?
- the family is not an important institution in education
  - all children are capable of being educated
  - teaching and learning should be a planned systematic process
  - everyone needs to learn how to read and write

98. One way not to grow in your profession as a teacher is to
- be willing to improve skills and increase knowledge
  - be willing to try new things
  - be willing to maintain an open-door policy in teaching
  - be willing to let seasoned teachers create your daily teaching schedule
99. Comprehension questions requiring pupil reactions, insights, judgements, or reflections before answering are
- highly effective
  - highly reflective
  - low in effectiveness
  - low in content
100. According to research on methods of improving comprehension, practice
- seems to be a poor use of instructional time
  - results in improved comprehension
  - does not consistently affect comprehension
  - has not been investigated adequately
101. Good comprehension instruction should be expected to
- restructure what a student already knows so he/she can transfer it to a new situation
  - increase a student's knowledge about main ideas
  - expand a student's ability to critic texts
  - increase a student's study reading
102. A good practice to follow when introducing a new comprehension method is
- to model it
  - to discuss it
  - to break it down into three steps
  - to make it very structured
103. In comprehension instruction, inferencing
- is a method of predicting the main idea
  - is a relationship between two or more units of information
  - is a means to answer all comprehension questions clearly
  - is primarily used for evaluative questions

104. Comprehension questions can be used for
- a. narrative and sequencing material
  - b. expository material
  - c. narrative and expository material
  - d. sequence and expository material



APPENDIX D

Raw Data

Pilot Test I

ITEM NO.	NO. COR.	% COR.	ITEM DISCRIMINATION			KEYED* RESPONSES				
			HI 27%	LO 27%	INDEX	1	2	3	4	5
1	28	93.3	7	7	+0.00	0	28*	0	2	0
2	29	96.7	7	8	-0.11	1	29*	0	0	0
4	30	100.0	8	8	+0.00	30*	0	0	0	0
13	28	93.3	7	7	+0.00	28*	1	1	0	0
15	30	100.0	8	8	+0.00	0	0	0	30*	0
19	24	80.0	6	6	+0.00	0	5	24*	1	0
21	2	6.7	0	1	-0.11	2*	0	1	27	0
22	23	76.7	6	7	-0.11	1	6	23*	0	0
23	27	90.0	8	6	+0.25	1	27*	1	1	0
24	27	90.0	7	7	+0.00	1	27*	2	0	0
25	27	90.0	8	6	+0.25	3	0	27*	0	0
26	26	86.7	8	6	+0.25	3	0	1	26*	0
27	25	83.3	8	6	+0.25	25*	3	1	1	0
28	29	96.7	8	8	+0.00	1	29*	0	0	0
31	19	63.3	6	3	+0.37	2	19*	3	6	0
32	6	20.0	2	1	+0.12	11	6*	9	4	0
33	29	96.7	8	7	+0.12	0	1	29*	0	0
34	24	80.0	6	5	+0.12	24*	0	2	0	4
35	26	86.7	7	7	+0.00	1	2	0	26*	1
36	2	6.7	0	0	+0.00	26	2*	0	2	0
37	21	70.0	8	1	+0.87	7	2	21*	0	0
38	18	60.0	7	5	+0.25	11	1	0	18*	0
39	22	73.3	8	8	+0.00	22*	0	0	8	0
40	23	76.7	6	5	+0.12	5	2	23*	0	0
41	22	73.3	7	6	+0.12	4	3	22*	1	0
43	28	93.3	8	7	+0.12	28*	0	1	1	0
45	28	93.3	8	7	+0.12	2	0	0	28*	0
48	2	6.7	1	1	+0.00	2*	0	2	26	0
52	14	46.7	6	1	+0.62	14*	0	5	11	0
53	19	63.3	7	2	+0.62	3	19*	2	6	0
55	29	96.7	8	7	+0.12	29*	1	0	0	0
57	27	90.0	8	6	+0.25	0	27*	3	0	0
60	24	80.0	7	5	+0.25	0	0	24*	6	0
61	28	93.3	8	6	+0.25	28*	2	0	0	0
62	0	0.0	0	0	+0.00	0	0	0*	30	0
63	28	93.3	8	6	+0.25	0	28*	2	0	0
66	30	100.0	8	8	+0.00	0	0	0	30*	0
67	28	93.3	8	6	+0.25	0	0	2	28*	0
68	30	100.0	8	8	+0.00	0	0	30*	0	0
69	12	40.0	6	3	+0.37	12*	18	0	0	0
70	22	73.3	7	4	+0.37	3	3	22*	2	0
71	27	90.0	8	7	+0.12	0	27*	1	2	0
72	29	96.7	8	8	+0.00	29*	1	0	0	0
74	29	96.7	8	7	+0.12	29*	0	0	1	0
76	22	73.3	7	3	+0.50	22*	2	3	3	0
78	23	76.7	8	5	+0.37	4	2	23*	1	0
79	21	70.0	7	4	+0.37	7	21*	2	0	0
80	30	100.0	8	8	+0.00	30*	0	0	0	0
81	29	96.7	7	8	-0.11	1	29*	0	0	0

ITEM NO.	NO. COR.	% COR.	ITEM DISCRIMINATION			KEYED* RESPONSES				
			HI 27%	LO 27%	INDEX	1	2	3	4	5
86	0	0.0	0	0	+0.00	3	23	0*	4	0
87	13	43.3	3	0	+0.37	5	13*	10	2	0
90	29	96.7	8	8	+0.00	0	29*	1	0	0
93	27	90.0	8	5	+0.37	1	1	27*	1	0
94	26	86.7	8	5	+0.37	1	26*	3	0	0
97	29	96.7	8	8	+0.00	0	29*	0	1	0
100	28	93.3	7	8	-0.11	0	0	2	28*	0
101	29	96.7	8	7	+0.12	1	0	29*	0	0
107	13	43.3	6	3	+0.37	13*	0	2	14	0
108	15	50.0	6	4	+0.25	13	15*	1	1	0
110	30	100.0	8	8	+0.00	0	0	30*	0	0
112	18	60.0	6	4	+0.25	0	1	11	18*	0
114	27	90.0	7	7	+0.00	0	27*	0	3	0
115	10	33.3	3	2	+0.12	6	10*	9	5	0
116	29	96.7	8	8	+0.00	0	0	1	29*	0
117	26	86.7	8	5	+0.37	0	26*	1	3	0
118	29	96.7	8	7	+0.12	0	0	29*	1	0
120	30	100.0	8	8	+0.00	0	30*	0	0	0
121	27	90.0	8	7	+0.12	0	3	27*	0	0
122	30	100.0	8	8	+0.00	0	0	0	30*	0
124	19	63.3	7	3	+0.50	2	9	19*	0	0
125	23	76.7	7	6	+0.12	2	23*	1	4	0
126	13	43.3	6	2	+0.50	6	13*	3	8	0
127	26	86.7	8	5	+0.37	4	26*	0	0	0
130	29	96.7	8	7	+0.12	0	29*	1	0	0
131	27	90.0	8	7	+0.12	2	27*	1	0	0
132	24	80.0	8	5	+0.37	2	24*	1	3	0

Pilot Test II

ITEM NO.	NO. COR.	NO. COR.	ITEM DISCRIMINATION			KEYED* RESPONSES				
			HI 27%	LO 27%	INDEX	1	2	3	4	5
1	11	45.8	4	3	+0.17	12	11*	0	0	1
2	22	91.7	5	5	+0.00	0	22*	2	0	0
3	15	62.5	6	5	+0.17	15*	4	5	0	0
4	23	95.8	6	6	+0.00	23*	0	1	0	0
5	21	87.5	6	4	+0.33	0	1	1	21*	0
6	17	70.8	5	4	+0.17	1	3	17*	3	0
7	11	45.8	3	3	+0.00	11*	2	7	4	0
8	23	95.8	5	6	-0.16	0	1	23*	0	0
9	18	75.0	6	5	+0.17	5	18*	1	0	0
10	22	91.7	6	6	+0.00	2	22*	0	0	0
11	17	70.8	4	5	-0.16	6	1	17*	0	0
12	22	91.7	6	5	+0.17	1	0	0	22*	0
13	21	87.5	6	5	+0.17	21*	3	0	0	0
14	5	20.8	2	1	+0.17	3	5*	15	1	0
15	16	66.7	5	3	+0.33	5	16*	0	3	0
16	3	12.5	2	0	+0.33	11	3*	7	3	0
17	17	70.8	5	2	+0.50	0	7	17*	0	0
18	24	100.0	6	6	+0.00	24*	0	0	0	0
19	20	83.3	5	5	+0.00	0	1	2	20*	1
20	22	91.7	5	6	-0.16	22*	1	0	1	0
21	22	91.7	6	5	+0.17	2	0	22*	0	0
22	20	83.3	6	2	+0.67	3	0	0	20*	1
23	22	91.7	6	4	+0.33	22*	0	1	1	0
24	15	62.5	4	3	+0.17	4	1	15*	4	0
25	16	66.7	5	5	+0.00	4	3	16*	1	0
26	15	62.5	3	4	-0.16	0	15*	2	7	0
27	17	70.8	6	3	+0.50	0	3	3	17*	1
28	12	50.0	4	1	+0.50	12*	11	1	0	0
29	16	66.7	5	2	+0.50	16*	2	2	3	1
30	14	58.3	5	1	+0.67	3	14*	3	3	1
31	20	83.3	6	4	+0.33	20*	3	1	0	0
32	14	58.3	4	1	+0.50	1	14*	3	6	0
33	20	83.3	5	5	+0.00	0	2	20*	2	0
34	23	95.8	6	6	+0.00	23*	0	0	1	0
35	23	95.8	6	5	+0.17	0	0	23*	1	0
36	19	79.2	6	4	+0.33	1	1	2	19*	0
37	21	87.5	6	4	+0.33	0	2	0	21*	1
38	23	95.8	6	5	+0.17	0	0	23*	1	0
39	17	70.8	5	3	+0.33	17*	2	0	5	0
40	20	83.3	6	4	+0.33	2	1	20*	1	0
41	21	87.5	6	5	+0.17	21*	0	0	3	0
42	18	75.0	5	4	+0.17	18*	3	3	0	0
43	16	66.7	6	2	+0.67	5	1	16*	2	0
44	17	70.8	5	3	+0.33	4	17*	2	1	0
45	23	95.8	6	6	+0.00	23*	0	1	0	0
46	10	41.7	2	1	+0.17	14	10*	0	0	0
47	23	95.8	6	5	+0.17	1	23*	0	0	0
48	17	70.8	5	3	+0.33	7	17*	0	0	0
49	20	83.3	6	4	+0.33	3	20*	0	1	0
50	23	95.8	6	5	+0.17	0	0	23*	1	0

ITEM NO.	NO. COR.	NO. COR.	ITEM DISCRIMINATION			KEYED* RESPONSES				
			HI 27%	LO 27%	INDEX	1	2	3	4	5
51	22	91.7	6	4	+0.33	0	22*	0	0	0
52	24	100.0	6	6	+0.00	0	24*	0	0	0
53	18	75.0	5	4	+0.17	2	0	18*	4	0
54	24	100.0	6	6	+0.00	0	0	24*	0	0
55	13	54.2	4	3	+0.17	13*	2	0	9	0
56	22	91.7	6	4	+0.33	0	22*	1	1	0
57	13	54.2	4	1	+0.50	9	2	13*	0	0
58	20	83.3	5	5	+0.00	0	0	3	20*	1
59	15	62.5	4	3	+0.17	0	15*	5	4	0
60	11	45.8	3	3	+0.00	8	11*	5	0	0
61	21	87.5	6	5	+0.17	0	2	0	21*	1
62	24	100.0	6	6	+0.00	0	24*	0	0	0
63	21	87.5	6	5	+0.17	2	0	21*	1	0
64	22	91.7	5	6	-0.16	1	22*	1	0	0
65	22	91.7	6	6	+0.00	1	1	22*	0	0
66	23	95.8	6	5	+0.17	0	0	0	23*	1
67	18	75.0	5	3	+0.33	2	4	18*	0	0
68	23	95.8	6	6	+0.00	0	23*	0	1	0
69	18	75.0	5	4	+0.17	4	18*	1	1	0
70	21	87.5	6	4	+0.33	2	21*	0	1	0
71	20	83.3	4	5	-0.16	0	20*	4	0	0
72	18	75.0	6	5	+0.17	1	18*	5	0	0
73	16	66.7	5	3	+0.33	0	16*	4	4	0
74	22	91.7	6	6	+0.00	22*	1	0	1	0
75	16	66.7	3	4	-0.16	5	16*	3	0	0
76	3	12.5	2	0	+0.33	3	2	3*	15	1
77	16	66.7	6	2	+0.67	3	4	0	16*	1
78	18	75.0	6	3	+0.50	2	2	1	18*	1
79	15	62.5	4	3	+0.17	15*	2	2	4	1
80	4	16.7	2	0	+0.33	0	4*	20	0	0
81	19	79.2	6	4	+0.33	0	5	19*	0	0
82	12	50.0	3	4	-0.16	4	8	0	12*	0
83	23	95.8	6	5	+0.17	23*	0	0	1	0
84	8	33.3	3	2	+0.17	16	8*	0	0	0
85	8	33.3	3	1	+0.33	0	8*	13	3	0
86	18	75.0	3	5	-0.32	0	6	18*	0	0
87	11	45.8	4	3	+0.17	9	0	11*	4	0
88	20	83.3	5	3	+0.33	20*	1	1	2	0
89	22	91.7	6	4	+0.33	0	0	1	22*	1
90	22	91.7	6	6	+0.00	22*	0	0	2	0
91	21	87.5	6	5	+0.17	2	21*	0	1	0
92	8	33.3	2	1	+0.17	2	6	8*	8	0
93	21	87.5	6	3	+0.50	0	21*	0	2	1
94	19	79.2	5	4	+0.17	0	4	0	19*	1
95	24	100.0	6	6	+0.00	0	0	24*	0	0
96	12	50.0	5	3	+0.33	10	12*	0	2	0
97	18	75.0	4	5	-0.16	18*	3	3	0	0
98	21	87.5	6	3	+0.50	0	1	1	21*	1
99	20	83.3	6	5	+0.17	20*	4	0	0	0
100	22	91.7	6	4	+0.33	0	22*	1	1	0

ITEM NO.	NO. COR.	NO. COR.	ITEM DISCRIMINATION			KEYED* RESPONSES				
			HI 27%	LO 27%	INDEX	1	2	3	4	5
101	21	87.5	6	3	+0.50	21*	2	0	1	0
102	18	75.0	6	3	+0.50	18*	3	0	3	0
103	9	37.5	2	2	+0.00	10	9*	0	5	0
104	12	50.0	2	3	-0.16	6	1	12*	4	1



DETERMINING THE RELIABILITY AND VALIDITY  
OF AN INSTRUMENT TO MEASURE BEGINNING TEACHER  
KNOWLEDGE OF READING INSTRUCTION

by

LESLIE TARBET

B. A., Northwestern University, 1961

---

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

College of Education

KANSAS STATE UNIVERSITY  
Manhattan, Kansas

1984

## ABSTRACT

The importance of identifying elementary teachers knowledge of teaching reading has become recognized increasingly in recent years. The purpose of this study was to construct a Knowledge of Teaching Reading in Elementary Grades test. The results of giving the test in a Pilot Study were then investigated to measure the reliability and validity of the test.

Items were constructed by a panel of reading professors and the author. They were written for three areas of knowledge identified in the literature: skills, comprehension, and theory/research. They were scrutinized for face validity by a panel of reading professors. Revisions to stems and/or foils were made. Pilot Test I was given to a 30-subject group of pre-service elementary teachers of reading. Test data was analyzed for reliability using the Kuder-Richardson formula and for validity by analyzing the difficulty and discrimination indexes for the items. Item corrections were made and new items written. Items were then classified to the three areas of knowledges by another panel of reading experts for inter-rater agreement. Inter-rater agreement on all items was 65 percent; while on three-fourths of the items, agreement was 90 percent.

A second test was given to a 24-subject group of pre-service elementary teachers of reading. Test data was analyzed in the same manner as the first Pilot Test.

Analyses of the results of the two Pilot Tests showed that this study is moving in a positive direction in attempting to measure beginning teachers knowledge of teaching reading. The reliability coefficient improved from .73 to .77. Validity improved 14 percent on the difficulty index and 30 percent on the discrimination index.

Although discrete guidelines exist for norm-referenced test item analysis, none are currently evident for criterion-referenced tests. Future work on this test should work towards improving the difficulty and discrimination of test items. One area that would prove helpful would be a study of the foils provided for each item. Continually updating the item pool will also improve the content validity of this test. Future work on this test should provide a useful tool in making decisions regarding the knowledge of pre-service elementary teachers of reading.