5-1991

The development of an instrument to measure the attitudes of middle school mathematics teachers toward low achievers in mathematics

Evelyn Elizabeth Dwyer

To the Graduate Council:

I am submitting herewith a dissertation written by Evelyn Elizabeth Dwyer entitled "The development of an instrument to measure the attitudes of middle school mathematics teachers toward low achievers in mathematics." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Education.

Donald J. Dessart, Major Professor

We have read this dissertation and recommend its acceptance:

Arnold Davis, Phyllis Huff,Tom Mathews

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Evelyn Elizabeth Dwyer entitled "The Development of an Instrument to Measure the Attitudes of Middle School Mathematics Teachers Toward Low Achievers in Mathematics." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Education.

Donald J. Dessart

Donald J. Dessart, Major Professor

We have read this dissertation
and recommend its acceptance:

Arnold R Davis

Phyllis E Huff

H T Mathews

Accepted for the Council:

C W Minkel

Vice Provost
and Dean of The Graduate School

THE DEVELOPMENT OF AN INSTRUMENT TO MEASURE THE ATTITUDES OF MIDDLE SCHOOL MATHEMATICS TEACHERS TOWARD LOW ACHIEVERS IN MATHEMATICS

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Evelyn Elizabeth Dwyer

May 1991

MY ATTITUDES

Love and gratitude to Eddie;

Pride and love for Christopher,
Julie Anne, and Andrew;

Sincere appreciation to Catana,
and all dear friends who have smoothed
my path.

ACKNOWLEDGEMENTS

I am grateful to Dr. Donald J. Dessart, my major professor, for encouragement, guidance, and good humor throughout my doctoral program. I am also appreciative of the support extended by Dr. Arnold Davis, Dr. Phyllis Huff, Dr. Tom Mathews, and Dr. Betty Heathington.

Special thanks are expressed to Dr. Charleen M. DeRitter, Mathematics Supervisor for Knox County, for her interest and support in this research project. Thanks are also extended to the two middle school travelling mathematics teachers in Knox County, Ms. Carol Russell and Ms. Lynda Kennedy, for their valuable assistance in collecting data. I also thank the 17 middle school mathematics department chairpersons, who graciously consented to administer the attitude scales at their school sites, the principals of the middle schools in which the study was conducted, and the many classroom teachers who were kind enough to respond to the research instruments.

ABSTRACT

This study was undertaken to develop and field test an instrument designed to measure attitudes of middle school mathematics teachers toward low achievers in mathematics. Three 15-item subscales designed to measure teacher beliefs, feelings, and intended behaviors toward low achievers in mathematics are contained in this 45-item Likert-type scale titled: Teacher Attitudes Toward Low Achievers in Mathematics Scale (TALAM).

The study was carried out in three phases. In Phase 1, comments about low achievers in mathematics were elicited from middle school mathematics teachers, experts in the field of mathematics and mathematics education, and from related literature. Statements were further validated by a panel of experts.

In Phase 2, the validated items were administered to 51 middle school mathematics teachers. The data generated were analyzed to estimate validity and reliability. Items were retained for the final scale if they had significant ($p < .01$) item-total correlations and the ability to discriminate between high and low criterion groups ($p < .05$). The three scales were found to be internally consistent (Cronbach's alpha range: .80 - .91). Principal component factor analysis of data resulted in three empirically distinguishable factors consistent with placement of statements within the three subscales and indicative of construct validity.

In Phase 3, the final scale was administered to 105 middle school mathematics teachers. The three scales were again found to be internally consistent (Cronbach's alpha range: .70 -.91) and stable (r coefficient range: .70 - .82) over a two-week interval. Concurrent and differential validity were also determined. The attitudinal components measured by the TALAM scales differed from those measured by the Revised Math Attitude Scale suggesting that teachers' attitudes toward mathematics in general are distinct from attitudes toward low achievers in mathematics. Further, the relationship between the TALAM scale and a previously developed semantic differential scale purporting to measure teacher attitudes toward low achievers in general was determined. Data analysis produced significant Pearson product-moment correlations ($p < .01$) between TALAM scales and the semantic differential scale (r coefficient range: .38 - .50).

Overall analysis of data yielded substantial support for the TALAM as a valid and reliable measure of attitudes of middle school mathematics teachers toward low achievers in mathematics. This contention, along with review of earlier research, provides support for further study of the relationship between teacher attitudes toward low achievers in mathematics and a variety of other variables.

TABLE OF CONTENTS

CHAPTER		PAGE
I.	INTRODUCTION	1
	Background	1
	Statement of the Problem	3
	Purpose of the Study	6
	Importance of the Study	8
	Assumptions	9
	Limitations of the Study	10
	Delimitations of the Study	11
	Definitions	11
	Organization of the Study	13
II.	REVIEW OF LITERATURE	15
	Definitions and Components of Attitude	15
	Definitions	15
	Components	16
	The Measurement of Attitudes	18
	Observational Methods	18
	Self-Report Methods	20
	Selected Techniques of Attitude Scale Construction	21
	Thurstone Technique	21
	Likert Scales	24
	Guttman Scales	28
	Semantic Differential Scales	31
	Methods for Estimating Reliability and Validity	33
	Reliability	33
	Validity	37
	Attitude Scales Related to Mathematics	42
	Mathematics Anxiety	43
	Attitudes Toward Mathematics	45
	Teacher Attitudes	47
III.	METHODS AND PROCEDURES	49
	Phase 1: Readiness	49
	Selection of Instrument Type and Format	49
	Writing the Potential Items	51
	Classification of Items by Judges	53
	Preparation of the Preliminary Attitude Scale	54
	Phase 2: Administration of Preliminary Scale	54
	Rationale	55
	Population and Sample	56
	Administration Procedures	56
	Data Analysis Procedures	58
	Construction of the Final Scale	59
	Phase 3: Administration of Final Attitude Scale	60
	Instrumentation	60
	Population and Sample	63
	Administration Procedures	63
	Data Analysis Procedures	66

CHAPTER PAGE

IV. ANALYSIS OF THE DATA 68
Phase 1: Readiness 68
Classification of Items by Judges 68
Preparation of the Preliminary Attitude Scale 69
Phase 2: Administration of Preliminary Scale 69
Data Obtained from Administration of 85-Item Scale 73
Data Obtained from Revision of Total Scale and Subscale 87
Phase 3: Administration of Final Scale 97
Scoring 99
Descriptive Statistics 99
Test of Significance 99
Multiple-Comparison Procedure 103
Internal Consistency Reliability 103
Correlations 103

V. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS 107
Summary 107
Phase 1 107
Phase 2 108
Phase 3 111
Conclusions 114
Recommendations for Uses of the Scale 116
Recommendations for Future Study 117
A Final Word 118

REFERENCES 120

APPENDICES 131

APPENDIX A 132

APPENDIX B 143

APPENDIX C 151

APPENDIX D 159

VITA 162

LIST OF TABLES

TABLE PAGE

1. Number of Teachers Contacted and Response Rate in Phase 2 . . 57
2. Number of Teachers Contacted and Response Rate in Phase 3 . . 64
3. Judgements of Ten Experts Regarding the Directionality of Items 70
4. Preliminary Scale Item Numbers, Subscale Placement, and Directionality 72
5. Subscales and Total Scale Score Distributions and Approximate Quartile Locations 74
6. Comparison of Scale Means Using Student's t Statistic 75
7. Rotated Factor Matrix Values for the Preliminary Total Scale 77
8. Item-Scale Correlations for Preliminary Total Scale Items . . 79
9. Item-Scale Correlations for Teacher Belief Preliminary Subscale Items 82
10. Item-Scale Correlations for Teacher Feelings Preliminary Subscale Items 83
11. Item-Scale Correlations for Intended Teacher Behavior Preliminary Subscale Items 84
12. Item-Scale Correlations for Cultural Belief Subscale Items 85
13. Internal-Consistency Estimates of Reliability (Coefficient Alpha) for the Preliminary Scale and Subscales 86
14. Inter-Correlation Matrix for the Preliminary Scale and Subscales 86
15. Determination of the Discriminating Quality of Preliminary Subscale Items Using Student's t test 89
16. Item-Scale Correlations for the Revised Total Scale (45 Items) 91
17. Item-Scale Correlations for the Revised Teacher Beliefs Subscale (15 Items) 93

TABLE PAGE

18. Item-Scale Correlations for the Revised Teacher Feelings Subscale (15 Items) 94

19. Item-Scale Correlations for the Revised Intended Teacher Behavior Subscale (15 Items) 95

20. Internal Consistency Estimates of Reliability (Coefficient Alpha) for the Revised Total Scale and Subscales 96

21. Inter-Correlation Matrix for the Revised Total Scale and Subscales 96

22. Factor Analysis Results for the Revised Attitude Scale (45 Items) 98

23. Subscale and Total Scale Score Distributions with Approximate Quartile Locations of Scales Shown 100

24. TALAM Data: Descriptive Statistics 102

25. Analysis of Variance for Means of Subscale and Total Scale Scores 102

26. Scheffe's Multiple Comparison of Means of Subscale and Total Scale Scores 104

27. Internal-Consistency Estimates of Reliability (Cronbach's Alpha) for the Final Scale and Subscales 104

28. TALAM Test-Retest Inter-Correlation Matrix (Pearson Product Moment) 105

29. TALAM Correlations with Other Instruments 106

CHAPTER I

INTRODUCTION

Background for the Study

Studies relative to the mathematical achievement of American school children indicate that serious problems exist (Kirsch & Jungeblut, 1986; Steen 1989). Difficulties in mathematics for many children begin when they are very young and persist from primary through secondary school (Stevenson, 1987; McKnight, Crosswhite, Dossey, Kifer, Swafford, Travers, & Cooney, 1987). In regard to performance in mathematics, the National Commission on Excellence in Education (1983) reported in A Nation At Risk that quantitative Scholastic Aptitude Test scores had dropped nearly 40 points from 1963 through 1980. A more recent indicator of difficulties with mathematics among American school children is the report of the National Assessment of Educational Progress cited by Brown, Carpenter, Kouba, Lindquist, Silver, and Swafford (1988). This NAEP report found major deficiencies in both computational and reasoning ability among American school children at both the elementary and secondary levels.

Many educators and psychologists believe mathematical achievement is effected most by student attitude toward their ability to do mathematics as well as the value they place on mathematics in general (Brophy, 1986; Confrey, 1986; Schunk, 1985; Buchanan, 1987; Frary & Ling, 1983). In this light, the National Council of Teachers of Mathematics (NCTM) (1989, 1991) proposed that student confidence and attitudes toward mathematics are critical components effecting achievement in mathematics. NCTM further advised educators to foster the development of positive mathematical disposition among school children at all grade levels as a means to maximize learning. Researchers, Carpenter, Corbitt, Kepner, Lindquist, and Reys (1980), Schoenfeld (1989), and Hart (1989) have also stressed the importance of affect in the mathematics classroom.

The interaction of student attitudes, motivation, and mathematical achievement has been extensively studied at all grade levels. However, researchers have consistently reported a substantial increase in pessimistic attitudes toward ability to learn mathematics among students from fifth to tenth grades with the sharpest attitudinal decline appearing at entrance to junior high school (Eccles, 1983; Eccles, Midgley, & Adler, 1984). Further, at the junior high school level, Pedersen, Bleyer, and Elmore (1985) found that negative attitudes towards ability to do well in mathematics correlated highly with student feelings of inability to have future careers in the natural, social, and medical sciences as well as in business operations. Although the apparent decline in positive attitudes toward mathematics in the middle grades might be due in part to changes commensurate with preadolescent development, researchers Midgley, Feldlaufer, and Eccles (1989) contend that changes in classroom environment as children move from lower elementary grades to the middle grade levels contributes heavily to decline in attitude, motivation, and subsequent performance in mathematics.

Statement of the Problem

Classroom teachers are of primary importance in determining classroom environment (Good & Brophy, 1987). Abundant research indicates that teacher attitudes and behaviors have a great deal of influence on student attitudes and achievement across the curriculum (Aiken, 1972; Good, 1981; Brophy & Good, 1970; Rosenthal & Jacobson, 1968; Rosenthal & Rubin, 1971; Rosenthal, 1974). Further, Aiken (1974) concluded that teachers are viewed as the prime determiners of both student attitude and performance in mathematics. In assessing student attitude toward mathematics and reasons for dislike of the subject, Quilter and Harper (1988) reported that students' identified the mathematics teacher as the greatest source of dissatisfaction with the subject. In an extensive review of research on attitudes toward mathematics reported shortly after his own empirical study, Aiken (1976) determined that findings suggest a strong relationship between the attitudes and behaviors of teachers and student attitudes and performance in mathematics.

During the past 20 years teacher attitudes and behaviors toward students perceived to be low achievers and students perceived to be high achievers have been extensively explored. The literature contains abundant empirical evidence documenting what appears to be inappropriate and differential treatment by teachers toward students perceived by them as low achievers (Brophy, 1979; Good, 1981; Rosenthal & Jacobson, 1968; Thorndike, 1968). Good (1981) provided a model, describing the effects of differential treatment of low and high achieving students by teachers:

1. The teacher expects specific behavior and achievement from particular students.

2. Because of these varied expectations, the teacher behaves differently toward different students.

3. This treatment communicates to the students what behavior and achievement the teacher expects from them and effects their self-concepts, achievement motivation, and levels of aspiration.

4. If this treatment is consistent over time, and if the students do not resist or change it in some way, it will shape their achievement and behavior. High-expectation students will be led to achieve at high levels, whereas the achievement of low expectation students will decline.

5. With time, students' achievement and behavior will conform more and more closely to the behavior originally expected of them.

(pg. 416)

Good (1981) described a study wherein a junior high school teacher taught a lower level mathematics class very differently from how he taught his upper level algebra class. Good found less teacher effort, enthusiasm, and positive reinforcement expended for the lower level class when compared with the algebra class. Not surprisingly, Good found lower levels of achievement gain among students in the general mathematics class.

Studies of this nature suggest that teacher attitudes and behaviors powerfully influence student attitude and achievement. All correlational evidence found in the research points to a strong relationship between teacher attitude and student achievement. Further, it is logical that a highly positive teacher attitude toward all levels of students would stimulate student performance. Therefore, the suggestion that teachers must develop and maintain positive attitudes and high expectations for all students including low achievers appears warranted. Specifically, low achievers in mathematics deserve and require the same degree of supportiveness, respect, response opportunities, and encouragement from their teachers that high achievers receive. Further, teachers are advised to treat all students as though they are expected to meet at least minimum specified achievement goals.

Compared to other industrialized nations, U.S. students rank low in mathematical achievement (Steen, 1989; Stevenson, 1987). In looking for insights into the source of Asian and Asian-American students apparent superior performance on standardized mathematics tests, Stevenson (1987) suggested that positive attitudes and high expectations held by their parents and their teachers, as well as the students themselves might be the key to unlocking student achievement potential.

In a conversation with Jaime Escalante, the hero-teacher portrayed in the movie **Stand and Deliver**, Meek (1989) asked Escalante what made him confident he would be able to teach mathematics to poor, inner-city, high school students. Escalante's reply is germane to this study:

> I don't think kids cannot learn. That's my own philosophy. Anybody, any kid can learn if he or she has the desire to do it. That's what ganas is about. The teacher plays an important role in education--we all remember the first teacher who really touched our lives, or gave us some encouragement, or at least appreciated our best. The teacher gives us the desire to learn, the desire to be Somebody.
>
> (pg. 47)

Like Escalante, Wallis Green (1991) appears well qualified to advise educators concerning classroom environments. Green won the 1988 Presidential Award for Excellence in Science and Mathematics Teaching presented by the National Council of Teachers of Mathematics. Green emphasized what he called the "personal approach" (pg. 32) in his high school classes to generate a mathematics classroom environment that "makes the individual students feel as though they are important parts of the class and integral players in solving the problems at hand" (pg. 32). Green stressed the importance of generating an environment where students "feel comfortable and confident when they ask questions and when they try to solve problems" (pg. 32). Green also stressed the importance of teaching students about the relevancy of mathematics in a variety of circumstances including applications in everyday life.

Purpose for the Study

The purpose for undertaking this study was to provide teachers, supervisors, and school administrators with a scale for measuring teacher attitude toward low achievers in the mathematics. The study was designed to develop an instrument possessing attributes beneficial to classroom teachers and other school system personnel. Consequently, the following criteria were established relative to determining contents of the instrument:

1. Little reading would be required in either administering or completing the instrument. Teachers usually have little time for completing tasks not directly related to instruction.

This is especially true at the beginning of the school year, the time when such an instrument could be most useful.

2. The instrument would require minimal time for scoring. Because of time restrictions during the school day, efficiency and simplicity in scoring is an essential quality. Further, such efficiency enhances the likelihood of continued use of the instrument.
3. The number of items contained in the instrument would be limited as well as the length of each item. Again, the time available to teachers is an essential consideration.
4. The instrument could be utilized by entire school systems as well as by individual teachers. The instrument would be designed with a focus on positive use, that is, school leaders could use the instrument to encourage more positive attitudes among teachers and as the basis for designing "affective" workshops for the benefit of the total community of learners.
5. The instrument would be valid and statistically reliable.
6. The instrument would be useful for research purposes. For example, there is an evident need for measuring attitudinal changes of teachers toward low achieving students in mathematics. The instrument could facilitate such research.

A comprehensive review of research indicates that no instrument exists combining the six criteria cited above. Such an instrument could be effectively and efficiently utilized by a variety of educators to promote both academic achievement and affective considerations.

Importance of the Study

A substantial body of research suggests that the attitude of the teacher toward students is the most critical factor influencing the classroom learning environment (Weiner, 1971). Not surprisingly, student motivation and performance levels in mathematics also appear to be profoundly influenced by teacher attitudes and behaviors (Haladyna, Shaughnessy & Shaughnessy, 1983; Kulm, 1980). That is, some teachers have demonstrated differential attitudes depending on whether students are perceived to have high or low ability in mathematics. For example, students perceived to be low achievers in mathematics might be seated farther away from the teacher, might receive fewer smiles, less eye contact, more criticism, less praise, and be interrupted more often when responding to instructor questions. Teacher attitudes can also be affected by stereotypic notions or biases toward groups. For example, teachers might expect students who are from lower socioeconomic status homes, from less educated families, or who are members of a particular minority group to do poorly in mathematics (Foster, Algozzine, Ysseldyke, 1980). Rosenthal (1973) reported that the academic performance of ghetto children worsens the longer they remain in school and that these children tend to have teachers who are convinced that the children cannot learn.

Good (1970, 1981) found many negative attitudes and unprofitable interactions occurring between teachers and students as the result of lack of awareness among teachers concerning their behavior. In other words, teachers might be unaware of their attitudes and behaviors toward those students perceived to be low achievers in mathematics.

Consequently, teachers, other school leaders, and the whole community of learners could, therefore, benefit from having access to an instrument designed to measure teacher attitudes toward low achievers in mathematics. Through the use of this instrument, teachers could become aware of negative attitudes and low expectations directed toward low achievers in mathematics and, as a result, engage in more appropriate and supportive behaviors. This contention is supported by conclusions reached by Pambookian (1976) and McNeil (1971) who determined that teachers are most likely to alter attitudes and behaviors when provided with information showing a discrepancy between what they think they are doing and what they are actually doing.

As mentioned above, mathematical achievement is a national concern. Since teacher attitudes play such a vital part in learning mathematics, an instrument is needed to measure teacher attitudes toward low achievers in that subject area. While this is a critical area for concern, there apparently is no measurement scale currently available. Consequently, the need for such a scale is essential.

Assumptions

1. Teachers will be honest in their responses. For an attitude measure to be valid, individuals taking the scale must be honest in their response to items. The assumption is accepted that teachers will respond truthfully based on a realization that the results of the measure might benefit them personally as well as the students they teach. Further, teachers will see no need to falsify responses when informed that results

will in no way be used as a way of evaluating their performance and no individual responses will be reported, only group data.

2. Attitudes are generally translated into behaviors. According to Good (1981), teachers have attitudes toward students leading them to expect specific behaviors and levels of achievement from particular students. Attitudes toward students vary; therefore, teachers behave differently toward different students. For example, Good further suggested that teachers smile less at low achievers than at high achievers.
3. The three basic assumptions relative to the measurement of attitudes proposed by Fishbein and Ajzen (1975) are accepted concerning this study : (1) an attitude is learned, (2) attitudes predispose action, and (3) actions based upon attitudes are consistently favorable or unfavorable toward the object.
4. Awareness of attitudes can effect change in behavior. Once aware of negative attitudes toward low achievers in mathematics at the middle school level and the probable differential treatment of students, teachers can effect change in those behaviors and begin to communicate more positively both in an academic sense and in the affective domain.

Limitations of the Study

Self-report measures of attitudes can only measure what individuals know about their attitudes and what they are willing to relate. However, Nunnally (1970) reported that self-report measures of attitudes and beliefs are much more believable than those asking for more complex self description inventories. Further, Nunnally determined that when used in a group situation, as long as anonymity of individual responses and scores is assured, self-report of attitudes should not be affected by a lack of frankness.

Delimitations of the Study

1. The sample involved in the study was restricted to middle school teachers of mathematics. Teachers were excluded if all or the major proportion of the students they teach were designated as "mentally retarded" or "emotionally disturbed."
2. Teachers who taught middle school mathematics in Washington, Carter, and Sullivan counties in northeast Tennessee provided data used for developing the final instrument. All the middle school mathematics teachers in the Knox County School System in Tennessee were invited to participate in providing data for determining validity and reliability for the final instrument, the Teacher Attitudes Toward Low Achievers in Mathematics Scale (TALAM).

Definitions

Attitude - a learned pre-disposition to respond in a consistently favorable or unfavorable manner with respect to a given object (Fishbein & Ajzen, 1975, pg. 6).

Attitude scale - an instrument which numerically measures the direction and intensity of an individual's attitude toward an object (Anastasi, 1976).

Attitude statements or items - statements made by middle school mathematics teachers, by experts in mathematics or mathematics education, or by researchers relative to teacher attitudes toward low achievers in mathematics.

Behavior - refers to overt behaviors of mathematics teachers toward low achievers in mathematics.

Beliefs - refers to opinions held by mathematics teachers about low achievers in mathematics. Often referred to as the cognitive component of attitude.

Feelings - refers to feelings of mathematics teachers toward low achievers in mathematics and feelings about working with low achievers in mathematics. Often referred to as the affective component of attitude.

Intended behaviors - refers to a mathematics teachers' intentions to perform various behaviors with regard to low achievers in mathematics. Such behavioral intentions are often referred to as conations.

Likert scale - a standard attitude scaling technique developed by Likert wherein respondents react to given statements by choosing one from among the following five response categories--(SD) strongly disagree, (D) disagree, (U) undecided, (A) agree, (SA) strongly agree.

Modified Likert scale - a modification of the Likert technique allowing for alternatives in the number of response categories and/or modifications in the statement format. The scale developed for this study contains six response opportunities for each attitude statement.

Perceived low achievers - those students perceived by mathematics teachers as performing at a low level of achievement in mathematics.

Perceived high achievers - those students perceived by mathematics teachers as performing at a high level of achievement in mathematics.

TALAM - a scale used to measure the attitudes of middle school mathematics teacher toward low achievers in mathematics: Teacher Attitudes Toward Low Achievers in Mathematics Scale.

Teacher expectations - judgements made by mathematics teachers about the present and potential mathematical achievement ability of their students.

Organization of the Study

This study is organized into five major chapters.

Chapter I provides an introduction to the study and contains the following components: (1) background for the study, (2) statement of the problem, (3) purpose of the study, (4) importance of the study, (5) assumptions, (6) limitations of the study, (7) delimitations of the study, (8) definitions and, (9) organization of the study.

Chapter II contains a survey of the literature relevant to the study and contains the following components: (1) definitions and components of attitudes, (2) the measurement of attitudes, (3) selected techniques of attitude scale construction, (4) methods for estimating reliability and validity and, (5) attitude scales related to mathematics.

Chapter III presents the methods and procedures used in the study in the three major phases of instrument development: Phase 1 - readiness, Phase 2 - administration of the preliminary scale and, Phase 3 - administration of the final scale.

Chapter IV contains an analysis of data obtained in the three phases of instrument development mentioned above.

Chapter V provides a summary of the study, conclusions, recommendations for uses of the scale, and recommendations for future research.

CHAPTER II

REVIEW OF LITERATURE

The literature was examined to find studies specifically concerned with the development of instruments designed to measure attitudes. Close attention was paid to procedures used to determine reliability and validity in the development and subsequent use of such instruments. An effort was made to select the most relevant literature from among the vast number of studies undertaken to determine attitudes toward a great many objects. This review is organized into the following categories: (1) Definitions and components of attitude, (2) The measurement of attitude, (3) Techniques for attitude scale construction, (4) Test construction statistics and, (5) Related attitude scales.

Definitions and Components of Attitude

Definitions

Among the more commonly accepted definitions of attitude are the following:

> An attitude is a mental and neural state of readiness, organized through experience, exerting a directive or dynamic influence upon the individual's response to all objects and situations with which it is related.
>
> (Allport, 1935, p. 810)

> An attitude can be defined as an enduring organization of motivational, emotional, perceptual, and cognitive processes with respect to some aspects of the individual's world.
>
> (Krech and Crutchfield, 1948, p. 152)

> An individual's social attitude is a syndrome of response consistency with regard to social objects.
>
> (Campbell, 1950, p. 31)

> An attitude is an idea charged with emotion which predisposes a class of actions to a particular class of social situations.
>
> (Triandis, 1971, p. 2)

The definition of attitude proposed by Triandis suggests that attitude has three components: (a) a cognitive component (the idea), (b) an affective component (the emotions), and (c) a behavioral component (the action). Further, the definition of attitude proposed by Krech and Crutchfield (1948) described above also implies three similar components of attitude. A discussion concerning components of attitude is presented below:

Components

The cognitive component of attitude was described by Triandis (1971) as the ideas or beliefs that subjects have about an attitudinal object, the object, in this context, being the focal point of attention. The affective component was described as the emotions or feelings about the attitudinal object, while the behavioral component was described as predisposition to action with regard to the same object. Triandis found the behavioral component measurable though direct observation of overt actions or through analysis of verbal statements concerning intended behavior. Triandis further indicated that, although the three components are closely related, the components can appear to be inconsistent with one another based on overall analysis of attitude scale responses from individuals.

Other researchers referred to three similar subcomponents of attitude and recommended attitude measurement approaches reflecting those subcomponents. In this light, Hassan and Shrigley (1984) categorized attitude scale components as (1) egocentered, (2) social-centered and, (3) action-centered. The three item types suggested by Hassan and Shrigley appear similar to the affective, cognitive, and behavioral components of attitude described by Triandis (1971) and Oskamp (1991). Likewise, Chein (1950) discussed attitudes in terms of three components. In a similar approach, Greenwald (1968) also described the three components of attitude as "affects, cognitions, and action tendencies" (p. 363).

Fishbein and Ajzen (1975) suggested a classification of four components of attitude rather than the more commonly used three. While maintaining affect (feeling), and cognition (belief), Fishbein and Ajzen divided the behavioral component into two parts: the actual behavior (observed overt acts) and the conation (behavioral intentions). Further, these researchers concluded that if attitude must be measured as a single dimension and reported in a single score, it is most accurately measured through the affective part of the attitude concept. The last contention of Fishbein and Ajzen is consistent with the apparent widespread agreement among researchers that, although affect cannot capture the full complexity of the attitude concept, it is the most essential, consistent, stable and reliable measure of attitude.

The Measurement of Attitudes

Information about attitude can be gathered in two basic ways: through observing subjects and/or by asking subjects what they believe. In this light, Anderson (1981) stated that information is gathered about attitude or any affective characteristic though observational methods and/or through self-report methods. The purpose of this section is to present information about observational and self-report methods of attitude assessment and to highlight advantages and problems inherent in each.

Observational Methods

Using observational methods for obtaining information about attitude is based on the assumption that it is possible to infer attitude from the observation of overt behavior or physiological reactions (Fox, 1969; Anderson, 1981). Three major problems are reportedly inherent in observational research methodology:

1. The problem of inaccurately inferring affective characteristics from overt behavior.

2. The problem of determining which behaviors to observe and how to accurately record those behaviors.

3. The problem of misinterpreting the behavior noted by the observer.

Anderson (1981) proposed potential solutions to the three problems inherent in observational methods of obtaining information about attitude. For the first problem related to making inferences, Anderson suggested that correct inferences are more likely to be made if multiple observations are made of the same behavior in a variety of settings or over time in the same setting. With regard to problem number two, observing relevant behaviors, Anderson suggested that appropriate inferences can be made if the affective characteristics are clearly defined at the outset and care is taken to observe only those clearly defined behaviors in an appropriate context. With regard to the third problem, that of misinterpreting behaviors, Anderson suggested using more than one carefully trained observer in the same setting to minimize misinterpretation.

Purkey, Cage, and Graves (1973) assessed affective characteristics of 357 students at two elementary schools. The researchers designed a measure they called the Florida Key. Teachers of the 357 subjects were asked to evaluate their students based on observations of the students behaviors. In the Florida Key, 18 behaviors were designated for evaluation and subsequent analysis. While the researchers reported only a modest relationship between the Florida Key and a self-report measure of affective characteristics, the study presents an interesting research design pairing observational research with quantitative research methodology. Further descriptions and presentations of data concerning data collected through observing subjects are presented by Cook and Sellitz (1964), Lemon (1973), and Crano and Brewer (1973).

The measurement of attitude through observation of physiological reactions was studied by Porier and Lott (1967), Westie and DeFleur (1959), Woodmansee (1970), and Mueller (1970). Such techniques are based on the assumption that there is a close relationship between physiological responses and affective states. Researchers noted that autonomonic responses might function as valid indicators of strong attitude but might be insensitive to less extreme attitudinal reactions. Further, researchers generally have noted that the ability to determine the directionality of response through analysis of physiological reactions is extremely limited. The two main types of physiological responses discussed in the literature are the Galvanic Skin Response (GSR), a calculation of the amount of electrical conductance of the skin, and pupillography, a measure of change in reaction of the pupil in the eye to various attitudinal stimuli.

Self-Report Methods

Self-report methods of attitude assessment are usually a series of questions, adjectives, or statements about an attitudinal object. Respondents are asked to read and react to each question, adjective, or statement about an attitudinal object in terms of agreement or disagreement. Responses are then scored in terms of positiveness toward the attitudinal object. In some instances, responses are summed to attain a total score.

According to Anderson (1981), the major difficulty associated with self-report methods of attitude assessment is that subjects may provide misinformation to the researcher. Anderson contends that misinformation is sometimes supplied to the researcher when individuals respond to a question, statement, or adjective in a way they think will be socially acceptable to the researcher or when they respond in an acquiescent manner. Acquiescence, in this instance, refers to the tendency of an individual to agree with a question, statement or adjective when they are actually unsure of their response. Thurstone and Chave (1929) considered the issue of misinformation and offered the following advice to researchers:

> All that we can do with an attitude scale is to measure the attitude expressed with the full realization that the subject may be consciously hiding his true attitude or that the social pressure of the situation made him really believe what he expresses. . . . All we can do is minimize as far as possible the conditions that prevent our subjects from telling the truth, or else to adjust our interpretation accordingly.
>
> (pg. 10)

Selected Techniques of Attitude Scale Construction

The four major types of attitude scales described in the literature were: Thurstone scales (Thurstone and Chave, 1929); Likert scales (Likert, 1932); Guttman scales (Guttman, 1944); and semantic differential scales (Osgood, Suci, and Tannenbaum, 1957). An overview of each of the four attitude measurement scale types is presented below:

Thurstone Technique

Thurstone and Chave (1929), developed the method of equal-appearing intervals to measure attitudes. According to Thurstone, the essential characteristic of the method of equal-appearing intervals is the series ". . . of evenly graduated opinions so arranged that equal steps or intervals on the scale seem to most people to represent equally noticeable shifts in attitude" (pg. 554). Edwards (1957) reported on the usefulness of the method of equal-appearing intervals especially when a large number of statements must be scaled. Edwards further described the method of equal-appearing intervals as much preferable to the earlier more laborious paired-comparison technique of attitude scale construction also introduced by Thurstone in 1927.

Procedures. Using the method of equal-appearing intervals developed by Thurstone and Chave (1929), opinions about an attitudinal object can be collected from designated samples and from related academic literature. The collected opinion statements about the object of focus can then be edited. The editing process is undertaken to select statements covering the widest possible range from the most intensely negative to the most intensely positive attitudes toward the object. The selected items are each printed on a separate slip of paper and subjects (sometimes called "judges" in the literature) are given a copy of each item.

The subjects are asked to sort the items into 11 piles representing an evenly graduated series of attitudes from extremely negative (pile 1) through extremely positive (pile 11) toward the attitudinal object. After sorting, data are tabulated to show how each subject placed every one of the statements. Figure 1 shows the method used by Thurstone and Chave to summarize the sorting of items by subjects. The first column gives the item number. The second and third column contain, respectively, the scale value and the Q value (see "Scale and Q Values" below). The remaining columns, progressing from left to right, give the cumulative frequency of times the specified item was placed in each pile by subjects.

Item #	Scale Value	Q	A 1	B 2	C 3	D 4	E 5	F 6	G 7	H 8	I 9	J 10	K 11
1	9.9	2.4	.00	.00	.00	.00	.00	.08	.17	.23	.33	.52	1.00

Figure 1. Hypothetical example of how an item is sorted using the equal-appearing intervals technique.

Scale and Q Values. A scale value for each item was determined graphically by Thurstone and Chave (1929). Considering each item separately, the cumulative proportion of responses to an item (Y axis) was plotted against the corresponding sorted scale values for the same item (X axis). An overall scale value for the item was then determined by locating the item's median scale value.

After, Thurstone and Chave located the scale value and the upper and lower quartile response scores of each item, an overall Q value was then determined for that item. The Q value was calculated by subtracting the lower quartile score from the upper quartile score. The Q value was considered to be a measure of ambiguity and also a measure of dispersion of judgments for an item. If the dispersion of judgements for a statement is high in comparison with other statements, the statement would be considered ambiguous and eliminated from consideration on the final scale. The Q value has also been referred to in the literature as the semi-interquartile range (Guilford, 1965). Guilford defined the semi-interquartile range as one-half the range of the middle 50 per cent of judgements about an item: $Q = (C_{75} - C_{25})/2$.

The Final Thurstone Scale. After considering and comparing the scale and Q value of every item, and after giving logical consideration to the content of every item, Thurstone and Chave (1929) selected items for inclusion on their final attitude scale. The statements selected approximated as closely as possible a uniformly graduated series of scale values. The scale was then presented to subjects who were asked to place a check mark beside each statement with which they agreed.

Thurstone and Chave (1928) described two methods of scoring the equal appearing interval attitude scale. The first method involved summing the scale-values of all items with which a subject agreed and then obtaining the arithmetic mean of those scale values. The second method of scoring a subject's set of responses consisted of assigning a numerical rank to each of the items on the scale. The rank values of all items with which a subject agreed were summed and an arithmetic average determined.

Likert Scales

Likert scales are an extremely popular method for measuring attitude (Oppenheim, 1966; Crano & Brewer, 1973; and Anderson, 1981). The researchers cited above suggested that the Likert method of scale construction is less laborious than the Thurstone technique. Further, the researchers suggested that it is the most efficient and effective method of developing highly reliable scales.

The Likert Scale was developed by Rensis Likert (1932). Likert's primary concern for such a scale was that it measure a unidimensional construct, that is, that all items measure the same thing. Edwards (1957) and Sellitz, Jahoda, Deutsch, and Cook (1959) referred to the

Likert scaling technique as the method of summated ratings because the total score for each subject is obtained by summing the subject's response to each item. The summated score, therefore, represents the degree of favorable or unfavorable attitude toward the object under consideration. Components and strategies for developing Likert scales are presented below:

Procedures. Items should be clearly favorable or unfavorable with regard to the attitudinal object. Likert (1932) determined it desirable to prepare and select more statements than are likely ever to be used, since many of the items would be found unsatisfactory for the intended purpose of an instrument. Years later, Lemon (1973) suggested using approximately the same number of positive and negatively stated items in a Likert scale. However, other researchers, including Hassan and Shrigley (1984), favored using more negative than positive statements because negatively stated items "are less prone to withstand the rigor of Likert's item analysis" (pg. 660).

After preliminary items on the Likert scale have been written, several judges are asked to classify each item as positive, negative, or neutral with regard to the attitudinal object. Items not classified by the majority of judges as either positive or negative with regard to the attitudinal object are eliminated from consideration for use in the final scale.

A decision must be made relative to the number of response alternatives for each statement. Likert originally used a five response format: (1) strongly disagree, (2) disagree, (3) undecided, (4) agree, and (5) strongly agree. However, modifications in the number of response alternatives are acceptable. A number of response categories ranging from two to seven are described by Anderson (1981) with the even numbered categories yielding a forced choice i.e. no neutral response is possible. Anderson further suggested increasing the number of response categories as a means to strengthen reliability.

The self-report instrument is then administered to a sample of the audience for whom the instrument is intended. Data are analyzed to estimate the validity and reliability of the scale. A revised final scale is then constructed based on conclusions drawn from the data.

Scoring. The respondent is asked to react to each item in terms of several degrees of agreement or disagreement; for example, (1) strongly agree, (2) agree, (3) undecided, (4) disagree, and (5) strongly disagree. The response alternatives are weighted so the most favorable response carries the highest weight. For example, if a statement is favorable regarding the attitudinal object, "strongly agree" carries the highest weight. On the other hand, if the statement is unfavorable toward the object, then "strongly disagree" carries the highest weight. Consequently, when scoring, the tallies on negative items would be reversed.

Likert's original method of weighted scoring (Edwards, 1957) was based on Likert's conclusion that a normal distribution often results when the five point response system is used. Likert determined the proportion of subjects falling into each of the five response categories for a favorable statement and then calculated the corresponding normal deviate weights i.e. Z score for each item. The overall score was obtained by summing the weights for all items. As mentioned above, the weights were reversed for unfavorable statements so that the strongly disagree category had the highest positive weight for those negative items.

Likert (1932) also devised a less complex method for assigning weights to the five response categories by eliminating the need for Z score transformation. In the simpler system, for favorable items, Likert assigned the "strongly agree" response a weight of 4, the "agree" response a weight of 3, the "undecided" response a weight of 2, the "disagree" response a weight of 1, and the "strongly disagree" response a weight of 0. For unfavorable items, the scoring was reversed. For each respondent, a total score was then obtained by summing all scores for all items.

Item Selection Criteria. The criterion of internal consistency is commonly used as a method of selecting items for inclusion on a final Likert scale (Likert, 1932; Ferguson, 1981; Crano & Brewer, 1973; Anderson 1981). The criterion of internal consistency is applied by correlating item scores with total scores. Any item with a nonsignificant item to total correlation is eliminated from consideration for use in the final scale. Researchers agree that high correlations between scores on a particular item and total test scores suggest the item represents the attitude under study.

According to Hassan and Shrigley (1984) and Edwards (1957), another test of the validity of a particular item is the discriminating quality of the item. A positively written item is valid only if those individuals with a generally positive attitude toward the attitudinal object agree or strongly agree with the item and if those with a generally negative attitude disagree or strongly disagree with the item. The researchers cited above suggested establishing positive and negative criterion groups composed of subjects having the highest and lowest 27% of scores within the overall group being considered. Student t scores would then be calculated comparing the mean score for each criterion group. A significant difference in the mean scores of the two criterion groups would suggest that the item has discriminating quality.

Guttman Scales

Guttman (1944) and Guttman and Suchman (1947) developed what Edwards (1957) suggested is more a procedure for evaluating a set of statements about an attitudinal object rather than an actual attitude scale. Nevertheless, the procedure has become known throughout the literature as the Guttman Scale. A description of the Guttman procedure follows.

Procedures. In constructing a Guttman scale, according to Crano and Brewer (1973), statements appearing to have the following characteristics are written or selected:

1. Statements have common content
2. Statements are ordered along a continuum from least positive to most positive
3. Agreement with a given statement implies agreement with every other less positive statement.

Given an instrument with statements about an attitudinal object meeting the criteria described above, subjects are then instructed to check each statement with which they agree. When a subject agrees with an attitude statement, the subject receives a score of 1 for the item. However, if the individual disagrees with the attitude statement, the subject receives a score of 0 for the item. The subject's total score is the sum of all his/her item scores on the scale. The overall score suggests the subject's degree of favorability toward the attitudinal object. Data are then submitted to Guttman scale analysis.

Scalogram analysis. Guttman scale analysis involves the computation of the coefficient of reproducibility (CR) (Guttman, 1944, 1947). The calculation of the coefficient of reproducibility is illustrated through the following hypothetical example: A scale with five statements is administered to a group of subjects. The five statements were written along a continuum from least positive (statement number one) through most positive (statement number five). If the Guttman assumption is met, several potential response patterns are acceptable. All acceptable response patterns illustrate that a person who agrees with statement number five (the most positive), must also agree with statements one through four. Through analysis of all response patterns, the number of errors due to inappropriate responses can be calculated. Figure 2 illustrates an acceptable response pattern for a five item Guttman scale.

Acceptable Pattern	Statement Numbers 1	2	3	4	5
A	D	D	D	D	D
B	A	D	D	D	D
C	A	A	D	D	D
D	A	A	A	D	D
E	A	A	A	A	D
F	A	A	A	A	A

Figure 2. Acceptable response patterns for a five item Guttman scale.

In calculating the coefficient of reproducibility (CR), the total number of errors (deviations from acceptable patterns) is counted for all subjects. A percentage of error is then computed by dividing the total number of errors made by all subjects by the total number of responses. The total number of responses refers to the number of subjects multiplied by the number of statements. The CR is then obtained by subtracting the error rate from 100 percent. For example, suppose 25 subjects were administered a five item Guttman scale. The total number of potential responses is 25 X 5 = 125. If the total number of errors made by all the subjects is 15, then the error rate is 15/125 = .12 or 12 %. The CR calculation: 100% - 12% = 88%. Guttman (1944) suggested that the error rate should be no larger than 10% for the set of statements to be considered an acceptable scale.

The Cornell technique (Guttman, 1947) and the Goodenough technique (Goodenough, 1944) are two prominent methods of scalogram analysis. Both the Cornell and Goodenough scalogram methods calculate the percent of accuracy that the data obtained from responses to a Guttman attitude scale can be reproduced from the total scores. For example, if the coefficient of reproducibility for a scale is .88, this means that 88% of the subjects' responses could be predicted from knowledge of total test scores alone.

Scalogram analysis can also be generalized to more than two categories of response. For example, three categories of response can be used such as: agree, undecided, and disagree with weights of 2, 1, and 0 assigned, respectively. A more comprehensive description of these procedures is found in Edwards (1957).

Semantic Differential Scale

The semantic differential technique was introduced by Osgood, Suci, and Tannenbaum (1957) for measuring attitude. This technique is adjective based and measures reactions of subjects to pairs of bi-polar adjectives with meanings as nearly opposite as possible (Osgood, 1952). Examples might include: good-bad, happy-sad, etc. The semantic differential (SD) measures directionality of a reaction and also intensity of reaction (Osgood & Suci, 1955). Heise (1967) reported that ratings on SD scales tend to be correlated around three basic dimensions of response accounting for most of the covariance in ratings: evaluation, potency, and activity (EPA). SD scales generally contain adjectives from all three dimensions. Examples of EPA types might include: Evaluation - good/bad, Potency - hard/soft, Activity - fast/slow. Lists of evaluative adjective pairs are included in a text by Osgood, Suci, and Tannenbaum (1957).

Procedures. In constructing a semantic differential scale, the name of the attitudinal object is placed at the top of the scale. Then, 5 to 10 emotion laden adjective pairs are chosen and a response sheet is constructed. The bi-polar adjective pairs are placed at different ends of a numerical continuum of seven equal segments. Figure 3 illustrates an example of a semantic differential response sheet.

```
           High Risk High School Students
     Good ___: ___: ___: ___: ___: ___: ___ Bad
Dishonest ___: ___: ___: ___: ___: ___: ___ Honest
     Weak ___: ___: ___: ___: ___: ___: ___ Strong
     Fast ___: ___: ___: ___: ___: ___: ___ Slow
```

Figure 3. Example of a Semantic Differential Scale.

After adjectives are selected and a response scale is constructed, it is administered to a sample. Subjects are instructed to place a check mark along the continuum at the point best describing how they feel about the object presented at the top of the sheet. A check mark near either end of the continuum indicates strong positive or negative feelings, while a center check mark indicates neutral feelings. Positive integer values of one through seven are assigned to each response option with the most favorable attitude toward an object given a weight of seven. The total score on the scale is the sum of the subject's response to each item.

Analysis of data. Analysis of data obtained through administration of the semantic differential scale is similar to analysis of data obtained from a Likert scale. Correlations between each adjective pair and the total scale score can be computed. Adjective pairs not correlating significantly with the total scale score are eliminated.

A comprehensive description of various statistical procedures available for processing data obtained from administration of a semantic differential scale is contained in a review of related research by Heise (1970). Further, reviews, methodological studies, and validity studies related to the semantic differential technique are found in Snider and Osgood (1969).

Methods for Estimating Reliability and Validity

Methods for estimating reliability and validity of tests will be discussed in this section. In part one, the concept of reliability and an overview of computational procedures related to calculating reliability will be reviewed. In part two, an overview of the types of validity and statistical procedures for calculating validity coefficients will be presented.

Reliability

The reliability of a test indicates the trustworthiness of scores obtained. The reliability of a test is an expression of both the stability and consistency of test scores (Cureton, 1951, 1965; Thorndike, 1966; Dick & Haggerty, 1971). Concerning stability, researchers determine whether the score obtained for a subject (S1) would be the same if S1 were tested again at a later date. The reliability coefficient then indicates whether the two test scores for S1 are stable indicators of S1's performance. Researchers also consider whether the reliability coefficient estimates the accuracy of S1's true score.

A reliability coefficient is represented by a numerical value between 0 and 1 reflecting the stability of the instrument. To compute reliability coefficients, four basic methods are generally used (Ferguson, 1981):

1. Test-retest method - the same test is administered twice to the same group of subjects with administrations separated by an interval of time.

2. Parallel-forms method - an alternative test form is administered to the same group after a period of time.

3. Split-half method - A test is divided into two parts and two scores are obtained. The paired observations are correlated.

4. Internal-consistency methods - based on the average correlation among items and the number of items on a test.

In all four of the basic methods mentioned above for approximating reliability, the calculation of correlation coefficients between paired observations is required. Many varieties of correlation have been developed for use with different types of variables and for data with special characteristics. An overview and discussion of all correlation coefficients is beyond the scope of this dissertation. However, a few of the more widely used measures of correlation will be briefly presented.

Product-moment correlations. The test-retest and alternate forms methods of estimating reliability are determined based on correlating two sets of test scores. Alternate formulas, derived from standard score form, exist for computing product-moment correlations between test scores. The most widely used product-moment correlation coefficient is the Pearson correlation coefficient (Ferguson, 1981). One form of the Pearson Product Moment, denoted by r, follows:

$$r = \frac{(\Sigma XY) - n\,\overline{X}\,\overline{Y}}{(n-1)\,s_x s_y}$$

where n is the number of cases, $\overline{X}$, $\overline{Y}$ are the means, and s_x and s_y are the standard deviations of the two variables.

The split-half method of reliability estimation requires the use of an additional formula (Guilford, 1965; Ferguson, 1981). As mentioned above, in the split-half method the same test is divided into two parts and the scores are correlated. The result is a correlation between scores on tests having half as many items as the original instrument. For example, on a 20 item test, 10 of the items would be correlated with the 10 other items with each set of correlated items having similar content. In effect, correlation would occur between paired scores based on scores from two 10 item tests. However, the reliability for the total 20 item test is needed. Therefore, the use of the Spearman Brown (SB) formula approximates the reliability for the total test. One form of the Spearman Brown formula (Ferguson, 1981) is shown below:

$$r_{tt} = \frac{n r_{11}}{1 + (n-1)\, r_{11}}$$

Where, n is the ratio of the number of items on the desired test to the number of items on the original test and r is the already obtained reliability for the partial test. The Spearman-Brown formula can also be utilized to estimate reliabilities obtained by the test-retest and alternate forms methods.

Kuder-Richardson. An internal-consistency measure commonly used to estimate reliability was derived by Kuder and Richardson (1937). The two assumptions underlying use of Kuder-Richardson formulas are: (1) the items are dichotomously scored, that is, items are scored 1 for a correct response and scored 0 for an incorrect response; and, (2) the items are unidimensional since they measure the same characteristic.

There are many ways a test can be split in order to compute half-test scores. For each possible split, a different reliability coefficient can be obtained. The Kuder-Richardson formula averages all the possible split half reliability coefficients of a particular test. The basic Kuder-Richardson formula (Guilford, 1954; Ferguson, 1981), referred to as formula 20 or KR-20, is shown below:

$$r_{tt} = \frac{k}{(k-1)} \quad \frac{\sigma^2{}_o \ - \Sigma p_i q_i}{\sigma^2{}_o}$$

where, k is the number of items in the test; p = the proportion of students responding correctly to item i; q = 1 - p, the proportion of students responding incorrectly to item i; σ^2_o = test variance, and $\Sigma p_i q_i$ = sum of p times q for all items.

When individual item statistics are not available, an alternative Kuder-Richardson formula can be used to give a conservative estimate of test reliability (Kuder & Richardson, 1937; Guilford, 1954; Ferguson, 1981). It is reasonable to assume that all test items have approximately the same level of difficulty; therefore, the term pq in the KR-20 formula can be replaced by kpq in the alternative Kuder-Richardson formula, where k is the number of test items.

A special case of the Kuder-Richardson formula, is Cronbach's coefficient alpha (α) (Cronbach, 1951). Coefficient alpha is the basic formula for determining the reliability of test scores based on internal consistency for items not dichotomously scores (Nunnally, 1967). According to Cronbach (1951), the coefficient alpha (α) is the mean of all possible split-half coefficients which can result from different splittings of a test and can be used as an index of inter-item homogeneity..

Validity

Test validity is an indication of how well a test measures what it was designed to measure (Garrett, 1947; Mehrens & Lehmann, 1980). Validity is always stated in reference to a given group, a given area, or a given circumstance. A test can be valid for one group but inappropriate for another. Validity involves gathering and evaluating information for determining how well a test measures what its authors purport it measures. Other definitions and discussions of validity can be found in works by Lindquist (1942), Guilford (1946), Cureton (1951), and Anastasi (1976).

Types of validity. Although there are many procedures for determining validity, all aspects of validity are interrelated. Types of validity usually considered when instruments are developed for measuring psychological traits are: (1) content, (2) concurrent, (3) construct, and (4) predictive (Wainer & Braun, 1988). Some of the other types of validity mentioned in the literature are: (1) face, (2) curricular, and (3) differential. The specific approaches for determining validity listed above will be described in the section that follows.

Content validity. The following definition of content validity was offered by the American Psychological Association (1966, p. 12):

> The test user wishes to determine how an individual performs at present in a universe of situations that the test situation is claimed to represent.

If test items are to have content validity, items should be representative of the characteristic being measured. For example, if teacher attitude toward low achievers in mathematics at the middle school level is being measured, items should be written based on middle school teachers' comments about low achievers in mathematics, on other scales measuring the same characteristic, or on relevant items found in the literature. In this way appropriateness of test content can be determined.

Predictive and concurrent validity. In describing predictive validity the American Psychological Association (1966, pg. 12) stated:

> The test user wishes to forecast an individual's future or to estimate an individual's present standing on some variable of particular significance that is different from the test.

When tests correlate highly with subsequent performance, the tests are said to have predictive validity. Validation of this type sometimes takes a long period of time. For example, the ACT mathematics scores of high school juniors might have predictive validity for grade point average in college freshman mathematics classes. There is no way to determine whether it does other than to wait and see how the subjects perform in college.

Concurrent validity, sometimes termed "immediate predictive validity," correlates a test in the process of being developed with scores obtained from previously established measures. For example, in establishing concurrent validity for an instrument designed to measure mathematics anxiety of pre-service elementary school teachers, a researcher might choose to correlate scores obtained on this measure with scores obtained from the same individuals taking the previously established Mathematics Anxiety Rating Scale (MARS) (Suinn, 1972). By obtaining a significant positive correlative between scores obtained on the two measures, researchers could infer that the anxiety scale written for pre-service teachers does indeed appear to measure mathematics anxiety.

Construct validity. In defining construct validity, the American Psychological Association (1966, pg. 12) stated:

> The test user wishes to infer the degree to which the individual possesses some hypothetical trait or quality (construct) presumed to be reflected in the test performance.

Construct validity involves formulating a theory of relationships and cannot generally be expressed in terms of one coefficient. Cronbach (1959) contend that the following types of evidence, among others, must be taken into consideration when attempting to achieve construct validity: content validity, interitem correlations, intertest correlations, studies of stability over time and after experimental intervention.

Face validity. This type of validity merely answers the question, "Does the test appear to measure what it purports to measure"? For example, the Math Anxiety Rating Scale (MARS) (Suinn, 1972) appears from the name of the instrument and a perusal of items therein to measure what it was designed to measure, mathematics anxiety.

Curricular validity. Cronbach (1960) introduced the term "curricular validity." This type of validity requires determining if tests are representative of instructional content and reflect goals of instruction. For example, the mathematics teacher who is concerned with students' achievement of specific objectives would make certain that his/her test measures those same objectives.

Differential validity. Anastasi (1986) defined differential validity as the difference between two correlation coefficients when one measure is correlated with two different measures This procedure is undertaken to determine what a test measures and what it does not measure. For example, as a classification test, an honors level high school calculus achievement test might be administered to all students in the honors calculus class. The results of the classification test could then be correlated with two separate criteria: (1) a test of creative ability and (2) a test of mechanical ability. If the classification test correlates .11 with the creative ability test and

.92 with the mechanical ability test, then the differential validity of the classification test would be .92 - .11 = .81.

<u>Computational Procedures</u>. In the preceding section entitled "Reliability," several methods were given for approximating the reliability of a test. Whether using statistical methods applicable to reliability established through the use of alternate forms, test-retest, split-half, or internal-consistency reliability, the correlation coefficient given was obtained through correlating a test in some manner with itself. Correlations can also approximate validity coefficients. When statistical procedures correlate a test (x) and some other external criterion (y), such as another test, then they become calculations of validity coefficients. Statistical procedures for calculating validity coefficients and considerations concerning the choice of statistical procedures are found in works by Ferguson (1981), Guilford (1965), Wainer and Braun (1988), Edwards (1972), Nunnally (1967), Guilford (1954) and, Mehrens and Ebel (1967).

Another procedure, factor-analysis, has been suggested by researchers as a useful indicator of the construct validity of scales (Oppenheim, 1966; Hassan & Shrigley, 1984; Gorsuch, 1974; and Mulaik, 1972). Through the use of factor analysis, researchers can test how well statistical clusterings of items match the intended construct groupings. The clusters of items that appear as a result of factor analysis can be examined to determine if they represent the component or subcomponents of the attitude under study.

Innovations. The Mantel-Haenszel procedure was proposed as a "practical and powerful way to detect test items that function differently in two groups" (Holland, 1985, pg. 129). This statistical application can be used to shed light concerning the effect of experiential background relative to subject reaction to test items. Similarly, other researchers have conducted studies relative to what has become known in the literature as differential item functioning (Thissen, Wainer, & Williams, 1984). Methodologies described by the researchers cited above are designed to investigate methods of locating test items likely to be responded to differently based on the characteristics of groups setting them apart from others.

Meta-analysis is another statistical innovation in validity assessment. In relationship to validity, meta-analysis is concerned with quantitative methods for combining evidence from different studies. Wainer and Braun (1988) presented information from a variety of sources concerning the calculation and merits of meta-analysis, including the empirical Baysian approach.

Attitude Scales Related to Mathematics

Analysis of the literature suggests a vast array of studies undertaken to determine attitudes among a variety of samples concerning countless areas of interest. On the other hand, the comprehensive review of the literature has not produced evidence of any substantial study in the realm of measuring teacher attitudes toward low achievers in mathematics, the focus of this study. Therefore, the attitudinal instruments presented in this section relate to the measurement of affective attributes related to mathematics.

Mathematics Anxiety

The Fennema-Sherman Mathematics Attitude Scales (1976) consist of a group of five instruments: (1) Mathematics Anxiety Scale, (2) Attitude Toward Success in Mathematics Scale, (3) Effectance Motivation in Mathematics Scale, (4) Usefulness of Mathematics Scale and, (5) Confidence in Learning Mathematics Scale. The Fennema-Sherman scales are designed for administration to high school students. Item responses for the five tests are obtained on a four point Likert scale. Each test consists of 12 items, half of which are positively worded while the other half are negatively worded. Split-half reliability for the five tests were given by the researchers with coefficients ranging from .87 to .93. The Fennema-Sherman studies were innovative in the suggestion that a psychological trait such as mathematics anxiety, might be a multi-dimensional construct. Investigators found relatively low intercorrelations among test scores obtained through administration of the five instruments mentioned above. The researchers, therefore, concluded that each scale measured a different construct.

In a factor-analytic study of mathematics anxiety, conducted by Ling (1982), the five Fennema-Sherman scales (1976) were administered to 500 college freshman in mathematics courses. In addition to the five Fennema-Sherman scales, subjects were also administered the Short-Form Dogmatism Scale (Troldahl & Powell, 1965), The Adjective Check List (Gough, 1952), and the Test Anxiety Inventory (Spielberger, 1978). The study was designed to investigate mathematics anxiety and the possibility that it might be a multi-dimensional construct related to a variety of personality characteristics. However, after analysis of data, Ling (1983) concluded that mathematics anxiety is a unidimensional construct strongly related to attitude toward mathematics in general but not related to other personality characteristics represented by the instruments administered in the study.

Richardson and Suinn (1972) developed the Mathematics Anxiety Rating Scale (MARS). The scale consists of 98 items describing situations producing varying levels of anxiety to numbers. In the original study, 397 secondary level students responded to the items in the scale. An internal-consistency measure yielded a coefficient alpha of .97, while a test-retest procedure yielded a reliability coefficient of .85. In additional studies, a numerical ability measure was compared with the MARS, producing correlation coefficients suggesting that high levels of mathematics anxiety appear to interfere with achievement in mathematics.

Sandman (1974) developed the Mathematics Attitude Inventory (MAI) designed to measure several constructs related to mathematics: (1) Anxiety Toward Mathematics, (2) Value of Mathematics in Society, (3) Self-Concept in Mathematics, (4) Enjoyment of Mathematics, (5) Motivation in Mathematics, and (6) Perception of the Mathematics Teacher. The total scale contains 48 Likert items with each of the above mentioned subscales represented by eight items. Factor analysis of data obtained from 2,547 eighth and eleventh grade students provided support for the validity of the subscale constructs represented in the total scale.

Attitudes Toward Mathematics

Aiken and Dreger (1961) developed the Math Attitude Scale and Aiken (1974) the Revised Math Attitude Scale (1974). There are 20 items on the scale written in a Likert format with 10 of the items stated positively and 10 stated negatively. In the original study (Aiken & Dreger, 1961), application of the test-retest procedure yielded a reliability coefficient of .94. In the Aiken and Dreger studies, the Math Attitude Scale was correlated with instruments designed to measure achievement in mathematics, experience with mathematics, and other personality variables. Researchers concluded that attitude toward mathematics appears to be related to achievement and ability in mathematics but not to temperament or other personality variables represented by instruments in the study.

The Dutton Scale (Dutton, 1954) was originally designed as a Thurstone type scale measuring feelings toward arithmetic. The scale was comprised of twenty-two statements with scale values ranging from 1.0 to 10.5 divided equally between favorable and unfavorable statements. In 1954, the test was administered to 289 education majors yielding a test-retest reliability coefficient of .94. The scale was revised by Dutton (1962) and its length reduced to 15 items. With a sample of 127 education majors, the test-retest reliability coefficient on the revised measure was .94. Dutton and Blum (1968) changed the scale again, this time to a Likert format having 25 items. The sample in the later study consisted of 346 middle school pupils from four socioeconomic groups. The Dutton-Likert scale yielded a split-half reliability coefficient of .84.

Gladstone, Deal, and Drevdahl (1960) developed a 12 item, modified Likert-type scale for use in studying the effects of remedial mathematics courses on attitude. The items were designed to measure attitudes toward mathematics as compared to attitudes toward other school subjects. No reliability estimates were found for the scale. However, some evidence of predictive validity of the scale items related to subjects' dispositions toward mathematics were found.

Aiken (1974) constructed scales designed to measure enjoyment of mathematics (E Scale) and the value of mathematics (V Scale). The scales were combined into a 40 item Likert-type scale and administered to 190 college freshmen. The internal-consistency reliability, coefficient alpha, for the instrument was found to be .95 for the E Scale and .85 for the V Scale. The correlation coefficient between the E and V scores was .64.

A Mathematics Attitude Inventory was constructed in two forms by Ellingson (1962) using Thurstone's method of equal appearing intervals. The two equivalent forms of the inventory, containing 25 items each, were administered to 755 students in 31 junior and senior high school mathematics classes. The scores were correlated yielding a coefficient of .77 . Teachers were asked to rate the attitude of those same students toward mathematics on a scale of one to nine with nine being the highest positive score. Data were also obtained relative to current grade in mathematics, overall grade point average, mental ability score, composite achievement and mathematics achievement scores, and percentiles. Teacher rating of student attitude toward mathematics and student scores on the Attitude Inventory correlated moderately (r = .48). However, Inventory scores were significantly correlated with composite achievement test percentile ranks (r = .64). Although other correlation coefficients were obtained, the reported relationships appeared minimal.

Teacher Attitudes

Bowling (1977) developed an instrument containing three subscales designed to measure attitudes of prospective teachers toward mathematics. Aiken's E and V Scales (measuring enjoyment and value of mathematics) were utilized along with a new N scale measuring prospective teachers attitudes toward the nature of mathematics. Bowling randomly organized 48 items from the three scales, Aiken's E and V scales and the N scale, and administered the resulting scale to 126 pre-service teachers. A revised 33 item scale was then administered to 328 prospective and inservice teachers. Coefficient alpha reliabilities ranged from .90 to .95 for the E scale portion, .70 for the V scale, and .85 for the N scale.

McCallon and Brown (1971) developed a semantic differential scale designed to measure attitudes of education majors toward mathematics. The researchers developed 15 items containing bi-polar adjectives placed at both ends of a continuum. The scores of 68 subjects were then correlated with the scores obtained from administration of the Aiken-Dreger Math Attitude Scale and a correlation coefficient of .90 was found.

Childress (1976) conducted studies investigating the relationship between college students attitudes toward mathematics and student ratings of teachers and courses in mathematics. A questionnaire containing 90 items was administered to 204 students enrolled in pre-calculus classes. Subscores were obtained from the 90 items measuring: (1) enjoyment of mathematics, (2) value of mathematics, (3) attitude toward mathematics, (4) teacher ratings, (5) course ratings, and (6) a combination of course and teacher ratings. Findings led Childress to conclude that general attitude toward mathematics was significantly related to course and instructor ratings.

Using the <u>Dutton Scale</u> (Dutton & Bloom, 1968), Phillips (1973) conducted studies relative to the effect of teacher attitude toward arithmetic on student attitude and achievement in mathematics. In the Phillips study, 306 seventh grade students and 59 teachers were tested. Analysis of data indicated that teacher attitude was significantly related to student attitude but not to student achievement. The study also provided evidence suggesting that the effect of teacher attitude on student attitude and achievement is cumulative. Students appeared to achieve higher in arithmetic if they had a sequence of three teachers with favorable attitudes toward mathematics.

CHAPTER III

METHODS AND PROCEDURES

This study had two major purposes. The major purposes of this study were (1) to develop an instrument to measure the attitudes of middle school mathematics teachers toward low achievers in mathematics, and (2) to establish the reliability and validity of the instrument. The methods and procedures utilized in the instrument development process are described in this chapter along with the three major phases of development: phase one - readiness, phase two - administration of the preliminary scale, and phase three - administration of the final scale.

Phase 1: Readiness

The readiness phase of this study consisted of (1) selection of instrument type and format, (2) writing of potential items, (3) classification of item directionality by judges, and (4) preparation of the preliminary attitude scale.

Selection of Instrument Type and Format

The following four types of affective scales were considered as methods for measuring teacher attitude toward low achievers in mathematics: Thurstone Scales, Likert Scales, Guttman Scales, and Semantic Differential Scales. Each of the four techniques was described in detail in Chapter II.

The strengths and weakness of each of the four scaling techniques were considered. The strengths of the Likert technique appeared to far outweigh the strengths of the other scales under consideration and the weaknesses of the Likert technique appeared to be minimal in comparison with the other type scales. Some of the strengths of the Likert scale are as follows:

1. Easy to administer and score,
2. Used more often than other scaling techniques,
3. Capable of being utilized by a wide variety of individuals in a variety of settings,
4. Capable of being administered to many individuals at the same time,
5. Adaptable to modification in response alternatives and statement format.

Further, the Likert technique appears to have the potential for exhibiting the important qualities of an attitude scale emphasized by Anderson (1981): communication value, objectivity, validity, reliability and interpretability.

Although the original Likert scale had five response categories ranging from "strongly agree" to "strongly disagree" with an option of "not sure," in the scale (TALAM) developed in this study, six response categories were chosen with the following response alternatives: strongly agree, agree, slightly agree, slightly disagree, disagree, and strongly disagree. An even number rather than an odd number of response alternatives was chosen in order to eliminate the "not sure" category. Further, an even number rather than odd number of responses was chosen in an effort to increase the reliability of the instrument.

According to Anderson (1981), the reliability of a Likert scale is likely to increase by increasing the number of response opportunities.

Writing the Potential Items

When writing potential items for inclusion in the teacher attitude toward low achievers in mathematics scale, the criteria for writing Likert type items described by Hassan and Shrigley (1984) were kept in mind. A summary of those recommendations follows:

1. Do not write in past tense,
2. Avoid factual statements or those appearing to be factual,
3. Do not use compound or complex statements,
4. Avoid the use of universals such as "always," "never," and "none,"
5. Items should be moderately negative or moderately positive rather than extreme in either sense,
6. More negatively stated items are needed than positively stated ones for they are "less prone to withstand the rigor of Likert's item analysis" (Hassan & Shrigley, 1984, pg. 660),
7. Limit the number of words in each statement to no more than 20 words,
8. The shorter the scale, the less likely respondents will experience fatigue.

Hassan and Shrigley (1984) further advised that items should be stated as egocentered, social-centered, or action-centered. Ego-centered items contain the words "I" or "me" while social-centered statements generally reflect a societal expectation. Action-centered statements are either action oriented or descriptions of intended behavior with regard to the attitudinal object. The researchers suggested that each of these categories might need to be analyzed as separate subcomponents in an attitude scale.

Fishbein and Ajzen (1975), stated that beliefs, feelings, and intended behaviors (conative behaviors) are closely inter-related components of attitude. However, the researchers suggested that attitude is better described by the "feeling" component than either beliefs or intended behaviors. Further, the researchers stated that beliefs, feelings and intended behaviors are not always consistent within an individual. For this reason, the preliminary items written for possible inclusion in the teacher attitude toward low achievers in mathematics scale contained items belonging in all three attitudinal categories: beliefs, feeling, and intended behaviors.

One method for generating scale items and for promoting content validity for an attitude scale is to elicit statements from members of the target population (Thurstone & Chave, 1928). In this study, 128 items were written based on input from middle school mathematics teachers, research literature, related scales, and from consultation with experts in mathematics and mathematics education. Approximately one-half of the items were worded favorably (positively) with regard to low achievers in mathematics while the rest were worded unfavorably (negatively) toward low achievers in mathematics. The 128 categorized items generated in "Phase 1: Readiness" are included in Appendix A. A breakdown of item categories and number of items in each category appears below:

1. Teacher BELIEFS about low achievers in mathematics:

Cultural characteristics	10 items
Cognitive ability	20 items
Work related behaviors	20 items
Affective characteristics	26 items

2. Teacher FEELINGS about working with low achievers in mathematics - 16 items
3. Teacher INTENDED BEHAVIORS with regard to working with low achievers in mathematics - 36 items

Classification of Items by Judges

All 128 items were studied by 10 judges who classified each statement as positive (+), negative (-), or neutral (?) with regard to low achievers in mathematics. The items were presented to the judges in three major categories: teacher beliefs about low achievers, teacher feelings about low achievers, and intended teacher behaviors with respect to low achievers in mathematics. The teacher belief category was further subdivided into beliefs about the cultural characteristics, cognitive characteristics, work related behaviors, and affective characteristics of low achievers in mathematics. The directions to the judges for each major category varied slightly depending on the focus of the section. An item was considered to be clear in directionality if nine of the 10 experts rated it as being clearly positive or negative with regard to low achievers in mathematics. The directions given to the judges and the 128 categorized items appear in Appendix A.

<u>Preparation of the Preliminary Attitude Scale</u>

As mentioned above, items were eliminated if not classified by at least 90% of the judges as being clearly positive or clearly negative with regard to low achievers in mathematics. Redundant items were also eliminated as were items appearing better suited to another category. Based on recommendations of the judges, seven new items, not among the original 128, were added to the teacher feeling scale and two were added to the intended teacher behavior scale based on expert opinion. This left a total of 85 items for inclusion, in random order, in the preliminary instrument.

On the preliminary scale, subjects were asked to indicate their response to each statement by selecting a "1" if they strongly disagreed with the statement, "2" if they disagreed, "3" if they slightly disagreed,"4" if they slightly agreed, "5" if they agreed and, "6" if they strongly agreed with the statement. Both the 85-item preliminary attitude scale and the directions to respondents appear in Appendix B.

Phase 2: Administration of Preliminary Scale

Phase 2 in the development of an instrument to measure the attitudes of mathematics teachers toward low achievers in mathematics involved the administration of the preliminary scale and the subsequent preparation of the final scale. This section contains a description of the methods and procedures utilized in Phase 2 and includes the following components: (1) a rationale for developing the preliminary attitude scale, (2) a description of the population and sample utilized in the preliminary scale, (3) administration procedures, (4) a summary of data analysis procedures and, (5) procedures used in constructing the final attitude scale.

Rationale

The purpose of developing the preliminary instrument was to evaluate the potential of all 85 items for possible inclusion in the final instrument. Through analysis of preliminary scale data, estimations of the internal-consistency of each item and the discriminating quality of each item could be evaluated.

Further, through analysis of preliminary scale data, a decision could be made relative to the need for one scale or several subscales. Likert (1932) suggested that the primary concern in scaling is unidimensionality, i.e. the scale should measure one construct. However, recent research has indicated that attitude is a complex multidimensional construct made up of beliefs, feelings, and intended behaviors. Consistency or inconsistency in the directionality of subjects' responses to belief, feeling, and intended behavior items, would be apparent through examination of preliminary scale data. In this light, Oppenheim (1966) cautioned that if a scale measures more than one construct and a summated score is obtained, two or more identical total scores could have entirely different meanings. For example, two subjects receive the same total score of 298 on a scale designed to measure attitude toward low achievers in mathematics. Expectation would be that the two subjects have approximately the same degree of positiveness or negativeness with regard to low achievers in mathematics. However, subject number one has generally negative beliefs about low achievers in mathematics but strong intentions to behave toward low achievers in mathematics in positive and encouraging ways in the classroom. Subject number two has generally positive beliefs about low achievers in mathematics but extremely negative feelings about working with low achievers in class. Thus, the overall scores are the same but attitudes of the two subjects are substantially different.

Population and Sample

The sample for developing the preliminary instrument was comprised of middle school mathematics teachers from the Tri-Cities region in Tennessee. All middle school mathematics teachers from the following city systems were asked to participate: Johnson City, Kingsport, and Elizabethton. Additionally, three Washington County schools were asked to participate. The sample selected consisted of 51 participants who completed the 85 item preliminary attitude scale "Statements About Low Achievers In Mathematics." The sample appeared representative of middle school teachers of mathematics in Northeast Tennessee and of middle school mathematics teachers in general. Table 1 provides a more detailed view of the number of respondents at each site.

Administration Procedures

Permission to collect data was granted by the superintendents office for each system and from the principal of each school involved and appointments made for the administration of the preliminary attitude scale at each middle school site. Prior to the appointment date, middle school mathematics teachers at each site were requested by

Table 1

Number of Teachers Contacted and Response Rate in Phase 2

System/ school	Teachers available	Teachers responding	Response per cent
Johnson City	13	13	100%
Washington County			
Location 1	6	6	100%
Location 2	4	4	100%
Location 3	5	5	100%
Kingsport City			
Location 1	8	8	100%
Location 2	7	7	100%
Elizabethton City	8	8	100%
TOTAL	51	51	100%

principals to participate in the study and informed of the appointment date for their school.

The preliminary attitude scale was administered at the conclusion of the regular school day at each middle school site. Directions and administration took approximately 20 minutes. Specific directions to respondents for completing the survey and the survey itself are presented in Appendix B.

Data Analysis Procedures

The data analyzed included the responses of all 51 participants who completed the preliminary attitude scale, "Statements About Low Achievers in Mathematics." Data are presented in Chapter IV. Data were analyzed using the Statistical Package for the Social Sciences (SPSS-PC+) (Norusis, 1988). Data analysis consisted of the following procedures:

1. Scoring of preliminary scale and subscales: The instrument was scored by summing the responses to all items on the scale after responses to all negative items had been reversed (e.g., a response of 6 became 1, a response of 5 became 2, etc.). Separate scores were tabulated for each subscale and the total scale.

2. Student's t (Ferguson, 1981) tests were conducted to determine if there were significant differences among the means of the three subscale scores and the mean of the total scale score.

3. Teacher responses on the 85 item preliminary attitude scale were factor analyzed using a principal components analysis (Hotelling, 1935; Harman, 1967) and a varimax method of rotation (Kaiser, 1958; Ferguson, 1981; Gorsuch,1974). Oppenheim (1966) suggested that a

factor analysis with a pool of only 20 items requires the computation of over 10 times that many correlation coefficients. Therefore, given the limited sample size (n = 51) as compared to the number of items (85) available for correlation in the preliminary study, the amount of information obtained through factor analysis was fairly limited.

4. Item-to-scale correlations were computed for each item and internal-consistency estimates of reliability (coefficient alpha) were calculated for each sub-scale and total scale. Inter-correlations among sub-scales and total scale were also analyzed.

5. The discriminating quality of each item was approximated. Item means for the highest 25% of total scores and for the lowest 25% of total scores were computed for each subscale and for the total scale. Differences among the means were then compared through the use of student's t statistic. This procedure provided a measure of how well a given item discriminated between high total scores and low total scores for each scale.

Construction of the Final Scale

Information obtained through the use of the statistical procedures described in items two through five above was the basis for identifying items to be included on the final attitude scale. A 45 item final scale containing three subscales of 15 items each was constructed. Consideration was also given to equal representation of items from each of the categories: beliefs, feelings, and intended behaviors. Further, a nearly equal mix of positively worded and negatively worded items was selected to comprise each of the subscales.

After the 15 items were selected for each of the three subscales, data were again analyzed taking into consideration only those items to be used in the final instrument. Item-to-scale correlations and coefficient alpha estimates of reliability were again computed on each subscale and for the total scale. Inter-correlations among sub-scales and total scale were again computed. As a further indication of validity for each of the subscales, the 45 item total scale was factor analyzed using a principal components analysis with varimax rotation.

Phase 3: Administration of Final Attitude Scale

Phase 3 involved the administration of the final scale and subsequent analysis of data. This section contains a description of the methods and procedures used in Phase 3 including the following components: (1) a description of instrumentation utilized in the study (2) a description of the population and sample, (3) administration procedures, and (4) data analysis procedures.

Instrumentation

Four instruments were used in this study: (1) The final form of the Teacher Attitudes Toward Low Achievers in Mathematics Scale (TALAM) developed for this study; (2) Revised Math Attitude Scale (RMAS) (Aiken and Dreger, 1963); (3) A semantic differential scale measuring teacher attitude toward low achievers (Steeg, 1983); (4) a single item concerning teaching preference regarding low, average or high achievers in mathematics. The final 45-item form of the TALAM appears in Appendix C along with scoring instructions for the scales. A brief description of each instrument follows:

<u>The Teacher Attitudes Toward Low Achievers in Mathematics Scale (TALAM)</u>. The TALAM is a 45 item Likert-type instrument developed after administration of an 85 item preliminary scale to a sample of 51 middle school mathematics teachers. Based on analysis of data obtained through administration of the preliminary scale, three subscales were developed for use in the final study designed to measure: (1) teacher beliefs, (2) teacher feelings, and (3) intended teacher behaviors with regard to low achievers in mathematics. Each subscale contains 15 items with some items worded positively and some worded negatively. All subscale items were randomly combined into one 45 item instrument (TALAM).

In the TALAM, subjects are asked to respond to each item by choosing one of six Likert-type alternatives: strongly agree, slightly agree, agree, disagree, slightly disagree, and strongly disagree. The response alternatives for positive items are weighted from 6 (strongly agree) to 1 (strongly disagree) with weights reversed for alternatives to negative items. A subject's total score is the sum of the weighted alternatives endorsed by the individual. High scores reflect positive attitudes toward mathematics.

Analysis of preliminary scale data resulted in elimination of all but 45 items. Data for the remaining 45 items yielded internal consistency reliability coefficients of .91 for the total scale, .88 for the beliefs subscale, .90 for the feelings subscale, and .80 for the intended behavior subscale.

<u>The Revised Math Attitude Scale (RMAS)</u>. The <u>Revised Math Attitude Scale</u> (Aiken & Dreger, 1963) is a 20-item Likert scale designed to

measure attitudes toward mathematics. Subjects respond to each item by choosing one of five alternatives: strongly agree, agree, undecided, disagree, and strongly disagree. The response alternatives for positive items are weighted from 5 (strongly agree) to 1 (strongly disagree) with weights reversed for alternatives to negative items. A subject's total score is the sum of the weighted alternatives endorsed by the individual. High scores reflect positive attitudes toward mathematics. The authors (Aiken & Dreger, 1963) reported a test-retest reliability coefficient of .94 and satisfactory indications of both content and discriminant validity.

Semantic differential scale. A 60-item semantic differential scale was developed by Steeg (1983) as a pretest and posttest to measure change in teachers' attitude toward students perceived to be low achievers. Steeg's purpose was to evaluate the effectiveness of a teacher in-service program designed to help teachers become more aware of stereotypic expectations and behaviors and to learn more appropriate and supportive behaviors toward low achievers. Analysis of pre and post test data showed a 23.9% gain for the experimental group in teacher attitudes while the control group gained 2.6% in teacher attitudes.

The scaling method used by Steeg (1983) was adapted from the more generalized technique of measurement developed by Osgood, Suci, and Tannenbaum (1957). In keeping with Osgood, Suci, and Tannenbaum's contention that three principal factors i.e. adjective types should be used in a semantic differential scale, Steeg chose adjective pairs reflecting the following types : (1) evaluative (e.g., "good-bad", "successful-unsuccessful"), (2) potency (e.g., "strong-weak") and (3) activity (e.g., "fast-slow").

In an earlier study, Osgood (1969) reported a test-retest reliability coefficient of .85 for a semantic differential scale of 1000 items. These researchers collected from a sample of 100 college students.

Single item response. Subjects were asked to read one statement and respond by placing an "X' on the line in front of the response most closely describing their feelings about working with low achievers in mathematics (either "True" or "False"). The statement: "I prefer to teach mathematics to average or high achievers in mathematics rather than to low achievers in mathematics."

Population and Sample

The sample selected consisted of 105 of the 109 possible respondents in 17 schools in Knox County, Tennessee. The sample appeared representative of middle school mathematics teachers in Tennessee and in the United States as a whole. Table 2 provides a more detailed view of the number of respondents at each site.

Administration Procedures

Permission to conduct research involving human subjects was obtained through the Research Administration of the University of Tennessee. Further, permission to conduct research in Knox County middle schools was granted by the Coordinator for Research and Evaluation of the Knox County Board of Education.

Table 2

Number of Teachers Contacted and Response Rate in Phase 3

Location number	Teachers available	Teachers responding	Response per cent
L1	8	8	100
L2	6	3	50
L3	7	7	100
L4	6	6	100
L5	4	4	100
L6	7	7	100
L7	12	12	100
L8	6	6	100
L9	7	7	100
L10	7	7	100
L11	6	6	100
L12	6	6	100
L13	6	6	100
L14	6	6	100
L15	3	3	100
L16	5	5	100
L17	5	5	100
L18	2	1	50
	109	105	96

A meeting was held with the Knox County Mathematics Supervisor, mathematics department representatives from each of the 17 middle schools, and the two middle school travelling mathematics teachers. At the meeting, mathematics department representatives consented to administer the instruments to mathematics teachers from their schools during a specified time. Further, the two travelling mathematics teachers volunteered to bring the materials to a central location after administration.

A sufficient number of materials for all mathematics teachers at each middle school site were distributed to the mathematics department representatives. Test materials included a numbered folder for each respondent containing the following numbered and color coded instruments: (1) Teacher Attitudes Toward Low Achievers in Mathematics Scale (TALAM), (2) Revised Math Attitude Scale (Aiken & Dreger, 1963), (3) semantic differential scale (Steeg, 1983), and (4) a statement concerning teaching different achievement level students in mathematics. In addition to primary test materials, a sufficient number of TALAM retest materials were also distributed. Administration procedures for the test and retest were discussed at the meeting. Copies of two handouts entitled: "Administration of Attitude Scales" and "Administrator Records" were distributed at the meeting further clarifying procedures. Copies of both handouts appear in Appendix D.

The four primary test instruments mentioned above were administered to participating middle school teachers during a time arranged by the mathematics department representative for the school.

The TALAM was readministered two weeks later. All instruments were picked up after administration by the travelling mathematics teachers and returned to a central location as planned.

Data Analysis Procedures

Data analyzed included the responses of all 105 participants who completed the following: (1) Teacher Attitudes Toward Low Achievers in Mathematics Scale (TALAM), (2) Revised Math Attitude Scale (Aiken & Dreger, 1963), (3) semantic differential scale (Steeg, 1983), (4) a single item concerning teaching preference regarding (a) average and high achievers or (b) low achievers, and (5) a retest of the TALAM. Data were analyzed using the Statistical Package for the Social Sciences (SPSS/PC+) (Norusis, 1988). Data analysis included the following procedures:

1. Scoring of instruments used was described in the Chapter III section titled "Instrumentation." Directions for scoring the final TALAM scale and subscales are presented in Appendix D.
2. Descriptive statistics were calculated after administration of the TALAM including summaries of score distributions, measures of central tendency, and variation.
3. Analysis of variance (ANOVA) was undertaken testing the significance of the differences among means of the three TALAM subscales and total scale.
4. A multiple comparison procedure using Scheffé's method (Scheffé, 1953) was computed to calculate the F ratios between all possible pairs of TALAM subscale and total scale means.

5. Pearson product moment correlations were computed between paired scores obtained through:
   (a) administration and readministration of the TALAM
   (b) administration of the TALAM and the Revised Math Attitude Scale,
   (c) administration of the TALAM and the semantic differential scale.
6. A Point biserial correlation was calculated between TALAM scores (a continuous variable) and responses to a single statement (a dichotomous variable) concerning teaching preference with regard to achievement level of mathematics students.
7. An internal-consistency estimate of reliability, Cronbach's alpha, was determined for TALAM subscales and for the total TALAM scale.

CHAPTER IV

ANALYSIS OF THE DATA

This chapter presents results of data analysis concerning the three phases of instrument development described in Chapter III. The phases are: (1) readiness, (2) administration of preliminary scale and, (3) administration of the final scale.

Phase 1: Readiness

As previously stated in Chapter III, the readiness phase of instrument development consisted of (1) selecting the instrument type and format, (2) writing the potential items, (3) classification of items by judges and, (4) preparation of the preliminary attitude scale. Selection of instrument type and format and the writing of Likert items in Phase 1 were discussed in Chapter III. Appendix A contains the 128 categorized items and the instructions given to the 10 judges. This section will describe the classification of the 128 categorized items conducted by the 10 judges and the resultant preparation of the preliminary attitude scale.

Classification of Items by Judges

Items were placed in three major categories: teacher beliefs about low achievers, teacher feelings about working with low achievers, and intended teacher behaviors with respect to low achievers in mathematics. The teacher belief category was further subdivided into beliefs about the cultural characteristics, cognitive characteristics, work related behaviors, and affective characteristics of low achievers in mathematics.

The 10 judges were asked to read the 128 categorized Likert style items and decide whether each item was stated in a positive, negative, or neutral directions with regard to low achievers in mathematics. Items were eliminated from each section unless 90% of the judges agreed to directionality (positive or negative) with regard to low achievers in mathematics. Further, items were eliminated if content of the item appeared virtually the same or highly similar to another item(s) within the scale. The results of the directionality analysis of the original 128 items conducted by the 10 experts is summarized in Table 3. As shown in Table 3, 47 items were eliminated from consideration in the final attitude scale.

Preparation of the Preliminary Attitude Scale

A total of 85 items were retained for use in the preliminary attitude scale and placed in three major subscale divisions: beliefs (n = 44), feelings (n = 20), and intended behaviors (n = 21). The categorized items were randomly renumbered one through eighty-five . Preliminary scale item numbers, subscale placements, and established directionalities are presented in Table 4.

Phase 2: Administration of Preliminary Scale

Phase 2 involved the administration of the preliminary scale and the subsequent preparation of the final scale. This section contains an analysis of the data obtained in Phase 2. Analysis of data was conducted at two levels of development. First, data obtained from

Table 3

Judgements of Ten Experts Regarding the Directionality of Items

	Item no.	% Agree	+/-		Item no.	% Agree	+/-
		Section: Beliefs: Cognitive Ability of Low Achievers in Mathematics					
**	1	90	+	**	11	100	-
	2	100	-		12	100	-
	3	100	+	**	13	100	+
	4	90	+	**	14	90	+
	5	90	+	*	15	60	-
**	6	90	+		16	100	+
	7	100	+	*	17	60	+
*	8	50	-	**	18	100	-
*	9	50	+	**	19	90	-
	10	90	+		20	90	-
		Section: Beliefs: Work Related Behaviors of Low Achievers in Mathematics					
**	1	90	-	*	11	80	-
	2	90	+	**	12	90	+
	3	90	+	*	13	70	-
	4	90	+	**	14	90	-
	5	90	-		15	80	+
*	6	70	-		16	90	+
	7	90	-	*	17	20	-
	8	90	+		18	90	-
	9	90	-	*	19	70	-
**	10	90	+	**	20	90	-
		Section: Beliefs: Affective Characteristics of Low Achievers in Mathematics					
	1	100	+	*	8	80	-
	2	100	+		9	90	+
	3	90	+		10	100	-
	4	90	-	**	11	90	-
	5	90	-		12	90	-
	6	90	-		13	100	-
*	7	60	-		14	90	-
	15	90	+		21	90	+
	16	100	-	**	22	90	-
*	17	60	-		23	90	-
	18	100	+		24	90	-
**	19	90	-	**	25	90	+
**	20	100	-	**	26	90	+

Table 3 (continued)

Item no.	% Agree	+/-	Item no.	% Agree	+/-
Section: Teacher Feelings about Working with Low Achievers in Mathematics					
* 1	80	-	9	100	-
2	100	+	10	100	-
3	100	-	11	100	-
* 4	70	-	12	100	-
5	100	+	13	100	-
6	100	+	14	100	+
7	90	+	15	90	-
8	100	+	* 16	0	?
Section: Intended Teacher Behaviors with Regard to Low Achievers in Mathematics					
* 1	100	+	19	100	+
* 2	80	-	** 20	80	+
3	90	-	o 21	item number missing	
** 4	80	-	o 22	item number missing	
** 5	80	-	o 23	item number missing	
o 6	item number missing		24	100	-
** 7	80	-	25	100	-
* 8	70	+	** 26	80	-
* 9	70	+	* 27	80	-
10	100	+	28	100	+
** 11	80	+	29	100	-
12	90	+	30	90	-
13	100	+	31	90	+
14	90	+	32	90	+
15	90	-	33	100	-
16	90	-	34	100	+
17	100	+	** 35	80	+
** 18	70	+	** 36	80	-

*Items deleted due to low directionality rating.

**Items deleted due to content redundancy.

oItem numbers were inadvertently left out.

Table 4

Preliminary Scale Item Numbers, Subscale Placement, and Directionality

Belief	Feeling	IntBeh	+/-	Belief	Feeling	IntBeh	+/-
	1		-	44			+
2			-	45			-
	3		-		46		-
4			+	47			+
5			-	48			+
		6	+	49			-
7			+	50			-
8			-		51		-
9			-	52			-
10			+			53	-
	11		-		54		+
		12	+		55		+
		13	-	56			-
	14		-	57			+
15			+	58			+
16			-			59	+
	17		-	60			-
	18		+	61			-
19			+			62	+
		20	+	63			+
21			-	64			-
22			-			65	-
23			+	66			+
		24	-	67			-
		25	+		68		+
26			-			69	+
27			+	70			+
28			-	71			+
		29	-	72			-
	30		-		73		+
	31		-		74		-
		32	-			75	+
33			+	76			+
		34	+	77			-
35			-		78		-
36			-			79	+
		37	+			80	+
		38	+			81	-
	39		+			82	-
40			+	83			+
	41		+	84			-
42			-		85		+
	43		-				

administration of the 85 item preliminary scale and subscales were analyzed. Second, when the 45 item final scale and 15 item subscale items were determined, data analysis was undertaken again focusing on each of the revised subscales and total scale. Results of data analysis are presented below:

Data Obtained from Administration of 85-Item Scale

Scoring of scales. Using the SPSS/PC+ statistical package, items were scored for the total scale and for each of the three subscales. Score distributions and approximate quartile position scores are presented in Table 5.

Comparison of scale means. Student's t-tests were conducted comparing the mean score for each of the three subscales and the total scale score. Results of t-test procedures are summarized in Table 6. Significant differences among means ($p < .001$) were found between each pair of subscales and between each subscale and the total scale.

Factor analysis of total scale. Principal factors extraction (Hotelling, 1935) with varimax rotation (Kaiser, 1958) was performed on the total 85 item scale to determine relationships among variables. According to Oppenheim (1966), factor analysis intercorrelates all the items with one another and enables the researcher to extract one or more primary factors among the pool of items. Factor-analysis can be used to eliminate items that do not belong in a scale and locate items that have high "loadings" (Oppenheim, 1966, pg. 142), generally >.30,

Table 5

Subscales and Total Scale Score Distributions and Approximate Quartile Locations

Beliefs		Feelings		Int beh		Total scale	
Score	Freq	Score	Freq	Score	Freq	Score	Freq
84	1	60	1	81	1	284	1
88	1	66	1	82	2	292	1
89	1	72	1	86	1	293	1
104	1	78	1	87	2	298	1
106	1	80	1	88	1	305	1
108	1	81	1	90	2	306	1
110	2	83	4	91	2	318	1
112	1	Q1 84	2	Q1 95	4	324	1
113	2	86	2	96	3	330	2
Q1 114	3	87	2	97	2	332	1
116	3	88	1	98	1	Q1 337	4
117	1	89	1	99	4	339	2
119	1	91	2	Q2 100	3	340	2
120	4	92	3	101	2	344	1
Q2 121	2	Q2 93	2	102	1	346	1
123	3	94	3	103	4	347	1
124	1	96	2	104	3	348	1
127	1	97	2	Q3 109	4	351	2
128	2	98	1	111	2	Q2 352	1
129	1	100	2	114	1	353	1
130	2	101	2	116	1	354	1
Q3 133	2	Q3 102	3	117	2	357	1
138	3	103	1	119	1	358	2
140	1	104	3	121	1	363	1
143	2	107	1	123	1	368	1
144	1	109	1			369	3
149	2	110	1			373	1
151	1	111	1			Q3 374	2
159	1	112	1			378	1
189	1	118	1			379	1
		120	1			385	1
						390	1
						393	1
						397	1
						405	2
						419	1
						421	1
						423	1
						487	1

Possible Scale Range: Beliefs: 44-264 Feelings: 20-120
Int Beh: 21-126 Total : 85-510

Table 6

Comparison of Scale Means Using Student's t Statistic

t-test pair	$\bar{X}$	s	$s/\sqrt{n}$	t-values
Beliefs	123.71	18.55	2.60	13.08**
Feelings	93.72	12.12	1.70	
Beliefs	123.71	18.55	2.60	-10.02**
Int. behav.	100.47	10.23	1.43	
Beliefs	123.71	18.55	2.60	-74.60**
Total dcale	356.43	37.80	5.29	
Feelings	93.72	12.12	1.70	4.27**
Int. behav.	100.47	10.23	1.43	
Feelings	93.72	12.12	1.70	-62.19**
Total scale	356.43	37.80	5.29	
Int. behav.	100.47	10.23	1.43	-48.00**
Total scale	356.43	37.80	5.29	

n = 51   Degrees of Freedom = 50

** 2-tailed level of significance: $p < .001$

on the factor being measured. "Loading" refers to the correlation between the item and the factor.

Oppenheim further suggested that data be obtained from, at least, 10 times more subjects than items. Therefore, given the limited number of subjects in the preliminary study (51) in comparison with the number of items (85), factor analysis did not appear to be a powerful statistical tool in the context of this study. Never the less, data indicated that the 85 item scale did break up into several independent categories. Factor analysis yielded a primary first factor, made up of teacher belief items, accounting for 18.5% of total instrument variance. Every item originally written as a teacher belief item loaded on factor 1 at .30 or higher. Further, 15 of the 44 total belief items loaded on factor 1 at .50 or greater. Data analysis also yielded a secondary factor, made up of items directed toward teacher feelings, accounting for 11% of total instrument variance. On factor 2, 12 teacher feeling items loaded with values greater than .50 while all other teacher feeling items loaded on factor 2 at values ranging from .30 to .40. Factors 3 through 8 were a mixture of teacher beliefs, feelings, and intended behaviors. Table 7 gives a view of the of the eight factors appearing in the rotated factor matrix and the loadings of each attitude statement on the factors.

<u>Internal consistency estimates of reliability</u>. Item-total correlations were computed for each item within each subscale and the total 85-item scale. Table 8 presents item-total correlations for items in the total scale and corrected item-total correlations, means, variances, and estimates of the scale's internal-consistency

Table 7

Rotated Factor Matrix Values for the Preliminary Total Scale

Attitude item no.	Factor numbers 1	2	3	4	5	6	7	8
1		.36	.73					
2	.32			.74				
3		.34		.48				
4	.62							
5	.30					.31		
6	.36							
7	.32			.38				
8	.39							
9	.69							
10	.31			.45				
11		.84						
12				.49	.35			
13				.51	.30			
14		.56						
15	.38			.46				
16	.31							
17		.70						
18		.79						
19	.35							.58
20					.50			
21	.44							
22	.32					.52		
23	.33							
24				.31	.30			
25					.47			
26	.56							
27					.36			
28	.52							
29					.42			
30		.38	.60					
31		.74						
32			.69		.31			
33	.33						.39	
34				.54	.50			
35	.48							
36	.32		.60					
37					.31		.41	
38					.33			
39		.77						
40	.40			.49				
41		.40						

Table 7 (continued)

Attitude item no.	Factor numbers 1	2	3	4	5	6	7	8
42	.50	.47						
43		.58						
44	.35							.55
45	.31					.46		
46		.55						
47	.36			.65				
48	.36							
49	.64							
50	.76							
51		.30						
52	.52							
53				.68	.31			
54		.31					.66	
55		.50						
56	.31			.35				
57	.35						.39	
58	.68							
59					.33		.38	
60	.32					.45		
61	.54							
62					.45			
63	.34						.61	
64	.30					.77		
65			.55		.38			
66	.39						.45	
67	.64							
68		.55						
69					.31			
70	.67							
71	.50							
72	.35					.76		
73		.73						
74		.37	.64					
75					.45			
76	.67							
77	.53							
78		.59						
79					.62			
80					.59			
81			.75		.29			
82			.78		.31			
83	.39							
84	.36							.44
85		.23	.25					

Table 8

Item-Scale Correlations for Preliminary Total Scale Items

	Item total statistics				
Item no.	Mean if item deleted	Variance if item deleted	Item/ total corr.	Corrected item/total corr.	Alpha if item deleted
Q1	351.33	1384.39	.51**	.48	.93
Q2	351.33	1396.63	.36*	.33	.93
Q3	351.71	1401.77	.31	.28	.93
Q4	352.76	1389.74	.45**	.42	.93
Q5	353.37	1421.56	.09	.06	.93
Q6	352.92	1416.67	.12	.07	.94
Q7	352.27	1401.00	.28	.25	.93
Q8	352.08	1392.19	.42**	.40	.93
Q9	352.29	1391.29	.46**	.44	.93
Q10	352.59	1402.81	.31	.28	.93
Q11	351.51	1391.97	.43**	.41	.93
Q12	351.25	1398.87	.31	.27	.93
Q13	351.98	1404.46	.29	.26	.93
Q14	352.06	1371.06	.56**	.53	.93
Q15	353.16	1389.49	.45**	.43	.93
Q16	352.29	1392.61	.41*	.39	.93
Q17	351.73	1401.80	.33*	.31	.93
Q18	351.67	1405.71	.40*	.38	.93
Q19	353.08	1408.67	.22	.19	.93
Q20	351.24	1421.54	.11	.08	.93
Q21	353.35	1388.59	.40*	.37	.93
Q22	351.41	1396.57	.46**	.44	.93
Q23	351.63	1387.04	.57**	.55	.93
Q24	352.53	1368.57	.50**	.47	.93
Q25	351.33	1411.67	.26	.23	.93
Q26	353.02	1391.98	.40*	.37	.93
Q27	351.67	1398.27	.46**	.44	.93
Q28	352.02	1395.58	.39*	.36	.93
Q29	352.06	1389.38	.42*	.39	.93
Q30	351.14	1393.40	.61**	.60	.93
Q31	351.51	1396.17	.41*	.38	.93
Q32	351.47	1396.29	.40*	.37	.93
Q33	352.65	1390.59	.50**	.48	.93
Q34	351.10	1410.73	.40*	.38	.93
Q35	353.43	1393.01	.44**	.42	.93
Q36	352.78	1377.89	.52**	.50	.93
Q37	351.08	1413.79	.27	.25	.93
Q38	350.86	1418.36	.20	.18	.93
Q39	351.84	1388.81	.49**	.46	.93
Q40	352.16	1382.29	.55**	.53	.93
Q41	352.39	1414.64	.17	.14	.93

Table 8 (Continued)

	Item total statistics				
Item no.	Mean if item deleted	Variance if item deleted	Item/ total corr.	Corrected item/total corr.	Alpha if item deleted
Q42	352.10	1365.53	.65**	.63	.93
Q43	351.45	1404.93	.32	.29	.93
Q44	353.31	1399.34	.34*	.31	.93
Q45	352.80	1388.28	.44**	.42	.93
Q46	352.76	1365.74	.65**	.63	.93
Q47	352.67	1377.67	.60**	.58	.93
Q48	353.71	1417.25	.14	.11	.93
Q49	352.53	1400.61	.33*	.30	.93
Q50	353.22	1407.25	.27	.24	.93
Q51	351.08	1413.19	.29	.27	.93
Q52	353.22	1389.77	.44**	.41	.93
Q53	353.28	1389.52	.38*	.34	.93
Q54	351.71	1418.93	.17	.15	.93
Q55	351.53	1414.25	.31	.30	.93
Q56	352.41	1391.28	.44**	.41	.93
Q57	352.27	1393.24	.44**	.41	.93
Q58	353.08	1383.55	.56**	.54	.93
Q59	351.20	1406.60	.39*	.37	.93
Q60	353.06	1417.02	.17	.14	.93
Q61	352.67	1377.55	.57**	.54	.93
Q62	351.49	1416.01	.29	.28	.93
Q63	352.39	1395.92	.44**	.42	.93
Q64	352.53	1386.25	.51**	.49	.93
Q65	351.49	1392.69	.56**	.54	.93
Q66	352.73	1388.60	.43**	.41	.93
Q67	353.84	1406.33	.32	.30	.93
Q68	352.25	1392.27	.45**	.42	.93
Q69	352.71	1398.01	.33*	.30	.93
Q70	353.71	1380.57	.60**	.58	.93
Q71	353.61	1399.28	.40*	.38	.93
Q72	352.96	1385.83	.53**	.51	.93
Q73	351.98	1378.18	.64**	.62	.93
Q74	351.55	1394.29	.50**	.48	.93
Q75	351.67	1432.35	-.03	-.06	.93
Q76	353.16	1399.33	.39*	.36	.93
Q77	353.29	1394.33	.42*	.40	.93
Q78	352.43	1388.77	.43**	.41	.93
Q79	351.06	1418.50	.15	.13	.93
Q80	351.07	1416.43	.21	.19	.93
Q81	351.71	1382.29	.47**	.45	.93
Q82	351.10	1393.73	.48**	.45	.93
Q83	352.98	1380.34	.63**	.61	.93
Q84	353.16	1376.33	.59**	.57	.93
Q85	351.27	1419.92	.15	.13	.93

1-tailed significance level: * $p < .01$ ** $p < .001$

reliability (coefficient alpha) if the item were deleted. Tables 9, 10, 11, and 12 present data concerning item-total correlations, respectively, for each of the following: (1) teacher beliefs about low achievers, 34 items, (2) teacher feelings about low achievers, 20 items, (3) intended teacher behavior with respect to low achievers, 21 items, and (4) teacher belief items related to the cultural characteristics of low achievers, 10 items. Table 13 provides a summary of item numbers in each scale and an estimate of the internal consistency reliability (Cronbach's alpha) for each scale. The reliability coefficients, ranging from .58 to .93, were all significant ($p < .001$), suggesting a high level of inter-item homogeneity. Table 14 presents an inter-item correlation matrix for the subscales and total scale. Analysis of data indicated that the cultural belief items (n=10) were inconsistent with the feeling and intended behavior subscales and were therefore eliminated from inclusion in the final scale.

Testing discriminating quality of items. Positively written statements can be considered valid if subjects with a generally positive attitude slightly agree, agree, or strongly agree with the item. If valid, subjects with a generally negative attitude would slightly disagree, disagree, or strongly disagree with positively written statements. Positive criterion groups and negative criterion groups were established for the total scale and for each of the following subscales: beliefs, feelings, and intended behaviors. Criterion groups were established by selecting subjects who scored in the top 25% and those who scored in the lowest 25% for each subscale

Table 9

Item-Scale Correlations for Teacher Belief Preliminary Subscale Items

	Item total statistics				
Item no.	Mean if item deleted	Variance if item deleted	Item/ total corr.	Corrected item/total corr.	Alpha if item deleted
Q4	120.04	319.84	.57**	.53	.89
Q7	119.55	327.33	.36*	.29	.89
Q10	119.86	332.40	.30	.24	.89
Q15	120.43	326.89	.42*	.36	.89
Q16	119.57	324.85	.46**	.41	.89
Q19	120.35	335.39	.22	.15	.90
Q22	118.69	328.82	.45**	.41	.89
Q27	118.94	333.21	.35*	.30	.89
Q28	119.29	326.65	.43**	.37	.89
Q33	119.92	323.11	.57**	.53	.89
Q35	120.71	325.41	.48**	.44	.89
Q36	120.06	323.53	.45**	.39	.89
Q40	119.43	329.85	.37*	.31	.89
Q42	119.37	320.51	.52**	.46	.89
Q44	120.58	328.37	.38*	.32	.89
Q45	120.08	323.51	.47**	.43	.89
Q47	119.94	317.97	.64**	.60	.89
Q48	120.98	332.37	.29	.22	.89
Q49	119.80	324.96	.47**	.41	.89
.Q50	120.49	327.97	.41*	.36	.89
Q56	119.67	326.14	.44**	.39	.89
Q57	119.55	328.57	.40*	.35	.89
Q58	120.35	317.67	.67**	.64	.89
Q60	120.33	333.90	.30	.25	.89
Q61	119.24	316.37	.64**	.59	.89
Q63	119.67	327.18	.48**	.43	.89
Q66	120.00	320.28	.54**	.48	.89
Q67	121.11	328.55	.45**	.41	.89
Q70	120.98	320.10	.62**	.58	.89
Q71	120.88	326.07	.51**	.47	.89
Q76	120.43	324.33	.53**	.49	.89
.Q77	120.56	328.85	.59**	.54	.89
.Q83	120.25	319.47	.66**	.63	.89
.Q84	120.43	315.65	.66**	.62	.89

n = 51

1-tailed significance level: * p < .01 ** p < .001

Table 10

Item-Scale Correlations for Teacher Feelings Preliminary Subscale Items

	Item total statistics				
Item no.	Mean if item deleted	Variance if item deleted	Item/ total corr.	Corrected item/total corr.	Alpha if item deleted
Q3	89.00	136.64	.40*	.31	.89
Q11	88.80	128.44	.69**	.64	.87
Q14	89.35	123.15	.74**	.68	.87
Q17	89.02	131.94	.59**	.52	.88
Q18	88.96	134.24	.69**	.66	.88
Q30	88.43	135.73	.62**	.58	.88
Q31	88.80	129.36	.70**	.64	.87
Q39	89.14	126.32	.79**	.76	.87
Q41	89.69	136.18	.40*	.31	.89
Q43	88.74	132.71	.60**	.54	.88
Q46	90.06	127.14	.67**	.60	.88
Q51	88.37	138.92	.47**	.43	.88
Q54	89.00	141.80	.28	.22	.87
Q55	88.82	138.03	.53**	.49	.88
Q68	89.55	133.09	.55**	.48	.88
Q73	89.27	127.24	.79**	.78	.87
Q74	88.84	138.93	.38*	.31	.88
Q78	89.73	129.32	.62**	.55	.88
Q85	88.57	143.57	.18	.11	.89

n = 51

1-tailed significance level: * $p < .01$ ** $p < .001$

Table 11

Item-scale Correlations for Intended Teacher Behavior Preliminary Subscale Items

	Item total statistics				
Item no.	Mean if item deleted	Variance if item deleted	Item/ total corr.	Corrected item/total corr.	Alpha if item deleted
Q6	96.96	98.83	.25	.09	.80
Q12	95.29	93.69	.45**	.34	.78
Q13	96.01	93.02	.53**	.44	.77
Q20	95.27	98.96	.32	.21	.78
Q24	96.57	87.61	.58**	.46	.77
Q25	95.37	98.24	.37*	.29	.78
Q29	96.10	95.45	.40*	.29	.78
Q32	95.51	93.57	.53**	.44	.77
Q34	95.14	96.44	.67**	.63	.77
Q37	95.12	97.03	.52**	.46	.77
Q38	94.90	98.13	.45**	.39	.77
Q53	97.31	95.18	.39*	.25	.78
Q59	95.23	96.42	.54**	.49	.77
Q62	95.53	99.17	.46**	.41	.77
Q65	95.53	93.09	.68**	.63	.76
Q69	96.75	91.67	.54**	.44	.77
Q75	95.71	101.65	.20	.11	.79
Q79	95.10	97.65	.39*	.30	.78
Q80	95.12	96.47	.52**	.46	.77
Q81	95.75	94.91	.41*	.30	.78
Q82	95.14	96.32	.45**	.36	.77

1-tailed significance level: * p < .01 ** p < .001

Table 12

Item-Scale Correlations for Cultural Belief Subscale (10 Items)

	Item total statistics				
Item no.	Mean if item deleted	Variance if item deleted	Item/ total corr.	Corrected item/total corr.	Alpha if item deleted
Q2	33.43	57.05	.35*	.20	.87
Q5	35.47	53.09	.54**	.41	.85
Q8	34.18	51.99	.63**	.53	.84
Q9	34.39	52.32	.65**	.55	.84
Q21	35.45	48.41	.72**	.62	.83
Q23	33.73	54.68	.56**	.46	.85
Q26	35.12	49.39	.73**	.64	.83
Q52	35.31	48.26	.81**	.74	.82
Q64	34.63	49.88	.77**	.71	.83
Q72	35.06	49.46	.81**	.76	.82

1-tailed significance level: * $p < .01$ ** $p < .001$

Table 13

Internal-Consistency Estimates of Reliability (Coefficient Alpha) for the Preliminary Scale and Subscales

Scale	No. of items	Coefficient Alpha
CULTURAL CHARACTERISTICS	10	.58**
BELIEFS	34	.89**
FEELINGS (TEACHER)	20	.88**
INTENDED BEHAVIORS	21	.78**
TOTAL	85	.93**

n = 51 ** 1-tailed significance level: $p < .01$

Table 14

Inter-Correlation Matrix for the Preliminary Scale and Subscales

	Cult.	Bel.	Feel.	Int.Beh.	Total
Cultural Bel.	1.00**	.65**	.13	.21	.63**
Beliefs	.65**	1.00**	.50**	.46**	.91**
Feelings	.13	.50**	1.00**	.50**	.73**
Intended Behav.	.21	.46**	.50**	1.00**	.70**
Total Scale	.63**	.91**	.73**	.70**	1.00**

n=51

1-tailed significance level: * $p < .01$ ** $p < .001$

and total scale. The mean score of the high and low criterion groups for each item were analyzed using student's t-tests to determine differences between the mean of the high and low groups. Two-tailed tests were conducted at the .05 level of significance. Items with a t-score equal to or greater than (+/-) 2.07 were considered to have sufficient discriminating quality to be considered for possible inclusion on the final attitude scale. However, Edwards (1957) recommended that a t-score as low as (+/-) 1.75 could be considered sufficiently discriminating for an item with a sample size of approximately 25. Figure 4 illustrates the process followed for testing the discriminating quality of items. Results of student's t-tests for the mean of each subscale item appear in Table 15.

Data Obtained from Revision of Total Scale and Subscale

After examining data obtained from the statistical analysis of responses to the original 85 item scale and subscales, a revised 45 item scale was designed with three 15-item subscales measuring the following components of attitude: beliefs, feelings, and intended behaviors. Items were considered for inclusion in the subscales if they (1) discriminated between high and low criterion groups, and (2) had a significant (.01 or .001) item/total correlation with other items in the subscale. This section presents results obtained from statistical analysis of data based upon subjects' responses to the 45 items in the final scale. Analysis of data obtained in response to the 45 items included the following: (A) item/total correlations for items and internal-consistency reliability analysis for each subscale and total scale and, (B) factor analysis of the entire 45 item scale.

Calculated t = - 4.21*

Scale: Teacher Belief			Item No. 4(+)					
			Frequency of Responses					
Group	X	s	1	2	3	4	5	6
Low	3.08	1.24	1	4	5	1	1	1
High	4.58	1.00	0	0	2	3	5	2

Degrees of Freedom: 22

* 2-tailed significance level: $p < .05$

Critical value of t: (+/-) 2.07

Figure 4. Sample of Student's t analysis for one item.

Table 15

Determination of the Discriminating Quality of Preliminary Subscale Items Using Student's t Test

SUBSCALE					
Beliefs		Feelings		Int. behaviors	
Item #	t-score	Item #	t-score	Item #	t-score
4	-4.21	1	6.13	6	3.85
7	-2.03***	3	3.51	12	3.95
10	- .91***	11	5.67	13	-3.85
15	-2.89	14	9.07	20	-1.49
16	-4.04	17	4.62	24	6.19
19	-1.78***	18	3.52	25	3.06
22	-2.45	30	4.70	29	3.39
27	-2.19	31	4.18	32	5.97
28	-2.96	39	4.94	34	5.65
33	-3.13	41	2.78	37	3.34
35	-2.31	43	5.19	38	3.85
36	-3.99	46	5.35	53	2.39
40	-2.78	51	4.75	59	4.78
42	-2.86	54	1.91***	62	2.83
44	-2.86	55	3.34	65	5.78
45	-2.09	68	3.76	69	3.99
47	-4.66	73	5.76	75	-1.48***
48	-1.32***	74	2.40	79	1.93
49	-3.05	78	3.24	80	2.94
50	-3.30	85	.23***	81	2.89
56	-2.53			82	4.32
57	-2.18				
58	-3.75				
60	-2.23				
61	-6.90				
63	-3.65				
66	-3.99				
67	-2.50				
70	-3.97				
71	-3.90				
76	-4.04				
77	-3.79				
83	-3.87				
84	-5.29				

2-tailed significance level: $p < .05$

***Items not considered for final scale due to lack of discrimination between high and low criterion groups.

Critical value of t: (+/-) 2.07

<u>Internal-consistency reliability</u>. Item/total correlations were computed for each individual item for the total 45 item scale and for each 15 item subscale. Within the teacher belief and teacher feeling subscales, every item chosen for inclusion had a significant correlation with the rest of the subscale at the .001 level of significance. In the intended behavior subscale, every item chosen for the subscale had a significant item/total correlation at the .01 or .001 level. When pooled together into a total scale, 76% of all items had a significant item/total correlation at the .01 or .001 level. All items in all scales, including the total scale, had significant item/total correlations at the .05 level of significance. Tables 16, 17, 18, and 19 summarize item/total correlations and present reliability coefficients for each subscale and for the total scale. Table 20 presents internal-consistency reliability estimates Cronbach's coefficient alpha, for each scale. Table 21 provides an inter-correlation matrix for each of the 15-item subscales as well as for the total scale.

<u>Factor analysis of total 45 item revised scale</u>. After discarding seven items because of insufficient discriminating quality between high and low criterion groups and after discarding 33 items due to nonsignificant item/total correlation coefficients, as mentioned above, a 45-item final attitude scale was determined. To further evaluate the validity of the 3 subscales, a principal factors extraction with varimax rotation was performed for the 45 item total scale.

The three factors: (1) teacher feelings, (2) teacher beliefs and, (3) teacher intended behaviors, accounted for 62% of the total

Table 16

Item-Scale Correlations for the Revised Total Scale (45 Items)

	Item total statistics				
Item no.	Mean if item deleted	Variance if item deleted	Item/ total corr.	Corrected item/total corr.	Alpha if item deleted
Q1	186.98	471.47	.55**	.51	.91
Q4	188.41	480.93	.38*	.34	.91
Q11	187.16	476.77	.49**	.45	.91
Q12	186.90	485.21	.28	.24	.91
Q13	187.63	487.16	.27	.23	.91
Q14	187.71	461.81	.64**	.60	.91
Q17	187.37	481.60	.41*	.37	.91
Q18	187.31	483.98	.50**	.46	.91
Q24	188.18	467.47	.48**	.44	.91
Q30	186.78	479.81	.62**	.58	.91
Q31	187.16	476.13	.53**	.49	.91
Q32	187.12	476.49	.49**	.45	.91
Q33	188.29	477.33	.52**	.49	.91
Q34	186.75	490.59	.38*	.34	.91
Q35	189.08	478.83	.45*	.42	.91
Q37	186.73	488.92	.36	.32	.91
Q38	186.51	492.65	.26	.23	.91
Q39	187.49	469.77	.63**	.60	.91
Q43	187.10	480.69	.45**	.41	.91
Q46	188.41	460.17	.71**	.68	.91
Q47	188.31	474.66	.53**	.49	.91
Q49	188.18	485.95	.31	.27	.91
Q50	188.86	487.12	.30	.26	.91
Q51	186.73	488.20	.40*	.36	.91
Q53	188.92	475.71	.36	.33	.91
Q55	187.18	490.95	.37*	.33	.91
Q58	188.73	477.00	.50**	.46	.91
Q59	186.84	486.61	.42*	.39	.91
Q61	188.31	473.54	.54**	.50	.91
Q62	187.14	492.32	.33	.30	.91
Q65	187.13	479.52	.57**	.53	.91
Q67	189.49	491.45	.28	.24	.91
Q68	187.90	477.77	.48**	.44	.91
Q69	188.35	478.79	.41*	.37	.91
Q70	189.35	470.63	.63**	.61	.91
Q71	189.25	481.87	.45**	.41	.91
Q73	187.63	466.00	.74**	.72	.91

Table 16 (continued)

Item no.	Mean if item deleted	Variance if item deleted	Item/ total corr.	Corrected item/total corr.	Alpha if item deleted
		Item total statistics			
Q76	188.81	486.61	.74**	.70	.91
Q77	188.94	485.02	.31	.28	.91
Q78	188.08	470.71	.56**	.52	.91
Q80	186.73	493.76	.22	.19	.91
Q81	187.35	473.79	.46**	.42	.91
Q82	186.75	477.47	.54**	.50	.91
Q83	188.63	475.72	.56**	.52	.91
Q84	188.80	475.96	.49**	.45	.91

n = 51

1-tailed significance level: * p < .01 ** p < .001

Table 17

Item-Scale Correlations for the Revised Teacher Beliefs Subscale (15 Items)

	Item total statistics				
Item no.	Mean if item deleted	Variance if item deleted	Item/ total corr.	Corrected item/total corr.	Alpha if item deleted
Q4	46.06	92.18	.60**	.52	.87
Q33	45.94	95.02	.55**	.47	.88
Q35	46.73	94.08	.56**	.48	.88
Q47	45.96	94.32	.53**	.45	.88
Q49	45.82	92.79	.59**	.50	.88
Q50	46.51	92.89	.60**	.52	.87
Q58	46.37	90.20	.73**	.68	.87
Q61	45.96	90.48	.66**	.58	.87
Q67	47.14	94.72	.59**	.52	.87
Q70	47.00	91.12	.70**	.64	.87
Q71	46.90	94.73	.57**	.50	.87
Q76	46.45	93.37	.62**	.55	.87
Q77	46.59	92.25	.63**	.56	.87
Q83	46.27	93.64	.61**	.54	.87
Q84	46.45	90.89	.65**	.57	.87

$n = 51$

1-tailed significance level: ** $p < .001$

Table 18

Item-Scale Correlations for the Revised Teacher Feelings Subscale (15 Items)

	Item total statistics				
Item no.	Mean if item deleted	Variance if item deleted	Item/ total corr.	Corrected item/total corr.	Alpha if item deleted
Q1	65.10	99.93	.48**	.38	.90
Q11	65.27	94.20	.72**	.66	.89
Q14	65.82	89.27	.78**	.72	.89
Q17	65.49	97.09	.62**	.55	.89
Q18	65.43	99.29	.72**	.68	.89
Q30	64.90	101.21	.61**	.56	.89
Q31	65.27	95.20	.71**	.66	.89
Q39	65.60	92.36	.82**	.78	.88
Q43	65.21	98.01	.62**	.56	.89
Q46	66.53	93.25	.69**	.61	.89
Q51	64.84	104.05	.46**	.40	.90
Q55	65.29	104.57	.48**	.43	.90
Q68	66.02	99.10	.54**	.46	.90
Q73	65.75	93.07	.82**	.78	.88
Q78	66.20	95.52	.62**	.54	.89

n = 51

1-tailed significance level: ** $p < .001$

Table 19

Item-Scale Correlations for the Revised Intended Teacher Behavior Subscale (15 Items)

	Item total statistics				
Item no.	Mean if item deleted	Variance if item deleted	Item/ total corr.	Corrected item/total corr.	Alpha if item deleted
Q12	66.98	63.46	.44*	.39	.80
Q13	67.71	62.45	.53**	.43	.79
Q24	68.25	59.75	.57**	.41	.80
Q32	67.20	60.84	.62**	.56	.78
Q34	66.82	65.43	.67**	.61	.78
Q37	66.80	64.92	.59**	.52	.79
Q38	66.59	65.69	.55**	.46	.79
Q53	69.00	61.08	.37*	.32	.79
Q59	66.92	65.35	.57**	.47	.79
Q62	67.22	67.49	.47**	.40	.79
Q65	67.22	62.57	.71**	.62	.78
Q69	68.43	63.25	.49**	.39	.80
Q80	66.80	65.68	.54**	.42	.79
Q81	67.43	62.49	.44*	.41	.80
Q82	66.82	63.82	.54**	.44	.79

n = 51

1-tailed significance level: * $p < .01$ ** $p < .001$

Table 20

Internal Consistency Estimates of Reliability (Coefficient Alpha) for the Revised Total Scale and Subscales

Scale	No. of items	Coefficient alpha
Beliefs	15	.88
Feelings (teacher)	15	.90
Intended behaviors	15	.80
TOTAL	45	.91

n = 51

Table 21

Inter-Correlation Matrix for the Revised Total Scale and Subscales

	Bel.	Feel.	Int. beh.	Total
Beliefs	1.00**	.34*	.24	.71**
Feelings	.34*	1.00**	.54**	.84**
Intended behaviors	.24**	.54**	1.00**	.73**
Total scale	.71**	.84**	.73**	1.00**

n = 51

1-tailed significance level: * $p < .01$ ** $p < .001$

instrument variance. Factor 1, teacher feelings, with an eigenvalue of 10.12, loaded at .49 or greater for 12 of the 15 items. Factor 2, teacher beliefs, with an eigenvalue of 5.31 loaded at .48 or greater for 14 of the 15 items. Factor 3, intended teacher behaviors, with an eigenvalue of 3.5, loaded at .40 or greater for 14 of the 15 items. Data results obtained from factor analysis suggest the validity of constructs measured by each of the 3 subscales as well as the validity of individual items placed in the subscales. Table 22 presents a summary of the factor analysis of the 45 items comprising the revised total scale.

Phase 3: Administration of Final Scale

Phase 3 consisted of administering the following instruments to a sample of 105 middle school mathematics teachers in Knox County, Tennessee: (1) Teacher Attitudes Toward Low Achievers in Mathematics Scale (TALAM), developed for this study, (2) Revised Math Attitude Scale (RMAS) (Aiken & Dreger, 1963), (3) a semantic differential scale measuring teacher attitude toward low achievers (Steeg, 1983), (4) a single statement: "I prefer to teach mathematics to average or high achievers in mathematics rather than to low achievers in mathematics." requiring respondents to answer "True" or "False," and (5) readministration of the TALAM after a period of two weeks. This section contains an analysis of the data obtained from administration of the instruments listed above. Data were analyzed using the Statistical Package for the Social Sciences (SPSS/PC+) (Norusis, 1988).

Table 22

Factor Analysis Results for the Revised Attitude Scale (45 Items)

Factor 1:Feelings		Factor 2:Beliefs		Factor 3: Int. Beh.	
Item #	Loading	Item #	Loading	Item #	Loading
31	.77	58	.70	34	.74
18	.74	84	.68	65	.63
73	.73	77	.68	13	.63
11	.72	70	.67	38	.63
17	.70	76	.65	37	.59
14	.69	67	.61	80	.54
43	.59	61	.59	32	.54
46	.56	50	.57	12	.49
55	.56	35	.56	53	.47
78	.56	49	.56	1	.46
68	.54	71	.54	62	.44
30	.49	83	.54	59	.44
		47	.48	24	.41
		33	.48	82	.41

Note: Loading values are the correlation coefficients between individual items and the primary factors.

Scoring

Total scores for the Teacher Attitude Toward Low Achievers in Mathematics Scale (TALAM), containing 45 items, had possible minimum and maximum scores, respectively, of 45 to 270. Each of the three 15-item subscales (Beliefs, Feelings, Intended Behaviors) contained within the total TALAM had a possible minimum and maximum scores of 15 to 90. In all scales, the higher scores indicated a more positive attitude toward low achievers in mathematics. Appendix D contains information relative to scoring the TALAM without the use of a computer. Table 23 summarizes the scores and frequency of occurrence of scores obtained from administration of the TALAM and establishes approximate quartile locations for each subscale and total scale.

Descriptive Statistics

Summary descriptive statistics are presented in Table 24 for TALAM subscale and total scale data. Descriptive statistics included are as follows: mean, standard deviation, variance, standard error of the mean, minimum and maximum scores, and the range.

Test of Significance

A one-way analysis of variance was used to test the significance of the differences among the means of TALAM subscales and total scale. As shown in Table 25, a significant difference among means was determined ($p < .01$) indicating greater variation among means than would be expected by chance.

Table 23

Subscale and Total Scale Score Distributions With Approximate Quartile Locations

Beliefs		Feelings		Int. beh.		Total	
Score	Freq	Score	Freq	Score	Freq	Score	Freq
22	1	39	1	42	1	125	1
26	1	40	1	55	1	133	1
28	1	44	1	56	1	134	1
31	1	46	1	57	1	142	1
33	1	47	1	58	2	151	2
34	2	48	1	59	1	161	2
35	1	52	1	60	1	162	2
37	1	56	1	61	1	165	3
40	3	57	3	62	1	166	1
41	3	58	2	63	3	167	2
42	5	59	2	64	4	168	1
43	1	60	3	65	3	171	1
44	3	61	4	Q1 66	4	172	1
Q1 45	3	62	3	67	7	173	1
46	5	Q1 63	3	68	3	174	3
47	1	64	4	69	7	175	2
48	2	65	2	70	10	Q1 176	1
49	5	66	2	Q2 71	5	177	1
50	6	67	6	72	7	178	1
Q2 51	7	68	2	73	5	180	3
52	5	69	3	74	4	181	4
53	4	70	4	Q3 75	5	182	1
54	3	Q2 71	4	76	5	183	2
55	7	72	5	77	6	186	5
56	3	73	9	79	1	187	1
Q3 57	5	74	3	80	4	188	1
58	2	75	5	83	1	190	3
59	1	Q3 76	4	84	1	191	1
60	4	77	6	85	4	Q2 192	4
61	3	78	3	87	2	193	3
62	1	79	3	88	1	194	2
63	2	81	2			195	1
64	2	82	1			196	3
65	1	83	1			197	1
67	4	84	2			198	3
68	1	86	1			200	2
71	1	87	1			201	1
72	2	90	4			202	3
73	1					203	1
						204	2
						205	2

Table 23 (Continued)

Beliefs		Feelings		Int. beh.		Total	
Score	Freq	Score	Freq	Score	Freq	Score	Freq
						206	2
						Q3 207	1
						209	1
						210	1
						211	2
						212	2
						214	2
						215	1
						219	2
						220	1
						222	2
						223	2
						224	1
						225	1
						226	1
						230	1
						231	1
						232	1
						238	1
						243	1
						248	1

Table 24

TALAM Data: Descriptive Statistics

Scale	# items	$\bar{X}$	s	s^2	$s/\sqrt{n}$	Min	Max	Range
Beliefs	15	51.36	10.06	101.25	.98	22	73	51
Feelings	15	69.35	10.35	107.08	1.01	39	90	51
Int. behav.	15	71.20	7.68	59.05	.75	42	88	46
Total	45	191.91	23.58	555.81	2.30	125	248	123

Table 25

Analysis of Variance for Means of Subscale and Total Scale Scores

Source of variation	Sum of squares	Degrees of freedom	Variance estimate	F Value
Between	1,304,309.99	3	434,769.99	
Within	95,611.24	416	229.83	
Total	1,399,921.23	419		1,891.70**

Critical Value of $F_{.01}$, 3, 416 = 3.83

**Two-tailed significance level: $p < .01$

Multiple-Comparison Procedure

As indicated above, the one way analysis of variance (ANOVA) lead to a significant F test, that is, significant difference(s) among means of TALAM subscales and total scale were determined. Scheffé's multiple comparison method was used to compare all possible pairs of means to find the location of significant difference(s). As shown in Table 26, significant differences among several pairs of means were identified at the .05 level of significance.

Internal Consistency Reliability

Cronbach's coefficient alpha was used as a measure of internal consistency, or homogeneity of scale items. Coefficient alpha, derived for each TALAM subscale and total scale, are presented in Table 27.

Correlations

The TALAM was administered twice to a sample of 105 subjects. Scores were compared using the Pearson product-moment correlation. Table 28 describes the approximate degree of relation between TALAM subscale/total scale scores and TALAM subscale/total retest scores.

Pearson product-moment correlations were also used to compare TALAM subscale/total scale scores with scores obtained from administrations of the Revised Mathematics Attitude Scale (RMAS) (Aiken & Dreger, 1963) and the semantic differential scale (Steeg, 1983).

Further, a point-biserial correlation was used to compare TALAM subscale/total scale scores with dichotomous responses ("True" or "False") of subjects to the following statement: "I prefer to teach mathematics to averages or high achievers in mathematics rather than to

Table 26

Scheffé's Multiple Comparison of Means of Subscale and Total Scale Scores

Comparison of paired means	Calculated F values
I, II	00.18
I, III	00.22
I, IV	10.85 *
II, III	00.01
II, IV	08.25 *
III, IV	08.00 *

I, II, III, IV $p < .05$

Where I = Mean of Teacher Belief Scores
II = Mean of Teacher Feeling Scores
III = Mean of Intended Behavior Scores
IV = Mean of Total Scores

Table 27

Internal-Consistency Estimates of Reliability (Cronbach's Alpha) for the Final Scale and Subscales

Scale	# Items	Coefficient alpha
Beliefs	15	.84**
Feelings	15	.86**
Int. behaviors	15	.70**
Total scale	45	.90**

n = 105

**1-tailed level of significance: $p < .01$

Table 28

TALAM Test-Retest Inter-Correlation Matrix (Pearson Product Moment)

	TALAM RETEST			
	Beliefs	Feelings	Int. beh.	Total
TALAM TEST				
Beliefs	**.77****	.60**	.32**	.70**
Feelings	.56**	**.80****	.52**	.75**
Int. beh.	.32**	.36**	**.70****	.48**
Total	.68**	.71**	.56**	**.82****

n = 105

**1-tailed significance level: $p < .001$

low achievers in mathematics." Table 29 summarizes both the calculation of Pearson product-moment and point-biserial coefficients.

Table 29

TALAM Correlations with Other Instruments

	Pearson product moment		Point biserial teaching level preference item
	R M A S	S D S	
TALAM SCALES			
Beliefs..........	.17	.50**	-.24*
Feelings.........	.19	.37**	-.24*
Int. behaviors	.11	.10	-.04
Total scale......	.19	.41**	-.28**

n = 105

**2-tailed significance level: * p < .05 ** p < .01

Where, "RMAS" is the Revised Mathematics Attitude Scale,
"SDS" is the semantic differential scale, and
"Teaching Level Preference Item" is the item requiring a dichotomous response from subjects.

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

This study had two major purposes. The first purpose was to develop an instrument to measure the attitudes of middle school mathematics teachers toward low achievers in mathematics. The second purpose was to establish the reliability and validity of the instrument. The study was carried out in three phases summarized below.

Summary

Phase 1

Phase 1 consisted of (1) selection of instrument type and format, (2) writing of potential scale items, (3) classification of item directionality by judges (positive, negative, or neutral with regard to low achievers in mathematics), and (4) preparation of the preliminary attitude scale.

After considering various affective scaling procedures, a modified Likert-type scale was determined as most appropriate for this study. Instead of the typical Likert scale allowing for five response categories, ranging from "strongly agree" to "strongly disagree" with a neutral category included, the number of response alternatives was modified to consist of six response categories ranging from "strongly agree" to "strongly disagree" without the neutral category option.

After determining instrument type and response alternatives, 128 statements about low achievers in mathematics were composed based on input from middle school mathematics teachers, research literature, related scales, and from consultation with experts in mathematics and mathematics education. The statements were categorized as follows: (1) teacher beliefs about low achievers in mathematics, (2) teacher feelings about working with low achievers in mathematics, and (3) intended teacher behaviors toward low achievers in mathematics. The categorized items were then studied by a panel of 10 judges who classified each statement as positive, negative, or neutral with regard to low achievers in mathematics. Items were eliminated if not classified by at least 90% of the judges as clearly positive or negative or if items appeared redundant.

Statements from each of the three categories mentioned above became the basis for an 85-item preliminary attitude scale. For administrative purposes, the preliminary attitude scale was not separated into three distinct sections. Rather, the items in each category were randomly distributed throughout the instrument. Further, a nearly even mix of positively and negatively worded items was selected from each category.

Phase 2

Phase 2 consisted of (1) administering the preliminary attitude scale to 51 middle school mathematics teachers from the Tri-cities region in Tennessee, (2) analyzing the data obtained and, (3) preparing the final attitude scale: Teacher Attitudes Toward Low Achievers in Mathematics Scale (TALAM).

The teacher response rate in Phase 2 was 100%. The data analyzed included the responses of all 51 teachers who completed the preliminary attitude scale: "Statements About Low Achievers in Mathematics." A total score and three subscale scores (Beliefs, Feelings, Intended Behaviors) were obtained for each subject. Through application of student's t-tests, significant differences were found among mean scores for the three subscales (Beliefs, Feelings, and Intended Behaviors) and total scale scores; thereby, indicating the need for three distinct subtest scores as opposed to one summated total score.

Item analysis was conducted for statements comprising the total 85 item preliminary scale and for statements comprising each of the three subscales: (1) beliefs - 44 items, (2) feelings - 20 items and, (3) intended behaviors - 21 items. All items were examined to determine the discriminating ability of the item. That is, when a positively written item is valid, subjects with a generally positive attitude respond "agree" or "strongly agree" to the item and those subjects with a generally negative attitude respond "disagree" or "strongly disagree" to the same item. The discriminating ability of items was calculated by establishing positive and negative criterion groups for each of the three subscales and the total scale. The mean score for each individual item was computed for high and low criterion groups and compared through the use of student's t statistic. Significant difference ($p < .05$) between high and low criterion group mean scores for each item was indicative of the ability of the item to discriminate adequately between positive and negative criterion groups.

Item analysis also included item-to-scale correlations for items within each subscale and within the total scale. High correlations between individual item scores and total scale scores suggested that the item represented the attitude under study. Items were eliminated from each subscale if item-to-scale correlations were not statistically significant.

Subscales and total scale analysis included Cronbach's alpha as a measure of internal-consistency reliability. Coefficient alpha's for the three subscales and the total scale ranged from .78 to .93, indicating significantly high ($p < .01$) inter-item correlation among scaled items.

After statements were eliminated on the basis of the logical and empirical criterion discussed above, 45 items were retained for use in the final form of the attitude scale: Teacher Attitudes Toward Low Achievers in Mathematics Scale (TALAM). The final 45-item scale was composed of three distinct subscales, each containing 15 items, designed to measure the following (1) teacher beliefs, (2) teacher feelings and, (3) intended teacher behaviors, all with respect to low achievers in mathematics. Each subscale and the total scale were again submitted to a test for internal-consistency reliability. The alpha on the final 45-item scale was .91 and the alpha for each of the three 15-item subscales ranged from .80 to .90.

For the purpose of further evaluating the construct validity for the three subscales, a principal components factor analysis was conducted with varimax rotation for all items comprising the 45-item final scale. This analysis yielded data indicating that three major factors accounted for 62% of total scale variance. The three primary factors emerging from factor analysis matched the original grouping of the items when they were written: beliefs, feelings, and intended behaviors. Before items were considered to "load" on a factor, that is, be assigned to a factor, they had to correlate with the factor with an r-value of .50 or higher on one factor and .40 or lower on the other two factors. Approximately 32 of the 45 items (71%) met these criteria. However, the 13 items failing to meet the criteria were retained because of their high item-total correlations ($p < .001$) and other favorable statistical data.

Phase 3

Phase 3 consisted of administering the following instruments to a sample of 105 (96%) of the middle school mathematics teachers in Knox County, Tennessee: (1) Teacher Attitudes Toward Low Achievers in Mathematics Scale (TALAM), developed for this study, (2) Revised Math Attitude Scale (RMAS) (Aiken & Dreger, 1963), (3) a semantic differential scale measuring teacher attitude toward low achievers (Steeg, 1983), (4) a single statement: "I prefer to teach mathematics to average or high achievers in mathematics rather than to low achievers in mathematics." requiring respondents to answer "True" or "False" and, (5) readministration of the TALAM after a period of two weeks.

Data treatment included scoring and analyzing the responses of the 105 subjects who completed all five of the instruments listed above. After TALAM administration, four summated scores for each subject were determined: the three 15-item subscales and the score obtained for the 45-item total scale. Analysis of variance (ANOVA) was undertaken to test the significance of the differences among mean scores for the three TALAM subscales and the total scale yielding a significant F value (p < .01). Subsequently, Scheffé's multiple comparison procedure was used to isolate the location of significant differences among mean scores for the TALAM subscales and total scale. Although no significant differences among means were found between any of the three subscales, the mean score for each subscale was significantly different from the mean score for the total scale (p < .05).

Test/retest scores were obtained for the TALAM. A Pearson product-moment correlation was employed for estimating the test-retest reliability of each of the three subscales and total scale. Significantly high (p < .01) coefficients, ranging from .70 to .82, resulted from comparison of the scores from the two administrations indicating the degree of reliability of the instrument.

Subscale and total scale scores from the first administration of the TALAM were compared to scores obtained from the same subjects on the Revised Math Attitude Scale (RMAS) (Aiken & Dreger, 1963.). Utilizing a Pearson product-moment correlation, weak coefficients ranging from .11 to .19 were found, suggesting little or no relationship between teacher attitude (beliefs, feelings, intended behaviors) toward low achievers in mathematics and their general attitude toward mathematics.

Subscale and total scale scores from the first administration of the TALAM were also compared to scores obtained from the same subjects on the semantic differential scale designed by Steeg (1983) to measure teacher attitude toward low achievers in general. Significant correlations (p < .01), ranging from .37 to .50, were found between semantic differential scale scores and scores obtained from the following TALAM scales: teacher belief, teacher feeling, and total TALAM score. Although statistically significant, correlations appeared moderate and not suggestive of a strong relationship between teacher attitude toward low achievers in general and toward low achievers in mathematics in particular. The TALAM subscale designed to measure intended teacher behaviors toward low achievers in mathematics did not correlate significantly with the semantic differential scale scores ($r = .10$). This result appears reasonable given that none of the 60 items on the semantic differential scale were designed to measure teacher behaviors toward low achievers.

Subscale and total scale scores from the first administration of the TALAM were correlated with subjects' response to the following statement: "I prefer to teach mathematics to average or high achievers in mathematics rather than to low achievers in mathematics." A point-biserial correlation was used to compare TALAM scores with the response of each subject to the teaching preference statement. Significantly negative correlations ($p < .05$) were found between the TALAM belief subscale and the teaching preference item and between the TALAM feelings subscale and the teaching preference item. A significantly negative correlation ($p < .01$) was also found between TALAM total scores and responses to the teaching preference item. No significant correlation was found between TALAM intended behavior subscale scores and responses to the teaching preference item. Although statistically significant negative correlations were found between three of the four TALAM scales and the teaching preference item, the correlations were weak, ranging from -.38 to -.24. If subjects stated a preference for teaching average or high achievers in mathematics rather than low achievers in mathematics, those same subjects tended to have lower scores on the TALAM than other individuals but not to a strong degree.

Employing Cronbach's alpha, internal-consistency estimates of reliability were computed for each of the three 15-item TALAM subscales and the total 45-item TALAM scale. High reliability coefficients ranging from .70 to .90 resulted indicating the reliability of the TALAM as a measure teacher attitudes toward low achievers in mathematics.

Conclusions

Analysis of data in the context of the research reviewed provides substantial evidence indicating that the following conclusions are warranted:

1. The principles of attitude scale construction in the social psychological literature can be applied to develop reliable and valid instrumentation to measure teacher attitude toward low achievers in mathematics. Moreover, the Likert-type scale, as suggested by several researchers cited earlier, appears highly suitable for assessing the affective characteristic of attitude.

2. The Teacher Attitudes Toward Low Achievers in Mathematics Scale (TALAM) was demonstrated to be a valid and reliable indicator of teacher beliefs, feelings and intended behaviors toward low achievers in mathematics; therefore, the TALAM appears to be a viable method for assessing teacher attitude toward such students.

2. Teacher attitude toward low achievers in mathematics appears to be multidimensional in nature. That is, teacher beliefs, feelings, and intended behaviors are highly related components of attitude but are also separate constructs that may or may not be consistent within individuals and among groups. Evidence for this contention is found in data obtained from administration of the preliminary attitude scale which indicated significant differences among every possible pair of subscale means and total scale mean. However, data obtained from administration of the final attitude scale indicated that significant differences between means occurred only between the total TALAM scale paired with each individual subscale.

2. Subjects appeared to respond to the TALAM scale in good faith and provide honest and serious reactions to items on the scale. This conclusion is based on data analysis indicating that subscale and total scale TALAM scores for the sample group appeared to approximate a normal distribution. Therefore, most of the middle school mathematics teachers involved in the study had moderately positive attitudes toward low achievers in mathematics. High test-retest correlation coefficients also indicate thoughtful responses to items.

3. Item-total correlation coefficients for the three TALAM subscales and total scale ranged from .42 to .75 leading to the conclusion that the TALAM is a valid scale.

4. When submitted to factor analysis, the majority of the statements clustered statistically into the same three groups (beliefs, feelings, intended behaviors) for which the statements had originally been written. Again, this is suggestive of a valid scale.

5. The three TALAM subscales and the TALAM total scale exhibited internal-consistency reliabilities ranging from .70 to .90, suggesting that the scales are reliable and verifying an interrelatedness within scales.

6. The TALAM scale and sub-scale scores were stable over time as demonstrated by test-retest reliability coefficients ranging from .70 to .82 with a two week interval, again supporting the conclusion that the scales are reliable.

7. An average inter-correlation coefficient of .68 between TALAM subscales and total scale verifies an interrelatedness between subscales and total scale.

Recommendations for Uses of the Scales

Throughout this report, the need for a scale designed to measure teacher attitude toward low achievers in mathematics has been stressed. The primary focus was to develop a much needed instrument for measuring attitudes and testing treatment effects in attitude research with mathematics teachers or preservice mathematics teachers. Although this need is thought to be the primary one, other uses of the scale are recommended.

1. TALAM subscale and total scale scores could be used as the basis for designing attitude profiles for individuals or groups. A profile sheet would contain a simple two-dimensional graph with the vertical axis labelled with the range of scores, over all scales, and the horizontal axis labelled with the attitude components measured by each of the subscales and total scale.

2. Administration of the TALAM scale could be used as the basis for integrating and stressing affective goals for students, teachers, and administrators as an integral part of the curriculum of the school.

3. The attitudes of faculty and staff toward low achievers in mathematics as measured by the TALAM can be important considerations in the adoption and implementation of school in-service or other instructional programs.

Recommendations for Future Study

1. Modification of wording of statements is an area for further consideration. For example, several variations of an item could be tried with sample populations to determine which wording appears most acceptable and to determine any significant differences between the variations. Further, test-retest studies could be conducted wherein positively worded items in the first testing could be changed to negatively worded items in the retest.

2. Administer the TALAM in a wide variety of geographic and socioeconomic areas, obtaining reliability and validity data for different groups of subjects. This broad based study could either further demonstrate the value of the TALAM or else indicate the need for revision under certain circumstances.

3. Determine differences in attitudes toward low achievers in mathematics from teachers with varying amounts of experience. Answer the following research question: What is the relationship between years of teaching experience and attitudes toward low achievers in mathematics?

4. Determine the relationship between teachers' attitudes toward low achievers in mathematics and low achievers' attitudes toward their mathematics teachers.

5. Further investigations into the usefulness of factor analysis in attitude scale construction is desirable, using large sample populations.

6. Further research is needed to determine the relationship between the various components of attitude. This research could help determine whether attitude is a unidimensional construct best measured though the affective component or multi-dimensional and best measured though separate subscale scores.

7. Further investigation into the relationship of the intended behavior component of attitude and overt behavior toward the attitudinal object is recommended to answer the following research question: Do the actions of subjects match their self-reported behavioral intentions?

A Final Word

Although scales can be devised to measure attitude, research suggests that attitude is complex and cannot be measured entirely by any quantitative calculation or index. Nevertheless, study must continue in an effort to more rigorously measure and determine attitudes, given the power of attitudes to effect not only the behavior of the holder of the attitude but, also, others under the influence of the individual. Consequently, attitude measurement must eventually become an integral part of teacher education programs, in-service study, and teacher evaluation. As a now anonymous child said many years ago when confronting teachers: "It's not how I look that's important, it's how you see me."

REFERENCES

REFERENCES

Aiken, L. R., Jr. (1972). Research on attitudes toward mathematics. Arithmetic Teacher, 19, 229-234.

Aiken, L. R., Jr. (1974). Two scales of attitude toward mathematics. Journal for Research in Mathematics Education, 5, 67-71.

Aiken, L. R., Jr. (1976). Update on attitudes and other affective variables in learning mathematics. Review of Educational Research, 46, 293-311.

Aiken, L. R., & Dreger, R. M. (1961). The effect of attitudes on performance in learning mathematics. Journal of Educational Psychology, 52, 19-24.

Aiken, L. R., Jr., & Dreger, R. M. (1963). Personality correlates of attitude toward mathematics. Journal of Educational Research, 56, 576-580.

Allport, G. W. (1935). Attitudes. In M. Fishbein (Ed.) (1967), Readings in attitude theory and measurement (pp. 1-13). New York: John Wiley & Sons, Inc.

American Psychological Association. (1966). Standards for educational and psychological tests and manuals. Washington, D.C.: Author.

Anastasi, A. (1976). Psychological testing (4th ed.). New York: The Macmillan Company.

Anastasi, A. (1986). Evolving concepts of test validation. Annual Reviews of Psychology, 37, 1-15.

Anderson, L. W. (1981). Assessing affective characteristics in the schools. Boston: Allyn and Bacon, Inc.

Bowling, J. M. (1977). Three scales of attitude toward mathematics. Dissertation Abstracts International, 37, 4927A-4928A.

Brophy, J. E. (1979). Teacher behavior and its effects. Journal of Educational Psychology, 71, 733-750.

Brophy, J. E. (1986). Teaching and learning mathematics: Where research should be going. Journal for Research in Mathematics Education, 17, 323-346.

Brophy, J. E., & Good, T. L. (1970). Teacher's communication of differential expectations for children's classroom performance: Some behavioral data. Journal of Educational Psychology, 61, 365-374.

Brown, C. A., Carpenter, T. P., Kouba, V. L., Lindquist, M. M., Silver, E. A., & Swafford, J. O. (1988). Secondary school results for the fourth NAEP mathematics assessment: Algebra, geometry, mathematical methods, and attitudes. Mathematics Teacher, 81, 337-347.

Buchanan, N. K. (1987). Factors contributing to mathematical problem-solving performance and exploratory study. Educational Studies in Mathematics, 18, 399-415.

Campbell, D. T. (1950) The indirect assessment of social attitudes. In M. Fishbein (Ed.)(1967), Readings in Attitude Theory and Measurement (pp.163-179). New York: John Wiley & Sons, Inc.

Carpenter, T. P., Cobitt, M. K., Kepner, H. S., Jr., Lindquist, M. M., & Reyes, R. R. (1980). Students' affective responses to mathematics: Secondary school results from national assessment. Mathematics Teacher, 73, 531-539.

Chein, I. (1950). Behavior theory and the behavior of attitudes: Some critical comments. In M. Fishbein (Ed.), Readings in Attitude theory and measurement (pp. 51-57). New York: John Wiley & Sons, Inc.

Childress, D. R. (1976). A study of the relationships between students' attitudes toward mathematics and their ratings of mathematics courses and mathematics instructors. Dissertation Abstracts International, 36, 4340A.

Confrey, J. (1986). A critique of teacher effectiveness research in mathematics education. Journal for Research in Mathematics Education, 17, 347-360.

Cook, S. W., & Sellitz, C. (1964). A multiple indicator approach to attitude measurement. Psychological Bulletin, 62, 36-55.

Crano, W. D., & Brewer, M. B. (1973). Principles of research in social psychology. New York: McGraw-Hill Book Company.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16, 297-334.

Cronbach, L. J. (1959). Construct validity in psychological tests. Psychological Bulletin, 52, 281-302.

Cronbach, L. J. (1960). Essentials of psychological testing. New York: Harper & Row, Publishers, Incorporated.

Cureton, E. E. (1951). Validity. In E. F. Lindquist (Ed.), Educational measurement (pp. 621-694). Washington, D.C.: American Council on Education.

Cureton, E. E. (1965). Reliability and validity: Basic assumptions and experimental designs. Educational and Psychological Measurement, 25, 327-346.

Dick, W., & Hagerty, N. (1971). Topics in measurement. New York: McGraw-Hill Book Company.

Dutton, W. H. (1954). Measuring attitudes toward arithmetic. Elementary School Journal, 55, 24-31.

Dutton, W. H. (1962). Attitudes of prospective elementary school teachers toward arithmetic. Arithmetic Teacher, 9, 418-424.

Dutton, W. J., & Blum, M. P. (1968). The measurement of attitudes toward arithmetic with a Likert-type test. Elementary School Journal, 68, 259-264.

Eccles, J. S. (1983). Expectancies, values, and academic behaviors. In J.T. Spence (Ed.), Achievement and Achievement Motivation (pp. 75-146). San Francisco: W.H. Freeman.

Eccles, J. S., Midgley, C., & Adler, T. F. (1984). Grade related changes in the school environment: Effects on achievement motivation. In J.G. Nicholls (Ed.), The Development of Achievement Motivation (pp. 283-331). Greenwich, CT: JAI Press.

Edwards, A. L. (1957). Techniques of attitude scale construction. New York: Appleton-Century-Crofts, Inc.

Edwards, A. L. (1972). Experimental design in psychological research (4th ed.). New York: Holt, Rinehart and Winston, Inc.

Ellingson, J. B. (1962). Evaluation of attitudes of high school students toward mathematics. Dissertation Abstracts, 23, 1604.

Feldlaufer, H., Midgley, C., & Eccles, J. S. (1988). Student, teacher, and observer perceptions of the classroom environment before and after the transition to junior high school. Journal of Early Adolescence, 8, 133-156.

Fennema, E., & Sherman, J. A. (1976). Fennema-Sherman mathematics attitudes scales: Instruments designed to measure attitudes toward the learning of mathematics by males and females. Catalog of Selected Documents in Psychology, 6, 31.

Ferguson, G. A. (1981). Statistical analysis in psychology and education (5th ed.). New York: McGraw-Hill Book Company.

Fishbein, M., & Ajzen, I. (1975). Belief, attitude, intention and behavior: An introduction to theory and research. Reading, Massachusetts: Addison-Wesley Pub. Co.

Foster, G., Algozzine, B., & Ysseldyke, J. (1980). Classroom teacher and teacher-in-training susceptibility to stereotypical bias. The Personnel and Guidance Journal, 59, 27-30.

Fox, D. J. (1969). The research process in education. New York: Holt, Rinehart and Winston, Inc.

Frary, R., & Ling J. (1983). A factor analytic study of mathematics anxiety. Educational and Psychological Measurement, 43, 985-993.

Garrett, H. E. (1947). Statistics in psychology and education. New York: Longmans, Green.

Gladstone, R., Deal, R., & Drevdahl, J. E. (1960). An exploratory study of remedial math. In M. E. Shaw & J. M. Wright (Eds.) (1967), Scales for the measurement of attitudes. New York: McGraw-Hill Book Company.

Good, T. L. (1970). Which pupils do teachers call on? Elementary School Journal, 70, 190-198.

Good, T. L. (1981). Teacher expectations and student perceptions: A decade of research. Educational Leadership, 38, 415-422.

Good, T. L., & Brophy, J. E. (1987). Looking in classrooms. New York: Harper & Row, Publishers.

Goodenough, W. H. (1944). A technique of scale analysis. Educational and Psychological Measurement, 4, 179-190.

Gorsuch, R. L. (1974). Factor analysis. Philadelphia: W. B. Saunders Company.

Gough, H. G. (1952). The adjective check list. Palo Alto, CA: Consulting Psychologists Press.

Green, W. (1991). Maintaining vitality. Mathematics Teacher, 81, 32.

Greenwald, A. G. (1968). On defining attitudes and attitude theory. In A.G. Greenwald (Ed.), Psychological foundations of attitudes (pp. 361-390). New York: Academic Press.

Guilford, J. P. (1946). New standards for test evaluation. Educational and Psychological Measurement, 6, 427-438.

Guilford, J. P. (1954). Psychometric methods (2nd ed.). New York: McGraw-Hill Book Company.

Guilford, J. P. (1959). Personality. New York: McGraw-Hill Book Company.

Guilford, J. P. (1965). Fundamental statistics in psychology and education (4th ed.). New York: McGraw-Hill Book Company.

Gulliksen, H. (1950). Theory of mental tests. New York: John Wiley & Sons, Inc.

Guttman, L. (1944). A basis for scaling qualitative data. Sociological Review, 9, 139-150.

Guttman, L. (1947). The cornell technique for scale and intensity analysis. Educational and Psychological Measurement, 7, 247-280.

Guttman, L., & Suchman, E. A. (1947). Intensity and a zero point for attitude analysis. In M. Fishbein (Ed.) (1967), Readings in attitude theory and measurement (pp. 267-276). New York: John Wiley & Sons, Inc.

Haladyna, T., Shaughnessy, J., & Shaughnessy, J. M. (1983). A causal analysis of attitude toward mathematics. Journal for Research in Mathematics Education, 14, 19-28.

Harman, H. H. (1967). Modern factor analysis. Chicago: University of Chicago Press.

Hart, L. E. (1989). Classroom processes, sex of student, and confidence in learning mathematics. Journal for Research in Mathematics Education, 20, 242-260.

Hassan, A. M., & Shrigley, R. L. (1984). Designing a likert scale to measure chemistry attitudes. School Science and Mathematics, 84, 659-669.

Heise, D. R. (1970). The semantic differential and attitude research. In G. F. Summers (Ed.), Attitude measurement (pp. 235-253). Chicago: Rand McNally & Company.

Holland, P. W. (1985). On the study of differential item performance and the Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), Test validity (pp. 129-145). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Hotelling, H. (1935). The most predictable criterion. Journal of Educational Psychology, 26, 139-142.

Hume, D. (1979). A scale to measure attitude toward one's mathematics teacher. (Doctoral dissertation, University of Tennessee, 1979). Dissertation Abstracts International, 40, 3171A.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. Psychometrika, 23, 187-200.

Kirsch, I. S., & Jungeblut, A. (1986). Literacy profiles of America's young adults. Princeton, N.J.: Educational Testing Service.

Krech, D., & Crutchfield, R. S. (1948). Theory and problems in social psychology. New York: McGraw-Hill Book Company.

Krech, D., Crutchfield, R. S., & Ballachey, E. L. (1962). Individual in society. New York: McGraw-Hill Book Company.

Kuder, G. F., & Richardson, M. W. (1937). The theory and estimation of test reliability. Psychometrika, 2, 151-160.

Kulm, G. (1980). Research on mathematics attitude. Journal for Research in Mathematics Education, 11, 356-381.

Lemon, N. (1973). Attitudes and their measurement. New York: John Wiley & Sons.

Likert, R. (1932). A technique for the measurement of attitudes. In G. F. Summers (Ed.) (1970). Attitude measurement, (pp. 149-158). Chicago, IL: Rand McNally & Company.

Lindquist, E. F. (1942). A first course in statistics. New York: Houghton Mifflin.

Lindquist, E. F. (Ed.). (1951). Educational measurement. Washington, D.C.: American Council on Education.

Ling, J. L. (1983). A factor-analytic study of mathematics anxiety. Dissertation Abstracts International, 43, 2266A. (University Microfilms No. 82-26,901)

McCallon, E. L., & Brown, J. D. (1971). Semantic differential instrument for measuring attitude toward mathematics. Journal for Experimental Education, 39, 69-72.

McKnight, C. C., Crosswhite, F. J., Dossey, J. A., Kifer, J. O., Swafford, J. O., Travers, K. J., & Cooney, T. J. (1987). The underachieving curriculum: Assessing U.S. school mathematics from an international perspective. Champaign, IL: Stipes Publishing Co.

McNeil, J. (1971). Toward accountable teachers: Their appraisal and improvement. New York: Holt, Rinehart and Winston.

Meek, A. (1989). On creating ganas: A conversation with Jaime Escalante. Educational Leadership, 46, 46-47.

Mehrens, W. A., & Ebel, R. L. (Eds.). (1967). Principles of educational and psychological measurement. Chicago: Rand McNally & Company.

Mehrens, W. A., & Lehmann, I. J. (1980). Standardized tests in education (3rd ed). New York: Holt, Rinehart and Winston.

Midgley, C., Feldlaufer, H., & Eccles J. S. (1989). Change in teacher efficacy and student self-and task-related beliefs in mathematics during the transition to junior high school. Journal of Educational Psychology, 81, 247-258.

Mueller, D. J. (1970). Physiological techniques of attitude measurement. In G. F. Summers (Ed.) (1970), Attitude measurement (pp. 534-552). Chicago, IL: Rand McNally & Company.

Mulaik, S. A. (1972). The foundations of factor analysis. New York: McGraw-Hill Book Company.

National Commission on Excellence in Education. (1983). A nation at risk. Washington, DC: Superintendent of Documents.

National Council of Teachers of Mathematics. (1991). Professional Standards for teaching mathematics. Reston, VA: Author.

National Council of Teachers of Mathematics. (1989). Curriculum and evaluation standards for school mathematics. Reston, VA: Author.

Norusis, M. J. (Ed.). (1988). SPSS/PC+ V2.0 base manual for the IBM PC/XT/AT and PS/2. Chicago: SPSS, Inc.

Nunnally, J. C. (1967). McGraw-Hill series in psychology: Psychometric theory. New York: McGraw-Hill Book Company.

Nunnally, Jr., J. C. (1970). Introduction to psychological measurement. New York: McGraw-Hill Book Company.

Oppenheim, A. N. (1966). Questionnaire design and attitude measurement. New York: Basic Books, Inc.

Osgood, C. E. (1952). The nature and measurement of meaning. In J. G. Snider & C. E. Osgood (Eds.) (1969), Semantic differential technique (pp. 3-41). Chicago: Aldine Publishing Company.

Osgood, C. E., & Suci, G. J. (1955). In J. G. Snider & C. E. Osgood (Eds.) (1969), Semantic differential technique (pp. 42-55). Chicago: Aldine Publishing Company.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). The measurement of meaning. Urbana, IL: University of Illinois Press.

Oskamp, S. (1991). Attitudes and opinions. (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Pambookian, H. (1976). Discrepancy between instructor and student evaluation of instruction: Effect on instruction. *Instructional Science*, *5*, 63-75.

Pederson, K., Bleyer, D. R., & Elmore, P. B. (1985). Attitudes and career interests in junior high school mathematics students: Implications for the classroom. *Arithmetic Teacher*, *33*, 45-48.

Phillips, R. B. (1973). Teacher attitude as related to student attitude and achievement in elementary school mathematics. *School Science and Mathematics*, *73*, 501-507.

Porier, G. W., & Lott, A. J. (1967). Galvanic skin responses and prejudice. In G. F. Summers (Ed.) (1970), *Attitude measurement* (pp. 489-496). Chicago: Rand McNally & Company.

Purkey, W. W., Cage, B., & Graves, M. (1973). The Florida Key: A scale to infer learner self-concept. *Educational and Psychological Measurement*, *33*, 979-984.

Quilter, D., & Harper, E. (1988). Why we didn't like mathematics, and why we can't do it. *Educational Research*, *30*, 121-134.

Richardson, F. C., & Suinn, R. M. (1972) The mathematics anxiety rating scale: Psychometric data. *Journal of Counseling Psychology*, *19*, 551-554.

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. New York: Holt, Rinehart & Winston.

Rosenthal, R., & Rubin, D. (1971). Appendix C: Pygmalion reaffirmed. In J. Elashoff & R. Snow (Eds.), *Pygmalion reconsidered*. Belmont, CA: Wadsworth Publishing.

Rosenthal, R. (1973). The pygmalion effect lives. *Psychology Today*, *7*, 56-63.

Rosenthal, R. (1974). *On the social psychology of the self-fulfilling prophecy: Further evidence for pygmalion effects and their mediating mechanisms*. New York: MSS Modular Publications.

Sandman, R. S. (1974). The development, validation, and application of a multidimensional mathematics attitude instrument. *Dissertation Abstracts International*, *34*, 7054A-7055A.

Scheffé, H. (1953). A method for judging all contrasts in the analysis of variance. *Biometrika*, *40*, 87-104.

Schoenfeld, A. H. (1989). Explorations of students' mathematical beliefs and behaviors. *Journal for Research in Mathematics Education*, *20*, 338-355.

Schunk, D. (1985). Self-efficacy and classroom learning. Psychology in the Schools, 22, 208-223.

Sellitz, C., Jahoda, M., Deutsch, M., & Cook, S. W. (1959). Attitude scaling. In M. Jahoda & N. Warren (Eds.) (1966), Attitudes (pp. 305-324). Baltimore: Penguin Books Inc.

Snider, J. G., & Osgood, C. E. (Eds.). (1969). Semantic differential technique. Chicago: Aldine Publishing Company.

Spielberger, C. D. (1978). Test anxiety inventory. University of South Florida.

Steeg, J. L. (1983). Behavioral and attitudinal changes of teachers toward low achieving students as a result of the TESA program. (Doctoral dissertation, United States International University, 1982). Dissertation Abstracts International, 43, 425A.

Steen, L. A. (1989). Teaching mathematics for tomorrow's world. Educational Leadership, 47, 18-22.

Stevenson, H. W. (1987). America's math problems. Educational Leadership, 45, 4-10.

Suinn, R. M. (1972). Mathematics anxiety rating scale. Fort Collins, CO: Rocky Mountain Behavioral Science Institute, Inc.

Suinn, R. M., Edie, C. A., Nicoletti, J., & Spinelli, P. R. (1972). The MARS, a measure of mathematics anxiety: Psychometric data. Journal of Clinical Psychology, 28, 373-375.

Thissen, D., Wainer, H., & Williams, D. M. (1984). Accounting for statistical artifacts in item bias research. Journal of Educational Statistics, 9, 93-128.

Thorndike, R. L. (1966). Reliability. In E. F. Lindquist (Ed.), Educational measurement (pp. 560-619). Menasha, WI: George Banta Publishing Company.

Thorndike, R. (1968). Review of pygmalion in the classroom. American Educational Research Journal, 5, 708-711.

Thurstone, L. L., & Chave, E. J. (1928). Attitudes can be measured. American Journal of Sociology, 33, 529-554.

Thurstone, L. L., & Chave, E. J. (1929). The measurement of attitude. Chicago: The University of Chicago Press.

Triandis, H. D. (1971). Attitude and attitude change. New York: John Wiley & Sons, Inc.

Troldahl, V. C., & Powell, F. A. (1965). A short-form dogamatism scale for use in field studies. Social Forces, 44, 221-224.

Wainer, H., & Braun H. I. (Eds.). (1988). Test validity. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Weiner, B. (1971). Achievement motivation and attribution theory. Morristown, NJ: General Learning Press.

Westie, F. R., & Defleur, M. L. (1959). Autonomic responses and their relationship to race attitudes. In G. F. Summers (Ed.) (1970), Attitude measurement (pp. 497-506). Chicago: Rand McNally & Company.

Woodmansee, J. J. (1970). The pupil response as a measure of social attitudes. In G. F. Summers (Ed.), Attitude measurement (pp. 514-533). Chicago: Rand McNally & Company.

APPENDICES

APPENDIX A

CATEGORIZED STATEMENTS (128) ABOUT LOW ACHIEVERS IN MATHEMATICS WITH LETTER AND INSTRUCTIONS FOR JUDGES

Dear Judges,

The statements appearing on the following pages are potential items for inclusion in a modified Likert scale which is being designed to measure:

1. Beliefs of middle school mathematics teachers about low achievers in mathematics
2. Feelings of middle school mathematics teachers toward low achievers in mathematics
3. Intended behaviors of middle school mathematics teachers with regard to low achievers in mathematics

Both the pilot instrument and the final scale will request middle grade mathematics teachers to read and react to statements appearing on the scale by choosing one of the following response categories: Strongly Agree, Agree, Slightly Agree, Slightly Disagree, Disagree, Strongly Disagree.

You can assist in the development of the attitude scale(s) by reading the 128 potential items and following the directions that appear for each section. Your participation will be extremely valuable in the development of appropriate items for inclusion in both the pilot study and the final scale(s).

When you have completed reading and reacting to the items, please mail your responses back to me at the Office of Field Studies, Claxton Addition, Room 216. Thank you for your time.

Sincerely,

Evelyn M. Dwyer

DIRECTIONS: The following items were designed to provide information relative to middle grade mathematics teachers' opinions about the cultural characteristics of low achievers in mathematics. This section requires only that you **read the items listed below and suggest** any alterations or additional items that you may feel apply. These items will not be part of the middle grade teachers' overall attitude score but will be used to establish a profile of beliefs for particular groups. For example: 80% of the group strongly agreed that . . . etc.

BELIEFS: CULTURAL CHARACTERISTICS

1. Females have less natural ability than males when it comes to learning mathematics.

2. African-American children have as much natural ability to do mathematics as Caucasian children.

3. Low achievers in mathematics usually have parents with very little formal education.

4. Low achievers in mathematics generally come from low-income families.

5. Low achievers in mathematics have parents who have low aptitudes for learning mathematics.

6. The parents of low achievers in mathematics do not value achievement in mathematics.

7. Low achievers in mathematics come from homes where they are neglected.

8. Students who live in inner city communities in large urban environments have less mathematical ability than students who live in other areas.

9. Low achievers in mathematics come from homes where there are low expectations concerning career possibilities for them.

10. Low achievers in mathematics have parents who do not believe their children will attend college.

DIRECTIONS: Please read the items listed below and classify each as positive (+), negative (-), or unsure (?). Place an (+) in front of those items you believe are clearly favorable statements about children who are low achievers in mathematics. Place a (-) in front of items you believe are clearly unfavorable statements about low achievers in mathematics. Place a (?) in front of items that appear neutral (neither positive or negative) in describing low achievers in mathematics.

Although statements might well apply to children other than low achievers in mathematics, the purpose here is to determine if the statements reflect positively, negatively, or in a neutral way towards children who are low achievers when it comes to the study of mathematics.

Please feel welcome to suggest additional items by writing them on the reverse side of the page. Also, feel free to make suggestions concerning the wording of any individual item by marking through or writing below each item.

Example: __+__ Low achievers in mathematics are "cool dudes."

BELIEFS: COGNITIVE ABILITY

____1. Low achievers in mathematics ask appropriate questions.

____2. Low achievers in mathematics have limited intellectual ability.

____3. Low achievers in mathematics can learn to solve application problems.

____4. Low achievers in middle grade mathematics can succeed in secondary school mathematics.

____5. Low achievers in mathematics have the capacity to think logically.

____6. Low achievers in mathematics have the ability to see how mathematical concepts are related.

____7. Low achievers in mathematics have the ability to remember what they have learned in mathematics class.

____8. Some children do not have the ability to succeed in mathematics regardless of how hard they try.

____9. Low achievers in mathematics usually do well in other school subjects.

____10. Low achievers in mathematics can successfully learn geometry and algebra.

____11. The most teachers should expect of low achievers is for them to learn basic arithmetic facts.

____12. Low achievers in mathematics are not capable of higher level thinking.

____13. Low achievers in mathematics have the ability to answer questions posed by the teacher during class.

____14. Low achievers in mathematics usually improve in their ability to do mathematics as they progress through school.

____15. Low achievers in mathematics learn new material slowly.

____16. Low achievers in mathematics have the ability to comprehend written directions for completing assignments.

____17. Low achievers in mathematics would achieve at higher levels if more were expected of them.

____18. One of the major problems with low achievers in mathematics is that they are disorganized in their thinking.

____19. Low achievers in mathematics cannot adequately explain how they arrived at answers to problems.

____20. Low achievers in mathematics usually do not comprehend what is explained in class.

BELIEFS: WORK RELATED BEHAVIORS

____1. Low achievers in mathematics do not stay on-task during mathematics class.

____2. Low achievers in mathematics are curious about how answers to mathematics problems are determined.

____3. Low achievers in mathematics expect too much help from the teacher during mathematics class.

____4. Low achievers in mathematics are good listeners in mathematics class.

____5. Low achievers in mathematics do not like challenging assignments.

____6. Low achievers in mathematics rely too much on calculators.

____7. Low achievers in mathematics lack the self-discipline necessary to study mathematics effectively.

____8. Low achievers in mathematics usually complete their homework.

____9. The work of low achievers in mathematics is usually messy and disorganized.

____10. Most of the mathematics homework of low achievers is done correctly.

____11. Low achievers in mathematics interrupt the teacher to ask questions too often.

____12. Low achievers in mathematics are conscientious about bringing materials (pencils, texts, etc.) to class.

____13. Low achievers in mathematics would make better grades if they would be quiet and listen in mathematics class.

____14. Low achievers are not willing to work hard to do well in mathematics.

____15. Low achievers in mathematics usually participate in class discussions.

____16. Low achievers in mathematics persevere when faced with difficult problems.

____17. Low achievers in mathematics would rather work alone than in a group.

____18. Low achievers in mathematics usually do not communicate well verbally.

____19. Low achievers in mathematics often try to divert class attention from mathematics to areas of interest to them.

____20. Low achievers get more than their fair share of help from mathematics teachers.

BELIEFS: AFFECTIVE CHARACTERISTICS

____1. Low achievers in mathematics are usually polite.

____2. Low achievers in mathematics are usually pleasant.

____3. Low achievers in mathematics are generally energetic.

____4. Low achievers in mathematics are too retiring.

____5. Low achievers in mathematics are frequently impatient.

____6. Low achievers in mathematics are not serious enough.

____7. Low achievers in mathematics have just as many friends as other students.

____8. Low achievers in mathematics dislike coming to mathematics class.

____9. Low achievers in mathematics value the learning of mathematics.

____10. Low achievers in mathematics are not interested in learning mathematics.

____11. Low achievers in mathematics are so afraid of making mistakes that they won't even try.

____12. Low achievers in mathematics do not like themselves.

____13. Low achievers in mathematics cannot handle criticism.

____14. Low achievers in mathematics are not sociable.

____15. Low achievers in mathematics are usually honest on mathematics tests.

____16. Low achievers in mathematics are inconsiderate of other class members.

____17. Low achievers in mathematics are inconsiderate of me.

____18. Low achievers in mathematics are usually proud of their efforts.

____19. Low achievers in mathematics do not appreciate what a mathematics teacher does for them.

____20. Low achievers in mathematics cannot be expected to enjoy mathematics class.

____21. Low achievers in mathematics are usually well-behaved in mathematics class.

____22. Low achievers in mathematics have little imagination.

____23. Low achievers in mathematics are usually creative individuals.

____24. Low achievers exhibit distrust and hostility toward mathematics teachers.

____25. Teacher comments and suggestions are generally welcomed by low achievers in mathematics.

____26. Low achievers in mathematics often exhibit concern for others.

DIRECTIONS: The following directions apply to the remainder of the items appearing on pages 8 - 11. These items focus on the middle grade mathematics teacher and his/her feelings and beliefs about working with low achievers in mathematics.

Please read each item and think:

"If I were a middle grade mathematics teacher and felt or believed the following statement, would that feeling or belief be a positive statement in regard to working with low achievers (+), a negative statement in regard to working with low achievers (-), or is the statement a neutral one (?)."

For example: __-__ I am unhappy when working with low achievers in mathematics.

After making your decision, mark each item as (+) or (-) or (?).

TEACHER FEELINGS ABOUT WORKING WITH LOW ACHIEVERS

____1. I don't have the skill to work with low achievers in mathematics.

____2. I have the patience to work with low achievers in mathematics.

____3. I cannot succeed when working with low achievers in mathematics.

____4. I find it fatiguing to work with low achievers in mathematics.

____5. I communicate well with low achievers in mathematics.

____6. I can identify with the difficulties that low achievers experience in mathematics.

____7. If I put in enough effort, I can get through to low achievers in mathematics.

____8. I can influence the self-confidence of low achievers in mathematics.

____9. It is very irritating to me to work with students who are slow in mathematics.

____10. In the long run, whatever I do with low achievers in mathematics will not make any difference.

____11. I find it boring to teach mathematics to low achievers.

____12. I seldom find teaching low achievers in mathematics enjoyable.

____13. I find it difficult to care about the success of low achievers in mathematics.

____14. Working with low achievers in mathematics is rewarding to me.

____15. Low achievers in mathematics do not like me.

____16. I have no particular like or dislike for low achievers in mathematics.

INTENDED BEHAVIORS WITH REGARD TO TEACHING LOW ACHIEVERS

____1. Teachers should take the time to provide low achievers with opportunities to experience success in mathematics class.

____2. A disproportionate amount of funds are expended on low achievers in mathematics.

____3. Having low achievers in the classroom is a burden to the mathematics teacher.

____4. The trouble with low achievers in mathematics is that they have been socially promoted too often.

____5. Having low achievers in the regular mathematics classroom is a major source of "teacher burnout."

____7. Teachers get little return for their efforts when they try to teach low achievers in mathematics.

____8. Low achievers in mathematics have it way too easy in class today.

____9. The best mathematics teachers should be working with the low achievers in mathematics.

____10. There is a good chance that low achievers in mathematics would increase achievement if they were taught study skills.

____11. Small group instruction is valuable for low achievers in mathematics.

____12. Mathematics teachers need to be patient and listen to the answers that low achievers give to questions.

____13. Low achievers in mathematics should get a great deal of individual attention from their teachers.

____14. Teachers should make time for low achievers in mathematics to work with computers.

____15. Low achievers in mathematics need to be given more practice sheets.

____16. Low achievers in mathematics would do better if they would just memorize the rules.

____17. Low achievers in middle grade mathematics can benefit from using manipulatives.

____18. The mathematics teacher should take time to discuss mathematics with the low achiever.

____19. Low achievers in mathematics should be encouraged to write problems based on everyday experiences.

____20. Low achievers in mathematics might improve if they could see how math relates to other curriculum areas.

____24. Enrichment activities are not suitable for low achievers in math.

____25. Sometimes it does a low achiever in mathematics good to be criticized in the presence of other students.

____26. Having low achievers in my mathematics class hinders the progress of the whole class.

____27. There are times when a mathematics teacher cannot be blamed for losing patience with a low achiever.

____28. By trying different teaching methods, teachers can help improve student achievement in mathematics.

____29. Pupils who are slow in mathematics should be reminded of their limitations when they try to tackle problems too difficult for them.

____30. Low achievers in mathematics should be placed in low ability groups.

____31. Low achievers in mathematics should have the same mathematics texts as other students.

____32. Low achievers in mathematics should experience the same mathematics curriculum as other students.

____33. Low achievers in mathematics should be placed in vocational tracks as soon as possible.

____34. The mathematics teacher should provide a great deal of positive reinforcement for low achievers in mathematics.

____35. Mathematics teachers should help low achievers become more confident in their ability to do mathematics.

____36. Low achievers in mathematics deliberately try to make it hard on the mathematics teacher.

APPENDIX B

PRELIMINARY ATTITUDE SCALE WITH DIRECTIONS TO RESPONDENTS

STATEMENTS ABOUT LOW ACHIEVERS IN MATHEMATICS

The purpose of this survey is to determine attitudes of middle school teachers toward students who are low achievers in mathematics. You will be responding anonymously. Please answer expressing your own opinion. Only numerical data will be obtained for research purposes. No individual or individual school will be identified in any way nor will any report be issued to administrators of your school. Thank you for taking time to provide this important information.

This survey consists of a series of statements about low achievers in mathematics or about teaching low achievers in mathematics. Circle the number that indicates the degree to which you agree or disagree with each statement.

1 I STRONGLY DISAGREE with the statement.
2 I DISAGREE with the statement.
3 I SLIGHTLY DISAGREE with the statement.
4 I SLIGHTLY AGREE with the statement.
5 I AGREE with the statement.
6 I STRONGLY AGREE with the statement.

1. In the long run, whatever I do with low achievers in mathematics will not make any difference in their achievement level.
1 2 3 4 5 6

2. Females have less natural ability than males when it comes to learning mathematics.
1 2 3 4 5 6

3. Low achievers in mathematics deliberately try to make it hard on me in class.
1 2 3 4 5 6

4. Low achievers in mathematics have the ability to remember what they have learned in mathematics class.
1 2 3 4 5 6

5. Low achievers in mathematics usually have parents with very little formal education.
1 2 3 4 5 6

6. Low achievers in mathematics should have the same mathematics texts as other students.
1 2 3 4 5 6

7. Low achievers in mathematics are usually proud of their efforts.
1 2 3 4 5 6

8. Students who live in inner city communities in large urban environments have less mathematical ability than students who live in other areas.

1 2 3 4 5 6

9. The parents of low achievers in mathematics do not value achievement in mathematics.

1 2 3 4 5 6

10. Low achievers in middle grade mathematics can succeed in secondary school mathematics.

1 2 3 4 5 6

11. I find it boring to teach mathematics to low achievers.

1 2 3 4 5 6

12. The mathematics teacher should provide a great deal of positive reinforcement for low achievers in mathematics.

1 2 3 4 5 6

13. Low achievers in mathematics should be placed in vocational tracks requiring little mathematics as soon as possible.

1 2 3 4 5 6

14. Having low achievers in my mathematics class is a burden.

1 2 3 4 5 6

15. Low achievers in mathematics have the ability to comprehend written directions for completing assignments.

1 2 3 4 5 6

16. Low achievers in mathematics cannot handle constructive criticism.

1 2 3 4 5 6

17. I seldom find teaching low achievers in mathematics enjoyable.

1 2 3 4 5 6

18. I have the patience to work with low achievers in mathematics.

1 2 3 4 5 6

19. Low achievers in mathematics usually participate in class discussions.

1 2 3 4 5 6

20. Low achievers in mathematics should get a great deal of individual attention from their teachers.

1 2 3 4 5 6

21. Low achievers in mathematics have parents who do not believe their children will attend college.

1 2 3 4 5 6

22. Low achievers in mathematics are not sociable.
1 2 3 4 5 6

23. African-American children have as much natural ability to do mathematics as Caucasian children.
1 2 3 4 5 6

24. Low achievers in mathematics need to be given more practice sheets.
1 2 3 4 5 6

25. Low achievers should be encouraged to explore and discover mathematics through the use of computers.
1 2 3 4 5 6

26. Low achievers in mathematics generally come from low-income families.
1 2 3 4 5 6

27. Low achievers in mathematics can learn to solve useful application problems.
1 2 3 4 5 6

28. Low achievers in mathematics are inconsiderate of other class members.
1 2 3 4 5 6

29. Low achievers in mathematics would do better if they would just memorize the rules and procedures.
1 2 3 4 5 6

30. I cannot succeed when working with low achievers in mathematics.
1 2 3 4 5 6

31. It is very irritating to me to work with students who are slow in mathematics.
1 2 3 4 5 6

32. Sometimes it does a low achiever in mathematics good to be criticized in the presence of other students.
1 2 3 4 5 6

33. Low achievers in mathematics are usually pleasant.
1 2 3 4 5 6

34. Mathematics teachers need to be patient and listen to low achievers verbalize their thought processes.
1 2 3 4 5 6

35. Low achievers in mathematics lack the self-discipline necessary to study mathematics effectively.
1 2 3 4 5 6

36. Low achievers in mathematics expect too much help from the teacher during mathematics class.
1 2 3 4 5 6

37. By trying different teaching methods, teachers can help improve student achievement in mathematics.
1 2 3 4 5 6

38. Teachers should provide opportunities for low achievers to experience success in mathematics.
1 2 3 4 5 6

39. I like the challenge of working with low achievers in mathematics.
1 2 3 4 5 6

40. Low achievers in mathematics can successfully learn geometry and algebra.
1 2 3 4 5 6

41. I can identify with the difficulties that low achievers experience in mathematics.
1 2 3 4 5 6

42. Low achievers in mathematics are not capable of higher level thinking.
1 2 3 4 5 6

43. I feel angry when assigned to teach low achievers in mathematics.
1 2 3 4 5 6

44. Low achievers in mathematics are usually creative individuals.
1 2 3 4 5 6

45. Low achievers in mathematics usually do not communicate well verbally.
1 2 3 4 5 6

46. Having low achievers in my mathematics class hinders the progress of the whole class.
1 2 3 4 5 6

47. Low achievers in mathematics are usually polite.
1 2 3 4 5 6

48. Low achievers in mathematics persevere when faced with difficult problems.
1 2 3 4 5 6

49. Low achievers exhibit distrust and hostility toward mathematics teachers
1 2 3 4 5 6

50. Low achievers in mathematics usually do not comprehend what is explained to them in class.

1 2 3 4 5 6

51. I find it difficult to care about the success of low achievers in mathematics.

1 2 3 4 5 6

52. Low achievers in mathematics come from homes where there are low expectations concerning career possibilities for them.

1 2 3 4 5 6

53. Low achievers in mathematics should be placed in low ability groups.

1 2 3 4 5 6

54. If I put in enough effort, I can get through to low achievers in mathematics.

1 2 3 4 5 6

55. I communicate well with low achievers in mathematics.

1 2 3 4 5 6

56. Low achievers in mathematics have limited intellectual ability.

1 2 3 4 5 6

57. Low achievers in mathematics are usually honest on mathematics tests.

1 2 3 4 5 6

58. Low achievers in mathematics are usually well-behaved in mathematics class.

1 2 3 4 5 6

59. Low achievers in middle grade mathematics can benefit from using manipulatives.

1 2 3 4 5 6

60. Low achievers in mathematics do not like themselves.

1 2 3 4 5 6

61. Low achievers in mathematics are not interested in learning mathematics.

1 2 3 4 5 6

62. Low achievers in mathematics should be encouraged to write problems based on everyday experiences.

1 2 3 4 5 6

63. Low achievers in mathematics have the capacity to think logically.

1 2 3 4 5 6

64. Low achievers in mathematics come from homes where they are neglected.
1 2 3 4 5 6

65. The most teachers should expect of low achievers is for them to learn basic arithmetic facts.
1 2 3 4 5 6

66. Low achievers in mathematics are curious about how answers to mathematics problems are determined.
1 2 3 4 5 6

67. Low achievers in mathematics are frequently impatient.
1 2 3 4 5 6

68. Most of my interactions with low achievers in mathematics are positive.
1 2 3 4 5 6

69. Low achievers in mathematics should experience the same mathematics curriculum as other students.
1 2 3 4 5 6

70. Low achievers in mathematics usually complete their homework.
1 2 3 4 5 6

71. Low achievers in mathematics are good listeners in mathematics class.
1 2 3 4 5 6

72. Low achievers in mathematics have parents who have low aptitudes for learning mathematics.
1 2 3 4 5 6

73. I find it rewarding to work with low achievers in mathematics.
1 2 3 4 5 6

74. Low achievers in mathematics do not like me.
1 2 3 4 5 6

75. There is a good chance that low achievers in mathematics would increase achievement if they were taught study skills.
1 2 3 4 5 6

76. Low achievers in mathematics value the learning of mathematics.
1 2 3 4 5 6

77. The work of low achievers is usually messy and disorganized.
1 2 3 4 5 6

78. I feel frustrated when trying to teach low achievers in mathematics.

1 2 3 4 5 6

79. Small group instruction is valuable for low achievers in mathematics.

1 2 3 4 5 6

80. Mathematics teachers should provide opportunities for low achievers to see the usefulness of mathematics.

1 2 3 4 5 6

81. Enrichment activities are not suitable for low achievers in mathematics.

1 2 3 4 5 6

82. Students who are low achievers in mathematics need to be reminded of their limitations when they try to tackle problems too difficult for them.

1 2 3 4 5 6

83. Low achievers in mathematics are generally energetic.

1 2 3 4 5 6

84. Low achievers in mathematics do not like challenging assignments.

1 2 3 4 5 6

85. I can influence the confidence that low achievers have in their ability to do mathematics.

1 2 3 4 5 6

APPENDIX C

TEACHER ATTITUDES TOWARD LOW ACHIEVERS IN MATHEMATICS SCALE (TALAM) AND SCORING INSTRUCTIONS

TEACHER ATTITUDES TOWARD LOW ACHIEVERS IN MATHEMATICS (TALAM)

Directions: This survey consists of a series of statements about low achievers in mathematics or about teaching low achievers in mathematics. Circle the number that indicates the degree to which you agree or disagree with each statement.

1 I **STRONGLY DISAGREE** with the statement.
2 I **DISAGREE** with the statement.
3 I **SLIGHTLY DISAGREE** with the statement.
4 I **SLIGHTLY AGREE** with the statement.
5 I **AGREE** with the statement.
6 I **STRONGLY AGREE** with the statement.

1. In the long run, whatever I do with low achievers in mathematics will not make any difference in their achievement level.
1 2 3 4 5 6

2. Low achievers in mathematics have the ability to remember what they have learned in mathematics class.
1 2 3 4 5 6

3. I find it boring to teach mathematics to low achievers.
1 2 3 4 5 6

4. The mathematics teacher should provide a great deal of positive reinforcement for low achievers in mathematics.
1 2 3 4 5 6

5. Low achievers in mathematics should be placed in vocational tracks requiring little mathematics as soon as possible.
1 2 3 4 5 6

6. Having low achievers in my mathematics class is a burden.
1 2 3 4 5 6

7. I seldom find teaching low achievers in mathematics enjoyable.
1 2 3 4 5 6

8. I have the patience to work with low achievers in mathematics.
1 2 3 4 5 6

9. Low achievers in mathematics need to be given more practice sheets.
1 2 3 4 5 6

10. I cannot succeed when working with low achievers in mathematics.
1 2 3 4 5 6

11. I find it irritating to work with students who are slow in mathematics.
1 2 3 4 5 6

12. Sometimes it does a low achiever in mathematics good to be criticized in front of other students.
1 2 3 4 5 6

13. Low achievers in mathematics are usually pleasant.
1 2 3 4 5 6

14. Mathematics teachers need to be patient and listen to low achievers verbalize their thought processes.
1 2 3 4 5 6

15. Low achievers in mathematics lack the self-discipline necessary to study mathematics effectively.
1 2 3 4 5 6

16. By trying different teaching methods, teachers can help improve student achievement in mathematics.
1 2 3 4 5 6

17. Teachers should provide opportunities for low achievers to experience success in mathematics.
1 2 3 4 5 6

18. I like the challenge of working with low achievers in mathematics.
1 2 3 4 5 6

19. I feel angry when assigned to teach low achievers in mathematics.
1 2 3 4 5 6

20. Having low achievers in my mathematics class hinders the progress of the whole class.
1 2 3 4 5 6

21. Low achievers in mathematics are usually polite.
1 2 3 4 5 6

22. Low achievers exhibit distrust and hostility toward mathematics teachers
1 2 3 4 5 6

23. Low achievers in mathematics usually do not comprehend what is explained to them in class.
1 2 3 4 5 6

24. I find it difficult to care about the success of low achievers in mathematics.
1 2 3 4 5 6

25. Low achievers in mathematics should be placed in low ability groups.
1 2 3 4 5 6

26. I communicate well with low achievers in mathematics.
1 2 3 4 5 6

27. Low achievers in mathematics are usually well-behaved in mathematics class.
1 2 3 4 5 6

28. Low achievers in middle grade mathematics can benefit from using manipulatives.
1 2 3 4 5 6

29. Low achievers in mathematics are not interested in learning mathematics.
1 2 3 4 5 6

30. Low achievers in mathematics should be encouraged to write problems based on everyday experiences.
1 2 3 4 5 6

31. The most teachers should expect of low achievers is for them to learn basic arithmetic facts.
1 2 3 4 5 6

32. Low achievers in mathematics are frequently impatient.
1 2 3 4 5 6

33. Most of my interactions with low achievers in mathematics are positive.
1 2 3 4 5 6

34. Low achievers in mathematics should experience the same mathematics curriculum as other students.
1 2 3 4 5 6

35. Low achievers in mathematics usually complete their homework.
1 2 3 4 5 6

36. Low achievers in mathematics are good listeners in mathematics class.
1 2 3 4 5 6

37. I find it rewarding to work with low achievers in mathematics.
1 2 3 4 5 6

38. Low achievers in mathematics value the learning of mathematics.
1 2 3 4 5 6

39. The work of low achievers is usually messy and disorganized.
1 2 3 4 5 6

40. I feel frustrated when trying to teach low achievers in mathematics.
1 2 3 4 5 6

41. Mathematics teachers should provide opportunities for low achievers to see the usefulness of mathematics.
1 2 3 4 5 6

42. Enrichment activities are not suitable for low achievers in mathematics.
1 2 3 4 5 6

43. Students who are low achievers in mathematics need to be reminded of their limitations when they try to tackle problems too difficult for them.
1 2 3 4 5 6

44. Low achievers in mathematics are generally energetic.
1 2 3 4 5 6

45. Low achievers in mathematics do not like challenging assignments.
1 2 3 4 5 6

Teacher Attitudes Toward Low Achievers in Mathematics Scale (TALAM)

Evelyn E. Dwyer

Description

Three 15-item scales are contained in the 45-items listed. All 45 items should be administered at one time. However, the scales should be scored separately in order to obtain information about the different dimensions of teachers' attitudes toward low achievers in mathematics. The following chart indicates the factor which each item measures.

	Feelings Factor	Beliefs Factor	Intended Beh.Factor
Item Number	1, 3, 6, 7, 8, 10, 11, 18, 19, 20, 24, 26, 33, 37, 40	2, 13, 15, 21, 22, 23, 27, 29, 32, 35, 36, 38 39, 44, 45	4, 5, 9, 12, 14, 16, 17, 25, 28, 30, 31, 34, 41, 42, 43

Scoring of the Scales

On each scale some items are worded positively ("Working with low achievers in mathematics is fun."), and some are worded negatively ("Low achievers in mathematics are behavior problems."). **Positive** and **negative** items are scored as follows:

POSITIVE

Marked	Score
Strongly Disagree	= 1
Disagree	= 2
Slightly Disagree	= 3
Slightly Agree	= 4
Agree	= 5
Strongly Agree	= 6

NEGATIVE

Marked	Score
Strongly Disagree	= 6
Disagree	= 5
Slightly Disagree	= 4
Slightly Agree	= 3
Agree	= 2
Strongly Agree	= 1

Scoring of items in the three scales is summarized below:

Teacher Belief Scale

Item	Str Dis	Disagree	Sl Dis	Sl Agree	Agree	Str Agree
2	1	2	3	4	5	6
13	1	2	3	4	5	6
15	6	5	4	3	2	1
21	1	2	3	4	5	6
22	6	5	4	3	2	1
23	6	5	4	3	2	1
27	1	2	3	4	5	6
29	6	5	4	3	2	1
32	6	5	4	3	2	1
35	1	2	3	4	5	6
36	1	2	3	4	5	6
38	1	2	3	4	5	6
39	6	5	4	3	2	1
44	1	2	3	4	5	6
45	6	5	4	3	2	1

Teacher Feelings Scale

Item	Str Dis	Disagree	Sl Dis	Sl Agree	Agree	Str Agree
1	6	5	4	3	2	1
3	6	5	4	3	2	1
6	6	5	4	3	2	1
7	6	5	4	3	2	1
8	1	2	3	4	5	6
10	6	5	4	3	2	1
11	6	5	4	3	2	1
18	1	2	3	4	5	6
19	6	5	4	3	2	1
20	6	5	4	3	2	1
24	6	5	4	3	2	1
26	1	2	3	4	5	6
33	1	2	3	4	5	6
37	1	2	3	4	5	6
40	6	5	4	3	2	1

Teacher Intended Behavior Scale

Item	Str Dis	Disagree	Sl Dis	Sl Agree	Agree	Str Agree
4	1	2	3	4	5	6
5	6	5	4	3	2	1
9	6	5	4	3	2	1
12	6	5	4	3	2	1
14	1	2	3	4	5	6
16	1	2	3	4	5	6
17	1	2	3	4	5	6
25	6	5	4	3	2	1
28	1	2	3	4	5	6
30	1	2	3	4	5	6
31	6	5	4	3	2	1
34	1	2	3	4	5	6
41	1	2	3	4	5	6
42	6	5	4	3	2	1
43	6	5	4	3	2	1

TALAM Scale Score Sheet

Beliefs		Feelings		Int. Behav.	
+ 2	____	- 1	____	+ 4	____
+13	____	- 3	____	- 5	____
-15	____	- 6	____	- 9	____
+21	____	- 7	____	-12	____
-22	____	+ 8	____	+14	____
-23	____	-10	____	+16	____
+27	____	-11	____	+17	____
-29	____	+18	____	-25	____
-32	____	-19	____	+28	____
+35	____	-20	____	+30	____
+36	____	-24	____	-31	____
+38	____	+26	____	+34	____
-39	____	+33	____	+41	____
+44	____	+37	____	-42	____
-45	____	-40	____	-43	____
B TOTAL	____	F TOTAL	____	I TOTAL	____

Optional: Total Scale Score (B+F+I) =

For **Negative Items**

Strongly Disagree	= 6
Disagree	= 5
Slightly Disagree	= 4
Slightly Agree	= 3
Agree	= 2
Strongly Agree	= 1

For **Positive Items**

Strongly Disagree	= 1
Disagree	= 2
Slightly Disagree	= 3
Slightly Agree	= 4
Agree	= 5
Strongly Agree	= 6

APPENDIX D

ADMINISTRATION PROCEDURE HANDOUTS

FOR FINAL ATTITUDE SCALE

ADMINISTRATION OF ATTITUDE SCALES

1. Please ask all mathematics teachers in your school to meet together for approximately 20 minutes.

2. Inform teachers of the following:

A. The purpose of the research is to determine teacher attitudes about low achievers in mathematics.

B. Participation of teachers will be greatly appreciated and will contribute in an important way to the research in the area of mathematics education. However, participation is strictly voluntary.

C. All responses will be anonymous. No individuals will be identified in any way nor will any individual school be identified. Only numerical data for the total group of approximately 115 subjects will be reported.

D. Total honesty in responding is critical. There are no "right" or "wrong" answers. What is needed is the expression of your own belief concerning each of the items presented.

E. Please respond to EVERY item!

F. Make an appointment for participants to meet together again in two weeks to retake the Teacher Attitude Toward Low Achievers in Mathematics Scale (TALAM). This re-test is critical to the success of the study. Scores on the first and second administration of the TALAM must be compared for each individual; consequently, subjects need to remember the code number used from the first administration of the TALAM. For example, the subject with folder number 24 in the first administration must retake the TALAM at the second administration using retest number 24.

2. It is desirable that all subjects in your school be administered the instruments on the same day at the same time. However, if this is not possible for a few, then the instruments may be completed and returned to you at the earliest opportunity.

3. A day for pickup of instruments will be arranged with you.

4. If you have any questions, please feel welcome to call me collect in Johnson City at 926-6822 or leave a message for me in the Office of Field Studies at The University of Tennessee (974-0868). On Wed. and Thurs. evenings, I can usually be reached in Knoxville at 558-8230 or a message can be left there at any time for me.

THANK YOU VERY MUCH FOR YOUR ASSISTANCE AND COOPERATION!

Evelyn M. Dwyer

ADMINISTRATOR RECORDS

School________________________________

Date of first test administration____________________

Date of re-test administration____________________

Number of mathematics teachers in your school____________

Number of mathematics teachers in your school who:

participated in first testing____________________

participated in second testing____________________

Folder numbers for first test:

used____________________________

unused____________________________

RE-TEST instrument numbers:

used____________________________

unused____________________________

VITA

Evelyn Elizabeth Dwyer was born in Erwin, Tennessee, on August 26, 1952. She holds a Bachelor of Arts Degree in Secondary Education from Berea College in Berea, Kentucky and a Master of Arts in Teaching Degree from East Tennessee State University in Johnson City, Tennessee. This dissertation completes requirements for a Ph.D. in Education with an emphasis in Mathematics Education at The University of Tennessee in Knoxville, Tennessee.

Ms. Dwyer has taught in public and parochial schools in Berea and Lexington, Kentucky; Greeneville, Johnson City, and Kingsport, Tennessee. She also has experiences teaching adult basic education and remedial classes for high risk secondary school students. She has taught developmental studies reading, study skills, and mathematics at the college level. As a graduate assistant for the Department of Curriculum and Instruction at The University of Tennessee, Ms. Dwyer supervised field-based students, micro-teaching simulations, and conducted seminars for pre-service teachers. As a graduate teaching assistant for the Office of Field Studies at The University of Tennessee, Ms. Dwyer supervised elementary and secondary school student teachers.