AN ABSTRACT OF THE THESIS OF

Denr	nis Lee Crawford	for the degree of	Doctor of Philosophy		
in	Education	presented on	June 7, 1976		
Title;	Title: DIFFERENTIAL CULTURAL AWARENESS:				
	DEVELOPING AN	INSTRUMENT FOR	TESTING		
	THE CULTURAL A	AWARENESS OF TE	CACHERS		
		Redacted for P			
Abstract approved:					
	-	Gerald Be	cker		

The original purpose of this study was the development of an instrument to measure the cultural awareness of teachers. It was expected that scores on this instrument would be significantly correlated and positively related to affective teaching behavior. However, the logic of this position failed to be supported by the research findings. Hence, the problem under investigation shifts from demonstrating an expected relationship to one of logically critiquing or examining both the conventional wisdom about, and quantitative measurement of cultural attitudes and affective teaching behavior.

The instrument developed was a special Likert type attitude inventory. A pretest was first designed with 160 questions. These questions were formulated by the author and other competent people in the field of minority education in Oregon. The questions were sorted into six broad sub-categories related to the research problem.

Following the pretest an internal consistency method of item analysis was employed to select the 58 items that were used in the final instrument. The instrument developed was called the "Differential Cultural Awareness" and referred to hereafter as the DCA.

The DCA was administered to a sample group of 30 elementary school teachers in the State of Oregon. Thirty minute time samples of the teaching behavior of each teacher subject were filmed using a video tape recorder. These films were rated by a panel of three competent, trained judges using the "Affective Rating Scale."

The initial guiding hypothesis for the study, stated in null terms was as follows: There will be no significant correlation between teachers' scores on the DCA and their behavioral rating on the "Affective Rating Scale." It was the expectation of the research, however, that a statistically significant correlation between the two variables would in fact be demonstrated and thus the null hypothesis would be rejected.

The data that resulted from the testing of the teachers with the DCA and the judges' ratings of their affective teaching behavior from the films was analyzed. Spearman rank-order correlations were made for the DCA and the three sub-scales of the "Affective Rating Scale." Spearman rank-order correlations were also made for each of the sub-categories of the DCA and the "Affective Rating Scale" sub-scales. Various measurements using a t-test formula

were made to study the DCA scores of beginning and experienced teachers. Inter-rater reliability was established using an analysis of variance formula to study the judges' ratings on the "Affective Rating Scale." A split half reliability analysis of the DCA was performed to establish its internal consistency.

The analysis of the data revealed the following:

15

- 1. There was no significant correlation between the composite DCA scores of teachers and their scores on the "Affective Rating Scale."
- 2. There was no significant correlation between the sub-category scores of the DCA and the scores for the sub-scales of the "Affective Rating Scale."
- 3. There was a significant difference between the DCA scores of beginning teachers and more experienced teachers. Beginning teachers scored higher on the DCA.
- 4. There was a significant difference in the "Affective Rating Scale" sub-scale scores for "use of expression" between beginning and more experienced teachers. However, there was no significant difference on the other two sub-scales of the "Affective Rating Scale" between these two groups of teachers.
- 5. There was no significant difference between the DCA scores of younger and older teachers.

Differential Cultural Awareness: Developing an Instrument for Testing the Cultural Awareness of Teachers

by

Dennis Lee Crawford

A THESIS

submitted to

Oregon State University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

June 1977

APPROVED:

Redacted for Privacy

Associate Professor of Education in charge of major

Redacted for Privacy

Dean of School of Education

Redacted for Privacy

Dean of Graduate School

Date thesis is presented June 7, 1976

Typed by Opal Grossnicklaus for Dennis Lee Crawford

FOR

Alta Lee

and

Denetta Pauline

and

Dalene Rae

and

David Joseph

and

Laurie Lee

ACKNOWLEDGEMENTS

I wish to express my appreciation to Professor Gerald Becker who served as the chairman and to the members of the committee, Professors Michael Colbert, Charles Cormack, William Fielder, and Charles Warnath.

The cooperation and assistance of the teachers who served as raters, Mr. Richard Arnold, Mr. Roland Davin, and Ms. Susan Taylor was essential and appreciated.

For the special help in arranging for the video taping of teachers,

I am particularly indebted to Ms. Anne Sondgroth.

I want to express my appreciation to my many colleagues at Oregon College of Education and in the Oregon Public Schools who helped and encouraged me with the study. I am particularly grateful to Professor Merle Kelley who helped me through the statistical labyrinth.

I thank my parents, David and Erma Crawford, who planted in me a thirst for knowledge and have inspired me with their love and faith throughout my education.

My special appreciation and love goes to my children, Denetta,

Dalene, David and Laurie who have made it all worthwhile.

I cannot fully express the deep respect and love I feel for my wife, Alta, who has suffered the frustrations of my work and has challenged me to continue to this end. TE AMO

TABLE OF CONTENTS

I.	INTRODUCTION	1
	Theoretical Context Statement of the Problem Research Hypothesis	4 8 8
II.	REVIEW OF RELATED LITERATURE	9
	Attitude Development and Change Attitude Measurement Reliability Validity Behavioral Observation Sampling	9 14 19 20 21 22
III.	METHOD	26
	Development of the Instrument Gathering and Analysis of Data	27 30
IV.	RESULTS	33
	Inter-rater Reliability Analysis Split Half Reliability Analysis of the DCA	29 36
v.	DISCUSSION	47
	Variables Affecting the Study Rating Scales Implications of the Study	47 52 55
VI.	SUMMARY AND CONCLUSIONS	58
	Findings Recommendations for Further Research	61 62
	BIBLIOGRAPHY	64
	APPENDIX I	69
	APPENDIX II	78

APPENDIX	III	83
APPENDIX	IV	84
APPENDIX	v	85
APPENDIX	VI	86
APPENDIX	VII	88
APPENDIX	VIII	90

LIST OF TABLES

<u>Table</u>		Page
1.	Analysis of Variance for the Affective Rating Scale (Sub-scale #1), Use of Expression	34
2.	Analysis of Variance for the Affective Rating Scale (Sub-scale #2), Reaction with Pupils	3 5
3.	Analysis of Variance for the Affective Rating Scale (Sub-scale #3), Style of Presentation	3 5
4.	Split Half Reliability Analysis of the DCA	36
5.	Spearman Rank-Order Correlations for the DCA and the Affective Rating Scale sub-scales	37
6.	Spearman Rank-Order Correlations of the DCA Sub- categories and the Affective Rating Scale sub-scale, Use of Expression	38
7.	Spearman Rank-Order Correlations of the DCA Sub- categories and the Affective Rating Scale sub-scale, Reactions with Pupils	39
8.	Spearman Rank-Order Correlations of the DCA Sub- categories and the Affective Rating Scale sub-scale, Style of Presentation	40
9.	Significance of the Difference between DCA Scores for Beginning and Experienced Teachers	41
10.	Significance of the Difference between Years of Teaching and Composite Affective Rating Scale Scores	42
11.	Significance of the Difference between Years of Teaching and Scores from the Affective Rating Scale sub-scale, Use of Expression	43
12.	Significance of the Difference between Years of Teaching and Scores from the Affective Rating Scale sub-scale, Reactions with Pupils	44

Table		Page
13.	Significance of the Difference between Years of Teaching and Scores from the Affective Rating Scale sub-scale, Style of Presentation	45
14.	Significance of the Difference between Teachers' Ages and DCA Scores	46

DIFFERENTIAL CULTURAL AWARENESS: DEVELOPING AN INSTRUMENT FOR TESTING THE CULTURAL AWARENESS OF TEACHERS

I. INTRODUCTION

One of the problems facing the American school system today comes in the form of a demand that the educational needs of the various cultural minorities be met more effectively. The essence of this demand is discussed in a recent article by Herman Hudson (1972), Vice Chancellor for Afro-American Affairs at Indiana University:

Black people today are uncompromisingly insisting on their right to participate in those decision-making processes that vitally affect their lives - all the way from national housing and education policies, to the control of local police, to semantics of self-designation (p. 297).

In the past decade, American Minority groups have strengthened the force of these demands considerably. They insist that the cultural heritage they represent be incorporated into the regular public school curriculum. As a result of this pressure, teachers are being asked to understand, and become aware of the culturally different children they serve (Jaramillo, 1973; Spindler, 1974).

All children bring to school the outlook of the culture to which they belong. If they are from a minority, they are usually confronted with a different cultural value system and mode of operation at the school. The language encountered there may be unfamiliar, their

dress may be different, and the procedures of the school may be in conflict with their training at home. Must the child from a culturally different background conform to the culture found in the classroom, or is it the teacher and the school which ought to make room for greater cultural diversity and learning?

What directions have been taken in changing the schools in the wake of growing dissatisfaction in American society? Roger Baty (1972) makes the following observation:

Many efforts during the late 1960's have been focused on improving the educational opportunities of 'disadvantaged' children. The assumption has been that the source of the difficulty lies in the cognitive deficit which disadvantaged children bring with them to the classroom. This deficit makes cumulative retardation and the ensuing gap between advantaged and disadvantaged inevitable. Much less attention has been given to the need for reeducation of adults in the educational system, especially teachers and administrators. Yet, without changes in attitude and behavior of those who are responsible for the education of the young it is difficult to see how any lasting changes can be brought about (p. 6).

How can educators meet the challenges of a culturally mixed school without forcing children into a particular cultural mold? Can the American public schools devise a means to meet the needs of a culturally pluralistic society?

Challenged by the needs of its minority population, the State of Oregon has established "... Rules for Certification of Teachers, Specialists and Administrators" (effective January 1, 1975) which require that candidates for both the elementary and secondary

certificates demonstrate teaching competency through experiences in "social and cultural foundation, including an understanding and appreciation of the role of minority groups in American society" (p. 25). In addition, the Oregon State Department of Education designed and instituted "The Inservice Cultural Awareness Training Program."

The aim of cultural awareness training was to develop in the teacher an awareness of how his or her own culture has led to certain attitudes about minorities and influences his or her performance as a teacher, and how the culture of students determines their attitudes and behavior. This training called for the development of an evaluative process.

Was the training program helping teachers to become more culturally aware? It was hoped that this research would provide data to confirm the value of such training. The original plan was to extend and expand the research of Baty, Swick, Spindler, and others through the development of an evaluative instrument for examining the cultural awareness level of teachers. It seemed likely that cultural awareness would correlate with affective teaching behavior.

However, as the investigation progressed, it became apparent that the whole construct of "cultural awareness" was much more complicated than first perceived. Furthermore, given the accepted framework of the research design, the measurement of attitudes

presented unique problems. The numerous and complex variables of teacher cultural awareness did not permit either adequate controls or effective reduction to a paper and pencil instrument. Such problems began to alter the direction of the research. Indeed, the outcome of this study raises serious questions about the validity of the root construct itself--cultural awareness.

Theoretical Context

"Culture" has many definitions as shown by Kroeber and Kluckhohn (1952). However, E. B. Tylor (1871) provides a classical definition of culture as:

. . . that complex whole which includes knowledge, belief, art, morals, law, custom and any other capabilities acquired by man as a member of society.

An underlying consistency in the various definitions given "culture" is that it is transmitted, that it is learned, and that it is shared (Talcott Parsons, 1952).

The public school is one of the most important points of contact between the larger society and its sub-groups. Baty (1971) discusses the influences of the school:

In the school, the values, beliefs, history and skills of the larger society are "offered" the young people of the subgroup or minority group by the teachers acting as the agents of the "encompassing society" (p. 9). The public school teacher becomes a major cultural transmitter and enculturation agent (Landes, 1965; Spindler, 1963).

When studying the relationship of education to culture, Stenhouse (1971) suggests that "... we ought to scrutinize not only the nature of culture, but also the process by which it is transmitted." In the U.S. this becomes quite complex since we are confronted with numerous cultural features. Cross-cutting the cultural differences that exist in this society are broad common socio-economic class groupings (lower class, middle class, etc.). It is impossible to disentangle the influences which mold and shape an individual within American society. Nonetheless, Americans have created "... specially organized groups whose task it is to transmit culture methodically and self-consciously." (Stenhouse, 1967:5). However, such transmission does not take place in a vacuum. There is continuous rearrangement of the societal and cultural priorities which must be incorporated into the enculturation of new members of the society.

Teachers are confronted with a variety of "different" cultural traditions in the society and often within the classroom. An understanding of the term "different" as applied to cultures is offered by Harry Rivlen (Stend et al., 1973: vii), "Different means different; it does not mean better or worse than." Nor does it mean that all "cultural differences" can be equated on the same scale of values.

There is a growing awareness in the public schools concerning

the need to offer information about the cultures of minority groups to all students. What this entails is outlined by Von Maltitz (1975) in a two-fold proposal:

... to instil in children from those minority groups, who have often been subject to prejudice, a pride in their own language and culture and to promote among persons from other segments of the population an understanding and a respect for Americans who may be different from themselves (p. 1).

"Cultural awareness" conveys the idea that not only is one aware of the processes of cultural transmission but it also implies, as Charles Cormack (1975) states, ". . . that I not merely recognize the Other but that I am responsible for the Other." This responsibility and caring leads to enhancing the freedom of the minority or culturally different person.

Literally the term "cultural awareness" implies neither a positive nor a negative response, but simply comprehension or apperceptive knowledge. However, as the term is currently being applied by the U.S. Office of Education through the funding of "Cultural Awareness Training Centers" for teachers, it has taken on a more specific meaning and direction. This direction is toward positive action to enhance the freedom and range of choice for the culturally different. Following this vein, Cormack (1975) states:

... there can then be degrees of cultural awareness which, in a positive sense, open up the possibilities for mutuality of action that is adaptive, i.e., where possibilities for change exist, which result in the organization (or integration) of cultural difference to universally beneficial ends.

Thus far very little research has been devoted to the evaluation of cultural awareness training. Addressing the problem, Swick (1972) urges ". . . that longitudinal efforts to improve the racial attitudes of teachers is needed if deeply held belief systems are to be changed" (p. 6). In addition, Baty (1972) observes:

There are few studies of attempts to increase the supply of effective teachers by reeducation of those teachers already on the job and improving their attitudes toward minority and/or poor people. There is a need for such studies, which would yield comparative data on changes in teacher attitudes with and without supplemental training programs.

Further emphasis on the need for evaluative investigations comes from Paul Devore (1971) in his study "Variables Affecting Inservice Teacher Education." Referring to the reeducation of teachers, Devore concludes, "There is vast literature on 'how to do it' but little research and evaluation to note whether 'it did do it'...."

This study, then, was predicated on the belief that it is useless to train teachers in cultural awareness unless this training can be shown effective in developing a more responsive, sensitive teacher. It was believed that research was essential in order to, one, determine if teacher cultural awareness is a measurable attitude and, two, whether or not this attitude affects teaching behavior.

Statement of the Problem

The original focus of this study was the development of an instrument to measure cultural awareness of teachers. "Differential Cultural Awareness," abbreviated as DCA hereafter, is the title chosen for the measurement instrument developed through this research. Furthermore, it was expected that scores on this instrument would be significantly and positively related to affective teaching behavior in the classroom. However, it will be shown later in the study that the logic of this position has serious flaws. Hence, in the closing chapter the "problem" under investigation shifts from demonstrating an expected relationship to one of logically critiquing or examining both the conventional wisdom about and quantitative measurement of cultural attitudes and teacher affective behavior.

Research Hypothesis

The initial guiding hypothesis for the study, stated in null terms was as follows: There will be no significant correlation between teachers' scores on the DCA and their behavioral rating on the "Affective Rating Scale." It was the expectation of the research, however, that a statistically significant correlation between the two variables would in fact be demonstrated and thus the null hypothesis would be rejected.

1

II. REVIEW OF THE RELATED LITERATURE

The attitudes of individuals and groups have become a source of interest and concern to government agencies, educators, social scientists and many others. As a result of this interest, there has developed a wide range of literature, research data and materials covering attitude development, attitude change and attitude measurement. This chapter is devoted to examining the literature in the above mentioned areas, especially as it relates to racial attitudes and cultural awareness among teachers and students. Particular attention is given to procedural techniques in test development.

Attitude Development and Change

The way we view other people, their ways, beliefs, values, etc., is dependent upon our own culturally generated attitudes.

Sherif (1967) describes attitudes, as that which a person learns as a member of a family, a group, and a society which allows him or her to interact rather consistently and survive within the world. A person's attitudes stem from choices or decisions which he or she has made consciously and unconsciously beginning with birth and continuing throughout life.

A widely used and accepted definition of attitude was proposed by Gordon Allport (1935): 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.

The variety of interpretations of the meaning of attitude have substantial areas of agreement. Summers (1970) outlines the following areas in which there is general consensus.

First, an attitude is a predisposition to respond. It is characterized by a readiness to behave. Allport's (1935) definition of attitude calls it a "mental and neural state of readiness." Oppenheim (1966:105) calls this readiness, "a tendency to act or react in a certain manner when confronted with certain stimuli."

Second, "attitude is persistent over time" (Summers, 1971). It is not immutable but attitudes are resistant to change, especially strong attitudes. This persistence of attitudes results in some consistency of behavior.

Consistency of behavior is the third characteristic of attitudes in which there is fairly common agreement. In discussing consistency related to attitudes Muzafer Sherif (1967) suggests it is the process of learning that establishes attitudes which are both consistent and characteristic, instead of transitory and haphazard.

This leads to the fourth factor which is common to the many interpretations of attitudes, their directional quality. This directional quality of attitudes is expressed by Sherif (1967):

We are talking about the fact that he is no longer neutral in sizing up the world around him. That his behavior toward other persons, groups, institutions and nations takes on a consistent and characteristic pattern as he becomes socialized. (p. 2).

Attitudes function to help people understand the world (Triandis, 1971).

This process involves organizing and simplifying the environmental input into a consistent framework.

It is well known that children begin early to absorb the culture and attitudes to which they are exposed. Parents, teachers and peers are primary models in the enculturation process (Goodman, 1970; Coles, 1964; Sherif, 1967).

Because so much emphasis is placed upon the first six years of life, parents are given credit and blame for the broad social attitudes of their children (Sherif, 1967). However it must be remembered that children are selective in what attitudes and behaviors they choose from among the social-cultural forces impinging on them.

As time passes and their world expands, children find others from whom they learn and assimilate attitudes. A child is subject to the values and attitudes of the teacher, the school, and the dominant community, as well as his or her own family and culture. These make their imprint upon the child's self concept and influence his or her response to others and the society.

The 180 days a year that children spend the majority of their waking hours at school with teachers and peers serve to shape their

attitudes considerably. Spindler (1963) speaks of the teacher as a "cultural transmitter," and education as a process of "cultural transmission." He states:

. . . It becomes clear that even the most fair-minded teachers are highly selective of the values they communicate to students and are equally selective with respect to what values they screen out from what students might potentially communicate to them. It also seems clear on the basis of such approaches that teachers, in fact whole educational programs, frequently communicate assumptions and outlooks about human relations that are not in agreement with their declared goals (p. 43).

The role of the teacher in the development of attitudes is contradictory and varied when viewed as a mediator of culture. Teachers are involved in transmitting values implicitly and explicitly, consciously or unconsciously (Landes, 1965; Chilcott, 1969; Spindler, 1974).

Jules Henry (1960) suggests:

A focus on values in the study of educational processes has implications not only for understanding polyphasic learning, since education, the fundamental organizing process, occurs always in a context of values, and teachers are usually teaching values by implication, regardless of the immediate subject matter (p. 273).

Recent research seems to indicate that teachers' attitudes can be directly related to a child's success or failure in school (Baty, 1972; Parmee, 1968; Reece, 1974). "Failure to recognize the overwhelming influence of culture on personality and behavior" is cited by Thomas Carter (1969) as a major teacher deficiency. Teacher attitudes may reflect prejudices as seen in the definition given by Thorndike (1969:382), "Attitudes relate to tendencies to favor or

reject particular groups of individuals, sets of ideas, or social institutions."

As a result of what appear to be deficiencies in cultural understanding and awareness among teachers, Title IV¹ desegregation programs and centers in many state and local districts have been established. At one such center in New Mexico, more than 2000 school personnel have received training in cultural awareness (Jaramillo, 1973).

Teachers are urged by Spindler (1963) to learn to analyze the cultural influences they bring to the classroom:

Cultural therapy is one direct measure we can take in our teacher-training programs to help reduce the self-defeating effect of cultural transmission in American schools (p. 171).

In describing a teacher reeducation project conducted by Claremont Graduate School, Thomas Carter (1969) reports that teachers were:

. . . encouraged or 'forced' to understand culture's influence on their own individual perceptions, attitudes, and behavior - to see culture as manifest within themselves (p. 11).

The findings of a study carried out by Swick and Lamb (1972) indicate that cultural awareness type training for teachers ". . . can provide student (teacher education students) gains in information about minority groups and initiate racial attitude change."

A change in attitude may take place in several ways. Triandis (1971) suggests that attitude change can occur by changing the cognitive component with new information, or the behavioral component

¹Title IV, Civil Rights Act, 1964.

through norm change, or through forced change. Changes in the cognitive area will tend to be reflected in the affective and behavioral areas (Rokeach, 1968).

A call for studies and tools to assess attitude change among teachers is made by Baty (1971): "There is a need for such studies, which would yield comparative data on changes in teacher attitude with and without supplemental training programs" (p. 8).

Attitude Measurement

When beginning to examine ways of measuring attitude change, Rokeach (1968) cautions:

... the classical paradigm employed in experimental studies of the opinion change - pretest, treatment, post-test - is not capable of telling us whether an expressed opinion change indicates an attitude change; it can only tell us whether an expression of opinion has or has not changed as a result of a particular experimental treatment (p. 140).

Attitude scales can only measure the attitudes which are expressed.

The subject may consciously or unconsciously hide his or her true attitudes because of social pressures. Attitude instruments attempt to minimize the interference of such factors (Summers, 1970).

The first and most important task in measurement of attitudes is to specify the problem or variable which is to be studied and limit the measurement to that (Thurstone, 1928; Anastasi, 1961; Kerlinger, 1973).

Attitude scales are constructed to provide a quantitative measure of the individual's responses to the variable in question which gives his or her relative position along a unidimensional continuum

(Anastasi, 1961). Weschler (1950) in describing attitude measures says their usefulness correlates with their ability to force the subjects to reveal attitude through strong enough stimulus.

Oppenheim (1966) discusses at length what he calls "pilot work" which must precede the development of an attitude instrument. This involves exploratory work to develop a feel for the problem through interviews with key informants, and gathering of information about the subject of inquiry. Finally it involves the construction of a pretest or trial questionnaire with approximately 100 to 150 statements.

Thurstone (1928) suggests that in developing a list of statements for a questionnaire some important criteria are: (1) that statements be as brief as possible, (2) that they allow the subject to either agree or disagree, and (3) that every statement be such that its acceptance or rejection indicates something regarding the subject's attitude.

Many types of attitude scales have been developed over the past

40 years; some of which are more reliable than others, and some of
which are more valid for certain purposes (Thurstone, 1929; Likert,

1934: Edwards, 1957).

In the measurements of social attitude Thurstone, Likert and Guttman scales have been widely used. These scales have at least two characteristics in common:

1. They all represent the individual's attitude toward an object

by a single preference score or average "most acceptable"
position on a continuum of positions ranging from highly favorable to highly unfavorable.

2. In every case, the individual is fully aware that his attitude on the issue in question is being measured (Sherif, 1967:141).

Thurstone equal-appearing interval scales assigns each item a scale value which indicates the strength of attitude of agreement response to the item. Different scale values are assigned each item in an ordered set with equal intervals between items. Scores for each item are registered on a continuum from positive to negative, with the median score considered neutral (Kerlinger, 1973).

Assignment of scale values for the "equal appearing intervals" method is made by a panel of judges. Judges are asked to place the attitude statements in 11 piles ordered according to their degree of favorableness to the attitude object. The piles of statement items are to be considered equally distant from each other, so that intervals between the piles are equal appearing intervals. A distribution of judgments from the judges is thus obtained for each of the statements (Triandis, 1971).

Edwards (1957) levels several criticisms at the Thurstone type scales. First, he indicates that the distribution of attitudes is not normal over the entire continuum. In some areas the results will be skewed in one direction because the majority of subjects will

respond in like manner on a particular topic area. A second difficulty is that neutral statements do not discriminate and are irrelevant and contribute greatly to error, "and the coefficient of reproducibility will be decreased accordingly." The third difficulty which Edwards cites is that not all statements with the same scale value are equally discriminating. Statements that fall within the same scale interval and with comparable values still differ considerably in their power to differentiate between high and low criterion groups. All of these factors are seen as interfering with reproducibility of these type attitude scales.

Triandis (1971) points out that the basic assumption of the

Thurstone methods are that one sample of judges give the same values
for statement items as another sample of judges. ". . . As long
as judges are not extremists on the particular attitude continuum,
this assumption is generally safe" (p. 41). In addition to this problem, in the use of judges, Triandis also says that it is uncertain how
much agreement is required by the judges to assign the appropriate
location for a test item. Finally, it is argued that the judges and
subjects differ in task assignments: The judges assign the location
of the item while the subjects indicate agreement with the item.
"That sort of shift in the nature of the task may well introduce
measurement errors" (Triandis, 1971:45).

The Likert summated rating scales are similar to the Thurstone.

Items related to the attitude to be measured are formulated, but instead of being sorted by judges, the items are administered to respondents similar to the target group. Responses to each item are on a five point scale: "strongly approve," "approve," "undecided," "disapprove," or "strongly disapprove." The sum of the item scores is the total score for an individual (Sidowski, 1966). Items are considered of approximately equal "attitude value" and intensity of attitude expression is allowed. As Kerlinger (1973) indicates, there are both advantages and disadvantages to this method. The advantage is that variance is allowed and accounted for. The disadvantage Kerlinger (1973) mentions is:

Individuals have differential tendencies to use certain types of responses: extreme responses, neutral responses, agree responses, disagree responses. This response variance confounds the attitude (and personality trait) variance (p. 496).

The Guttman or the cumulative scale is made up of a relatively small set of homogeneous items that are supposed to be unidimensional. They are to measure one variable only. A cumulative relation is established between items and the total scores of individuals (Kerlinger, 1973). Given a set of attitude statements in the Guttman scale, an individual who has a higher rank or score than another individual, must score as high or higher than the other person on each one of his answers. The Guttman scales are criticized for being difficult to construct (Kerlinger, 1973; Triandis, 1971;

Sidowski, 1966).

In conclusion, Oppenheim (1966) points out that a great deal depends on the puprose for which the attitude scale is built. It is easier to build a scale that is descriptive and divides a sample into several groups with regard to a given attitude than to predict people's actions. "Failure to predict a particular action does not constitute proof that the attitude scale was invalid" (p. 153).

The conclusion reached by Kerlinger (1973) in discussing the Thurstone, Likert, and Guttman methods is:

Of the three types of scales, the summated rating scale seems to be the most useful in behavioral research. It is easier to develop, and as indicated, yields about the same results as the more laboriously constructed, equalappearing interval scale (p. 499).

It is pointed out by Thorndike (1969: p. 415) that if we recognize and are willing to work within the limitation that an attitude scale represents verbalized attitudes that an individual may be willing to express and doesn't have to back up, then an "attitude scale appears to be a useful research tool or tool for experimental evaluation of educational objectives lying outside the domain of knowledge and skills."

Reliability

Reliability is the accuracy or precision of a measuring instrument (Kerlinger, 1973). There are two ways of expressing reliability or precision of a set of measurements or the variation within the set. One approach involves the amount of variation within a set of repeated measurements of one individual. There is a frequency distribution that occurs in repeated testing, and we may obtain an average value which is an approximation of the true score. There is a standard deviation describing the spread of the scores obtained. In educational and psychological measurement we seldom obtain a whole series of scores for each individual. Having a pair of measurements for each individual it is possible to estimate the scattering of scores that would occur if we had been able to make the measurement again and again (Mehrens, 1973; Thorndike, 1969).

Secondly, reliability implies that the individual stays in about the same place in his group after repeated measurements. There are three common ways of obtaining this kind of data: test-retest, equivalent tests, and split half tests (Thorndike, 1969; Kerlinger, 1973; and Mehrens, 1973). (See Appendix VI for a further discussion of Reliability.)

Validity

The most important question in the development of any test concerns its validity. Validity provides a check on whether we are measuring what we think we are measuring.

Three types of validity have been classified by the joint committee of the American Psychological Association, the American

Educational Research Association, and the National Council on Measurements Used in Education. ² These three types include: content validity, criterion-related validity, and construct validity. For a further discussion of validity, see Appendix VII.

Behavioral Observation

Basically there are two modes of observation as outlined by Kerlinger (1973:537). "We can watch people do and say things and we can ask people about their own actions and the behavior of others." The goal of behavioral observations is either verbal or quantitative behavior descriptions (Ryans, 1960).

The observer becomes one of the main problems in the assessment process. The skill of the observer is crucial for the basic weakness is that incorrect inferences are often made from observations. Another problem associated with observer objectivity has to do with bias. When rating affective teaching, bias may invalidate the observations being made by teachers who are used as observers.

Observers for instance may be biased against public schools in favor of parochial schools. On the other hand if we select non-teachers to rate affective teaching we may eliminate biased observations but the

²Standards for Educational and Psychological Tests and Manuals, Washington, D. C. American Psycyological Association, 1966.

observations may be inadequate. Kerlinger (1973:538) in discussing this problem states, "Observation of human behavior requires competent knowledge of that behavior and even of the meaning of behavior."

A further problem discussed by Kerlinger (1973) regarding observations is that the process of observing will cause those being observed to act differently. This is thought not to be a significant problem if the observer is careful to be unobtrusive as possible (Kerlinger, 1973; Thorndike, 1969).

A major consideration in the observation process is the selection of an appropriate criterion measure or rating scale. Ryans (1960) states:

It is important that a criterion measure be both comprehensive, or inclusive, and also that it not measure behaviors extraneous to the criterion dimensions under consideration (p. 37).

There are a number of types of observational rating scales.

Those most appropriately considered for the research under discussion were: Category rating scales, numerical rating scales, and graphic rating scales. While similar they differ mainly in details (Kerlinger, 1973). (See Appendix VIII for a discussion of rating scales.)

Sampling

Sampling behavior is a way of obtaining observations. When proposing to observe classroom behavior of teachers, it must be

decided how such behavior is to be observed. There are two ways of sampling such behavior: event sampling and time sampling (Kerlinger, 1973).

Event sampling entails selection of particular events or behavior for observation. Examples of such are specific classroom interactions between teacher and pupils, fights, games, etc., in which the whole drama of any specific event is observed. Event sampling has several advantages. First, the events are true-life situations. Second, there is a continuity to the behavior from beginning to end. It is not a cut or piece of behavior as one might obtain in a time sample (a thirty minute period). Third, event sampling captures what may be infrequent or rare yet important behavior (Kerlinger, 1973).

Obviously the problem with event sampling is the enormous amount of time that it may take to observe a particular form of behavior.

Time sampling involves the selection of units of behavior for observation at different points in time. The observation of samples of behavior may be selected systematically or randomly. "Time samples," according to Kerlinger, "have the important advantage of increasing the probability of obtaining representative samples of behavior" (1973:546). Representative samples of usual or ordinary behavior are more apt to be obtained when several observations are made over a given period of time.

The disadvantage with the time sample is that infrequent behavior may be missed. Nonetheless both Ryans (1960) and Kerlinger (1973) suggest this method of sampling is probably the most usable approach in making behavioral observations.

There are many types of numerical rating scales for observing time samples of teacher-pupil behavior, two examples of which are Ryans' (1960:106) "Classroom Observation Record," and Garrison and Kersh's (1969) "Affective Rating Scale." The Ryans scale consists of four pupil behavior categories and eighteen teacher behavior categories. These are rated along a seven-point continuum. The "Classroom Observation Record" is rather long and requires thorough and extensive training of observers (Ryans, 1960).

The "Affective Rating Scale" (see Appendix III) developed by Garrison and Kersh (1969) consists of three variables which are rated on a ten-point scale. These variables are: (1) use of expression, (2) reactions with pupils, and (3) style of presentation. This rating scale also requires that the observers be well trained (Garrison and Kersh, 1969).

In making the observations and rating teacher-pupil behavior using a scale such as one of the above, Ryans (1960) cautions:

... once the observer has undergone a program of training it cannot be assumed that his observational procedures will remain uniform over a period of time without occasional check-ups and retraining: rather, it is necessary to

reinstitute training at regular intervals to ensure that there has been no shift in the definition and perception of the behavior being observed (p. 73).

III. METHOD

The idea for an evaluative instrument that would test the attitudes of teachers related to cultural awareness began to emerge in the spring of 1972 when this researcher attended an in-service training program in cultural awareness for teachers. The training was being conducted by members of the Oregon State Department of Education under the auspices of the Compensatory Education Division for a local Salem, Oregon, school. Help was asked of this researcher to develop an evaluation process for the cultural awareness training program.

After attending the in-service training in Salem, field interviews were conducted in Pendleton, Oregon, with teachers from the local schools who had recently taken cultural awareness training.

In these interviews, some teachers contended that the training had helped them change their attitudes toward the culturally different.

Others felt the training had made no difference and may have even reinforced some of their prejudices. A means of measuring cultural awareness levels of teachers seemed to be needed.

A search was undertaken to see if an adequate instrument for testing the cultural awareness level of teachers already existed which would meet the need for the Oregon training program. In the research conducted by Roger M. Baty (1971) in Northern California,

a cultural awareness attitude questionnaire was used. This questionnaire was designed for use with teachers who were currently teaching
in classrooms with large numbers of Mexican-American students.

Baty's instrument was not suitable for the present need in Oregon
since the questions were specific to Mexican-American culture and
designed for only those engaged currently in teaching MexicanAmerican students.

In Oregon there is a scattering of minorities in the schools with few concentrations of any one group. An instrument was required that was designed for minority cultures in general rather than being applicable to only one cultural group. The instrument needed to be usable with teachers who had not taught minority students as well as those who had. The DCA instrument was developed in an attempt to meet this need.

Development of the Instrument

A Likert-type scale was selected as the model for the DCA instrument. Such a scale provides an adequate range of attitude responses, has the advantage of simplicity of design and provides quantitative data (Likert, 1934).

The Likert scale is composed of an item pool. Items are designed to elicit either negative or positive responses with as few neutral items as possible. The scale has a range of five scores

from strongly agree to strongly disagree. The middle score is given the category of "uncertain." Items are scored from 1-5 or 5-1 depending upon whether they are stated positively or negatively.

Oppenheim reports of Likert scales:

... reliability of Likert scales tends to be good, and partly because of the greater range of answers permitted to respondents, is often higher than that of corresponding Thurstone scales (p. 140).

Poppleton and Pilkington (1964), and Tittle and Hill (1967), also attribute high reliability with Likert scales in their studies.

Kerlinger (1973) referring to Thurstone and Likert scales states

''. . . both types of scales yield about the same results as far as reliability and the placing of individuals in attitudinal rank orders are concerned' (p. 499).

For the purposes of this study, a pretest was first designed with 160 questions (see Appendix I). Questions for the pretest instrument were formulated by this researcher and three competent people in the field of minority education in Oregon. Each of these people had many years of experience and training in teaching minority students. Criteria for the selection of the questions for inclusion in the instrument were similar to that suggested by Kerlinger (1973):

- 1. Does the question relate to the research problem?
- 2. Is the question able to be applied to minorities generally rather than to a specific cultural group?

3. Is the question clear and unambiguous?

The questions were constructed in order to elicit definite attitude responses, whether positive or negative, with minimal room for ambivalence. The questions were then sorted into six broad subcategories related to the research problem:

- 1. minority language usage
- 2. minority students
- 3. minority parents
- 4. teachers relationship to minorities
- 5. schools relationship to minorities
- 6. general questions related to minorities in America

 Questions which did not fit one of these categories were eliminated.

 It was recognized that in the pretest as constructed the number of items was far too large for the final instrument.

The pretest model was given to 106 graduate and upper division level students at Oregon College of Education. Many of these were teachers or preparing to be teachers. They were asked to answer the questionnaire and write comments regarding the wording, appropriateness, fairness, etc. of the questions. In addition, several colleagues as well as the training team from the State Department of Education reviewed the model and provided helpful suggestions.

It was determined from the trial testing using the pretest model that some of the questions would discriminate attitude responses, and that others were not significant. An internal consistency method of item analysis was employed. Those items which were found to discriminate in the desired direction were retained for the DCA instrument. Fifty-eight items were thus retained for the final instrument. Some questions underwent slight modifications to achieve a balance of 29 negative and 29 positive item statements. Questions were assigned a number from a random numbers table. The questions were then selected at random from the table to be placed in sequential order within the final DCA instrument.

Each question on the DCA instrument received a score from one to five with the higher score indicating an attitude that was considered more "culturally aware." Each subject's total score was obtained by adding the scores for the individual questions.

Gathering and Analysis of Data

During the construction of the DCA pretest instrument, three competent judges were used in the selection of the items for use in the instrument. These judges were selected on the basis of their experience and training in the area of minority education. Each of the judges had five or more years teaching experience and possessed at least a masters degree. Each judge represented a different minority background which included Native American, Black, and Mexican-American. All the judges were employed by the Oregon State Department of Education. They were instructed to select items which they

judged to be specifically significant in determining cultural attitudes of teachers toward minority group people.

Criterion-related validity (Ryans, 1960; Kerlinger, 1973) was to be assessed by correlating the criterion of teacher behavior in the classroom with the teachers' scores on the DCA. Thirty teachers were selected from a pool of teachers who were administered the DCA instrument during 1975 and 1976. All of those selected were regular classroom instructors in Oregon schools. Thirty minutes of video tape were made of each teacher's teaching activity. Criteria for the selection of the 30 minutes that were filmed were as follows:

- 1. representative of a fair sample of teaching behavior
- 2. involve teacher-pupil interaction in groups
- show the teacher engaged in several different aspects of teaching.

Selection of the teachers to be observed included criteria for controlling three specific variables which were considered to be essential. To establish a homogenous student population with whom the teachers would be observed to interact, only elementary teachers were selected. Second, to avoid contamination by teachers having varying degrees of experience teaching culturally different students, only teachers with no experience of this type were selected. Third, teachers who were engaged in their first year of teaching were eliminated from consideration because of their lack of teaching

experience.

The "Affective Rating Scale" developed by Garrison and Kersh (1969) was selected for use in rating the filmed classroom teaching behavior of the teachers who were video taped. This rating scale was selected on the basis of its applicability to this research, and its high reliability in identifying affective teaching among teachers. An added factor in the selection of this rating scale was its simplicity of design and ease in use (see Appendix III).

Inadequacies in the validation of the "Affective Rating Scale" are similar to other such rating scales and discussed in the chapter on Discussion.

Observation and assessment of the teachers' behavior as recorded on video tape was made independently by three graduate students at Oregon College of Education who were experienced, practicing teachers. The observers were trained initially in the use of the "Affective Rating Scale" by Dr. Jesse H. Garrison, the developer of the instrument. They were retrained monthly by this researcher throughout the three months in which they viewed the tapes and made their ratings. Inter-rater reliability coefficients are presented in Tables 1, 2, and 3. An analysis of variance was employed to test for inter-rater reliability.

By analysis of variance (Winer, 1962:131).

IV. RESULTS

The present study provides evidence that despite careful instrumentation and testing the investigation of cultural attitudes of teachers is fraught with considerable difficulty and innumerable variables. As evidenced by the following results of this study, these attitudes remain illusive and continue to challenge scientific analysis. These research findings will form a small portion of the data that is needed to better understand the nature of cultural awareness and its relationship to classroom behavior of teachers. The results of this study represent the culmination of four years of research and the beginning of the study of cultural awareness among teachers in the State of Oregon.

Inter-rater Reliability Analysis

Inter-rater reliability was established separately for each of the three sub-scales on the "Affective Rating Scale" using an analysis of variance (Winer, 1962: p. 128). Three judges were used to rate video taped time samples of behavior which had been filmed of the 30 teacher subjects in the study. Thirty minute video tapes of each subject's behavior were rated separately by each of the three judges for each sub-scale on the "Affective Rating Scale" (see Appendix IV).

Analysis of variance statistical procedures were used to establish the reliability of the judges' ratings. The results of this analysis are presented in Tables 1, 2, and 3. The estimate of reliability of the raters on the first sub-scale of the "Affective Rating Scale" is (r=. 86). The estimate of reliability on the second sub-scale is (r=. 93). The estimate of reliability on the third sub-scale is (r=. 89). These are adequate correlations to substantiate the reliability of observations between judges.

Table 1. Analysis of Variance for the Affective Rating Scale (Subscale #1), "Use of Expression"

Source of variation		SS		df ————	MS
Between people		286. 76		29	9.89
Within people		81.34		60	1.35
Between judges	1.06		2		. 53
Residual	80. 28		58		1. 38
TOTAL		368. 10		89	

$$r = 1 - \frac{MS \text{ within people}}{MS \text{ between people}} = 1 - \frac{1.35}{9.89} = .86$$

Table 2. Analysis of Variance for the Affective Rating Scale (Sub-Scale #2), "Reactions with Pupils"

Source of variation		SS	-	df	MS
Between people		308.33		29	10. 63
Within people		44.67		60	. 74
Between judges	1. 76		2		.88
Residual	<u>42.</u> 91		58		. 74
TOTAL		353.00		89	

$$r = 1 - \frac{MS \text{ within people}}{MS \text{ between people}} = 1 - \frac{.74}{10.63} = .93$$

Table 3. Analysis of Variance for the Affective Rating Scale (Sub-Scale #3), "Style of Presentation"

Between people 316. 22 29 Within people 71. 34 60 Between judges 2. 89 2			SS		Source of Variation
Within poople	10.90	29	316. 22		Between people
Between judges 2.89 2	1.19	60	71.34		Within people
	1.44		2	2.89	Between judges
Residual 68.05 58	1. 17		58	68.05	Residual
TOTAL 386. 56 89		89	386. 56		TOTAL

$$r = 1 - \frac{MS \text{ within people}}{MS \text{ between people}} = 1 - \frac{1.19}{10.90} = .89$$

Split Half Reliability Analysis of the DCA

A Pearson product-moment correlation coefficient for an oddeven pairing of the DCA item scores was calculated. The value
yielded was . 69. In order to correct the value of the Pearson
product-moment correlation coefficient and make it applicable to
the whole DCA instrument, an additional computation was required
(Bruning and Kintz, 1968). This computation yielded a split half
reliability coefficient applicable to the whole DCA test of . 81 (see
Table 4). Bruning and Kintz (1968) have stated that to achieve reliability in measuring the characteristic which a test has been designed to measure a reliability value of (. 70 or higher) must be
achieved.

Table 4. Split Half Reliability Analysis of the DCA

	Odd # Items	Even # Item
Total # of Items	29	29
Total Sum of Scores	3175	3090
Value of the Pearson product-moment correlation	r =	. 69
Corrected Value of r applicable for the whole DCA test		.81

Spearman Rank-Order Correlations for the DCA and the "Affective Rating Scale" Sub-scales

As part of a test of validity, the Spearman's rho formula was used to compute rank-order correlations between the DCA scores of teachers and each of the three sub-scale ratings of teacher behavior from the "Affective Rating Scale." The results for the sub-scale "use of expression" and the DCA scores was (rho = .12). A critical-ratio z-test (Bruning and Kintz, 1968: p. 158) was used to test the significance of rho. Rho was not significant (z=.65). For the sub-scale "reactions with pupils" and the DCA scores (rho = .08). Rho was not significant (z=.42). For the sub-scale "style of presentation" and the DCA scores (rho = .10). Rho was not significant (z=.53). These results are summarized in Table 5.

Table 5. Spearman Rank-Order Correlations for the DCA and the "Affective Rating Scale" Sub Scales

rho = .12	z = .65, N.S.
	·
rho = .08	z = .42, N.S.
rho = .10	z = .53, N. S.
	rho = .10

Spearman Rank-Order Correlations for the DCA Sub-categories and the "Affective Rating Scale" Sub-scales

The Spearman's rho formula was used to compute rank-order correlations between the six sub-categories of the DCA and the three sub-scales of the "Affective Rating Scale." The results of these 18 rank-order correlations are summarized in Tables 6, 7, and 8. No significant correlations were found to exist between the sub-categories of the DCA and the sub-scales of the "Affective Rating Scale."

Table 6. Spearman Rank-Order Correlations of the DCA Subcategories and the "Affective Rating Scale" Sub-scale, Use of Expression

	Use of Exp	oression
DCA Sub-categories	rho values	Critical-ratio
minority language		
usage	rho =07	z = .37
minority students	rho =10	z = .54
minority parents	rho = .14	z = .78
teachers relationship		
to minorities	rho = .31	z = 1.68
schools relationship	tos,	
to minorities	rho = .15	z = .79
general items related		
to minorities in America	rho = .005	z = .03

Table 7. Spearman Rank-Order Correlations of the DCA Subcategories and the "Affective Rating Scale" Sub-scale, Reactions with Pupils

	Reactions with Pupils		
DCA Sub-categories	rho values	Critical-ratio	
minority language			
usage	rho =03	z = .16	
minority students	rho =25	z = 1.33	
minority parents	rho =04	z = .21	
teachers relationship			
to minorities	rho = .13	z = .70	
schools relationship			
to minorities	rho = .06	z = .30	
general items related			
to minorities in America	rho =18	z = .97	

Table 8. Spearman Rank-Order Correlations of the DCA Subcategories and the "Affective Rating Scale" Sub-scale, Style of Presentation

	Style of Presentation		
DCA Sub-categories	rho values	Critical-ration z values	
minority language	rho = .06	z = .34	
45 4 50	1110 .00	2 .31	
minority students	rho = ~. 22	z = 1.20	
minority parents	rho = .04	z = .24	
teachers relationship			
to minorities	rho = .29	z = 1.58	
schools relationship			
to minorities	rho = .21	z = 1.14	
general items related			
to minorities in America	rho =04	z = .20	

Significance of the Difference between DCA Scores for Beginning and Experienced Teachers

A t-test formula was used to compute the performance difference on the DCA between teachers who have taught from 2-8 years and teachers who have taught nine or more years. The results of these computations are shown in Table 9. A significant difference at the .05 level was found to exist between the DCA scores of the two groups. Teachers who have taught from 2-8 years had higher DCA scores than those who have taught nine or more years.

Table 9. Significance of the Difference between DCA Scores for Beginning and Experienced Teachers

	t-Test	
Teachers' years of teaching	2-8 years	9 or more years
Sample size	n = 11	n = 15
Composite DCA raw scores	2405	3072
Squared DCA scores	528, 271	632, 534
Degrees of freedom	df = 2	4
t value for the two groups	t = 2.	33
t value required (. 05 level)	t = 2.	064

Significance of the Difference between Years of Teaching and Composite "Affective Rating Scale" Scores

A t-test formula was used to compute the performance difference on the "Affective Rating Scale" between teachers who have taught from 2-8 years and teachers who have taught nine or more years. The results of these computations are shown in Table 10.

No significant difference was found to exist between the ratings of the two groups on the "Affective Rating Scale."

Table 10. Significance of the Difference between Years of Teaching and Composite "Affective Rating Scale" Scores

	t-Test	
Teachers' years of teaching	2-8 years	9 or more years
Sample size	n = 11	n = 15
Composite ARS raw scores	407	718
Squared ARS scores	16, 523	40,120
Degrees of freedom	df = 2	24
t value for the two groups	t = 1.	58
t value required (. 05 level)	t = 2.	064

Significance of the Difference between Years of Teaching and Scores from the "Affective Rating Scale" Sub-scale, Use of Expression

A t-test formula was used to compute the performance difference on the "Affective Rating Scale" sub-scale, use of expression, between teachers who have taught from 2-8 years and teachers who have taught nine or more years. The results of these computations are shown in Table 11. A significant difference at the .05 level was found to exist between the two groups. Teachers who have taught from 2-8 years were rated more affective in the use of expression than teachers who have taught nine or more years.

Table 11. Significance of the Difference between Years of Teaching and Scores from the "Affective Rating Scale" Sub-scale, Use of Expression

	t-Test	
Teachers' years of teaching	2-8 years	9 or more years
Sample size	n = 11	n = 15
Raw scores for Use of Expression	135	258
Squared scores	1,859	5,024
Degrees of freedom	df =	24
t value for the two groups	t = 2	. 16
t value required (. 05 level)	t = 2	. 064

Significance of the Difference between Years of Teaching and Scores from the "Affective Rating Scale" Sub-scale, Reactions with Pupils

A t-test formula was used to compute the performance difference on the "Affective Rating Scale" sub-scale, reactions with pupils, between teachers who have taught from 2-8 years and teachers who have taught nine or more years. The results of these computations are shown in Table 12. No significant difference was found to exist between the ratings of the two groups for "reactions with pupils."

Table 12. Significance of the Difference between Years of Teaching and Scores from the "Affective Rating Scale" Sub-scale, "Reactions with Pupils"

	<u>t-Test</u>	
Teachers' years of teaching	2-8 years	9 or more years
Sample size	n = 11	n = 15
Raw scores for reactions with pupils	128	227
Squared scores	1,656	4,097
Degrees of freedom	df =	24
t value for the two gorups	t = 1	. 56
t value required (. 05 level)	t = 2	2. 064

Significance of the Difference between Years of Teaching and Scores from the "Affective Rating Scale" Sub-scale, Style of Presentation

A t-test formula was used to compute the performance difference on the "Affective Rating Scale" sub-scale, style of presentation, between teachers who have taught from 2-8 years and teachers who have taught nine or more years. The results of these computations are shown in Table 13. No significant difference was found to exist between the ratings of the two groups for "style of presentation."

Table 13. Significance of the Difference between Years of Teaching and Scores from the "Affective Rating Scale" Sub-scale, "Style of Presentation"

	t-Test	
Teachers' years of teaching	2-8 years	9 or more years
Sample size	n = 11	n = 15
Raw scores for Style of Presen- tation	144	252
Squared scores	2, 030	4,896
Degrees of freedom	df = 24	
t value for the two groups	t = 1.61	
t value required (.05 level)	t = 2.064	

Significance of the Difference between Teachers' Age and DCA Scores

A t-test formula was used to compute the performance difference on the DCA between teachers from 20-40 years of age and teachers from 41-65 years of age. The results of these computations are shown in Table 14. No significant difference was found to exist between the DCA scores for the two age groups.

Table 14. Significance of the Difference between Teachers' Age and DCA Scores

<u>t-Test</u>		
20-40 yrs. of age	41-65 yrs of age	
n = 20	n = 10	
4220	2045	
895, 204	421,113	
df = 28		
t = 1.01		
t = 2.048		
	20-40 yrs. of age n = 20 4220 895, 204 df =	

V. DISCUSSION

This study was developed on the heels of an extensive cultural awareness training project conducted by the Oregon State Department of Education. Teachers from all over the state were thus involved in an instructional program that lacked an evaluative process. Were the teachers changed in any measurable way by this process? Were minority or majority culture children being helped or hindered by the teachers participating in this landmark attempt? The results of this study support the null hypothesis: no significant correlation was found between the DCA scores measuring cultural awareness and affective teaching behavior as rated on the "Affective Rating Scale." The results obtained reflect both the base from which the research was launched and the standard research procedures that guided the study.

Variables Affecting the Study

Evidence gained from this research indicates that it is extremely difficult, if not impossible, to control the variables affecting cultural attitudes. This is a problem faced in all studies of attitudes but its severity may be even greater when investigating the cultural awareness of teachers.

Teacher Variables

The subjects of this study came from many different geographic and socio-economic backgrounds as well as many different college and teacher training experiences. Some of the teachers had wide experience in various parts of the country and others were Oregonians who had never lived or studied outside the state. Some of the teachers had taught in other states. Others had only taught in one school with one socio-economic group of children. The multi-dimensional aspects of the teacher's personal background, education and training, teaching experience, psycho-physical nature, social situation, professional goals, etc., all contributed to a unique set of variables. These variables were reflected in the interpretations made by the teachers when reading and answering the questions on the DCA. Words, terms, and concepts were understood differently based on the type and range of experience of each teacher. Their inclination or disinclination for participation in the study itself was reflective of their experiencial background and training.

Variables Associated with the Setting

In addition to the above factors, when correlating attitude with behavior, there exists a whole dimension of variables within the behavioral setting which interact with, and affect attitudes and behavior of teachers. In this study several of the teacher subjects were

teaching in small, rural, isolated communities such as Paulina,
Oregon in the central part of the state. In such schools the student
population is small and grade levels are often combined. The community itself is dispersed on ranches as distant as 50 miles from the
school. The social experiences of the students, parents and teachers
in such settings is much different from that in urban areas such as
Salem, Oregon. School policy, supervising personnel, the children,
the classroom, the natural environment, population density, the
community, and the political atmosphere are all factors which influence attitude and behavior. Such factors constitute variables which
continually eluded control in this study.

<u>Demographic Variables</u>

This study was restricted to classrooms where minority children comprised less than a "significant" percentage of the class (five or less per classroom). Although this may have limited the scope of the investigation, it would be extremely difficult to find 30 elementary classrooms in Oregon with a high percentage of racially different students with teachers who would be willing to participate in cultural awareness testing. Furthermore, that kind of sample would not prove representative of the demographic distribution in this state. Finally, as this study indicates, it is important to examine and compare classrooms which serve less than a

"significant" degree of diverse minority backgrounds. However, in other situations different teacher responses and behavior might be expected as a result of the social pressure (misbehavior, alienation, parent concern, etc.) from larger numbers of minority children. In the present study the teacher could afford the leisure of pretending that no cultural differences exist.

Observational Variables

Different behavioral characteristics were observed with each teacher in the study. Some teachers were using inquiry methods during the observation period while others were not. Some teachers were observed dealing with student behavior problems while such problems did not arise during the observation period in other situations. Since behavior may be erratic and is dependent upon the large array of factors already discussed, long-term observation would be necessary for a thorough analysis of a teacher's behavior patterns.

Video taped time samples of behavior, as used in this study, are considered acceptable research procedure since they provide one of the few practical methods of observing numerous subjects. Nonetheless, such observations may miss behavior that is infrequent yet valuable for understanding the full scope of a teacher's behavior with children. For example, in one instance during this study a teacher was observed responding to a child who had lost a tooth during the

class session. This event did not occur again. In another instance the camera failed to capture a short but important fight between two students in the classroom because it was turned toward the teacher who was engaged in writing on the blackboard. Such problems are unavoidable when video taped time samples of behavior are used as part of the rating process.

Treatment Variables

The video taping of the teachers' behavior was carried out according to their teaching schedules. Some sessions were taped early in the morning, others at mid-day, and some in the afternoon. It is speculated that the time of day affected both the teacher and the childrens' behavior. For instance, it was noticed by this observer that the pupils were more ready to settle down and engage in their studies in the early morning than later in the day.

Teachers in the study ranged in age from those in their early twenties to those nearing the retirement age of 65. The ability and willingness of the teachers to examine their attitudes and behavior as it relates to cultural awareness may be connected with age. When approached, it did appear that older teachers were more reluctant to take part in the study. They wanted more assurance of their anonymity than did the younger teachers.

Varying degrees of pressure and persuasion to participate in

study, training and evaluation related to cultural awareness came from school administrators, other teachers, and pupils themselves. In this study it was noted that if the principal was not favorable toward the study, teachers were not allowed to have their classrooms video taped. Most of the teachers who participated in the study were recruited by teacher peers interested in the study. In schools where there was no positive, supportive teacher to encourage others to participate in the study, this researcher was unsuccessful in gaining entrance. In one instance where a large number of teachers were participating in the study, the pupils convinced one of the reluctant non-participating teachers to take part.

Rating Scales

While there are a number of rating scales used in making behavioral observations of teachers, all are limited to only a few types of behavior such as use of expression, reactions with pupils, and style of presentation. At the present time, there is no known method of taking all aspects of teacher behavior and incorporating them into an acceptable list of criteria sufficient for objective analysis. Selection of those behavioral characteristics deemed essential to a particular study of teaching behavior depends on the nature of the study.

In the area under investigation, no criteria have been

ascertained for specification of teacher cultural awareness. For instance, does cognitive knowledge of another culture constitute cultural awareness? Is a personal experience with an individual from another culture necessary for cultural awareness? Is overt behavior a valid expression of attitudes or is it, too, contaminated by cultural conventions and habits?

The "Affective Rating Scale" was selected as the most appropriate teacher behavioral rating scale available for this study after having investigated various other types of rating scales (see discussion under Method). All the scales appeared to have some limitations.

Limitations of the "Affective Rating Scale"

The criteria used in this scale may have been too broad. The categories may lack the discreteness required for any meaningful interpretation of the behavior being observed. On the other hand, the criteria used may be too limited in scope as to catch the full range of behavior encompassed in the area of affective teaching. It is also possible that the criteria may be inappropriate altogether for the nature of this investigation. However, the most flagrant limitation of the "Affective Rating Scale" is to be found in the validation process.

The "Affective Rating Scale" was validated on the basis of the use of "competent" judges.

The validity of the ratings is supported by the experience background of the judges. The three judges all were qualified supervisors who were well trained in the behavioral definitions of the three teaching variables rated ('use of expression, 'reactions with pupils', and 'style of presentation')." (Garrison and Kersh, 1969: p. 15).

The judges were education professors with PhD degrees and each had several years of teaching experience. The common practice of establishing validity on the basis of judge's competency raises some important questions:

Do judges give comparable ratings for the same reasons? Is it important that they do so? If similar rationale for making judgments is necessary, how can it be achieved?

Who selects the judges and how is the selection made? Is it sufficient to pick people with the same educational and philosophical background? Does such criteria tend to skew the ratings the judges will give?

How much is the construction of a rating scale of affective teaching the product of the values and thinking of those who construct it?

In the development of the "Affecting Rating Scale" Garrison and Kersh (1969) appear to have ignored these types of questions. They fail to discuss any limitations which these factors may pose. The use of "expert" judges as the criterion for attesting the validity of rating scales is common practice and certainly not limited to the "Affective Rating Scale." However, such validation procedure is questionable and reflects upon the adequacy of any such scales. Perhaps a method of validating the rating scale which combines the use of judges,

student performance levels, and teacher self-ratings would be more accurate and less disputable.

Implications of the Study

The results of this study did not fulfill the original expectations of the researcher. The null hypothesis was not rejected and no "significant" relationship could be demonstrated between cultural awareness and affective teaching behavior. However, the investigation did successfully bring new and vital clarity to the murky area of cultural awareness research. Furthermore, it serves to raise serious questions regarding on-going cultural awareness training programs.

For example, this study provides evidence that the expression of attitudes or values in one setting (on a paper and pencil test) is unlikely to correlate significantly with behavior in another setting (the classroom). Attitudes are psychological variables and teaching behaviors are objective or action variables. In each setting psychological variables are present. The problem is that they represent two different experiences having different settings and do not necessarily become congruent simply because they have some logical association or relationship. Rokeach (1968) in discussing these variables states, "It is meaningless to speak of two concepts that represent different universes of discourse as 'interacting' with one

another" (p. 127).

In addition, even though cultural awareness training may be desirable, it has yet to be demonstrated as an effective way of stimulating greater understanding and acceptance of minorities. Field interviews with some teachers who had taken Oregon's cultural awareness training revealed that the experience may in fact perpetuate stereotypes about minority people and reinforce prejudice. The training, while tending to heighten awareness of minority demands was unfortunately not as successful in dispelling fear and prejudice toward the culturally different. Many teachers claimed to have come away from the training sessions with less willingness to make changes which would accommodate minority demands and the special needs of minority children.

Finally, since federal and state agencies are spending thousands of dollars to provide cultural awareness training, it is reasonable to expect that this effort receive careful study. However, having demonstrated, through the DCA, the inadequacy of present standard methods, this research exposes a critical need for immediate innovative procedures. The most promising possibilities seem to lie in the direction of longitudinal studies which would provide continued monitoring and feedback between the teacher, trainer, and evaluator. Only by involving individuals from these and, possibly, other categories in the shared responsibility for development of cultural awareness will

it come to have any value as a principle of right action.

If teachers, among others, are engaged responsibly in the process of planning, implementing, monitoring, and evaluating a program of culture awareness education the probabilities would seem to be greater that they would benefit from such education, becoming effective agents for desirable change at least at the classroom level. Furthermore, such longitudinal studies would not only disclose the long-term effects of training but they would also serve to unearth the type of attitudes and behavior which, when linked together, work to stimulate understanding among people of different cultural backgrounds.

SUMMARY AND CONCLUSIONS

The original puprose of this study was the development of an instrument to measure the cultural awareness level of teachers.

Later the testing of the instrument led to a shift in focus from the instrument to examining the conventional wisdom about cultural awareness and teacher affectiveness.

The instrument developed in this study for testing the cultural awareness levels of teachers was the "Differential Cultural Awareness."

The DCA is a Likert-type scale with 58 items selected by competent judges and chosen from a large item pool. The item pool was formulated by this researcher and other competent people in the field of minority education in Oregon. A split half reliability test was used to establish the internal consistency of the DCA.

Thirty elementary school teachers from the State of Oregon were tested with the DCA. A sample of their classroom behavior was observed and filmed. Scores on the DCA were compared with ratings of the teachers' behavior that had been observed in their classrooms. The films of the teachers' behavior were rated by three competent, trained judges using the "Affective Rating Scale" (Garrison and Kersh, 1969). The judges' ratings were tested for inter-rated reliability using an analysis of variance (Winer, 1962).

⁴Hereafter, the "Differential Cultural Awareness" instrument will be referred to as the DCA.

A high level of inter-rater reliability was found between the judges' ratings of teacher behavior on the "Affective Rating Scale."

For the main data analysis, rank-ordered DCA scores and rank-ordered "Affective Rating Scale" scores were correlated using the Spearman rho formula. Analysis indicates that no significant correlation exists between the cultural awareness level of teachers and their teaching behavior.

A second analysis using the sub-category DCA scores with the behavioral ratings was correlated using the Spearman rho formula. Here again no significant correlation was found between the cultural awareness levels of teachers and their affective teaching behavior.

A third type of analysis using data from the DCA questionnaires for the age of the teachers and the number of years they had taught was computed by means of t-tests.

There was no significant correlation between DCA scores and teacher age. However, DCA scores do seem to correlate with the number of years a teacher has taught when comparing those having taught 2-8 years and those having taught nine or more years. The less experienced teachers received higher DCA scores. This is similar to the finding reported by Baty (1972) for teachers of Mexican-American children in northern California. This finding, while not conclusive, does point to a promising area for further research.

A comparison of DCA scores for each of the two groups of teachers based on years of teaching with the "Affective Rating Scale" scores was made. It was found that a significant correlation occurred on only one of the three sub-scales of the "Affective Rating Scale," the sub-scale for "use of expression." The importance of this correlation is questionable in the absence of other substantial supportive data. A chance correlation in one such treatment of the data is a likely explanation in the light of the large number of statistical calculations which were performed.

Although the null hypothesis predicted no correlation between teacher performance on the DCA test instrument and the affective teaching behavior, it was expected that the null hypothesis would be rejected. Since the results of the research failed to reject the null hypothesis, it has led to examination of the relationship between a paper and pencil test of cultural awareness and affective teaching behavior in the classroom. Here we have two different sets of attitudes and behavior—one, in responding to the paper and pencil test: The other, in teaching in the classroom. The problem is that they represent two different experiences having different settings. It does not follow that these two settings will become congruent simply because they may have some logical association or relationship.

This study provides evidence to suggest that the variables are too numerous and uncontrollable to provide a reasonable expectation

that a significant correlation can be found between the cultural awareness levels of teachers on a paper and pencil test and their affective
teaching behavior in the classroom. This research does point out
that efforts to train teachers in cultural awareness as a means of
enhancing their ability to teach in multi-ethnic situations, as presently
being conducted in the State of Oregon, are questionable.

In conclusion, the results of this study demonstrate the need for a new approach to the study of cultural awareness, one that combines both research and evaluation with a system for monitoring and providing feedback between trainers, teachers, evaluators and others interested in culture awareness education.

Findings

The analysis of the data from the study revealed the following:

- 1. There was no significant correlation between the composite DCA scores of teachers and their scores on the "Affective Rating Scale" (see Table 5).
- 2. There was no significant correlation between the sub-category scores of the DCA and the scores for the sub-scales of the "Affective Rating Scale" (see Tables 6, 7, and 8).
- 3. There was a significant difference between the DCA scores of beginning teachers and more experienced teachers. Beginning teachers scored higher (see Table 9).

- 4. There was a significant difference in the "Affective Rating Scale" sub-scale scores for "use of expression" between beginning teachers and more experienced teachers. Beginning teachers scored higher (see Table 11). However, there was no significant difference on the other two sub-scales of the "Affective Rating Scale" between these two groups of teachers (see Tables 12 and 13).
- 5. There was no significant difference between the DCA scores of younger and older teachers (see Table 14).

Recommendations for Further Research

There needs to be further research in all aspects of cultural awareness as it pertains to education. Studies that would provide longitudinal data on the effects of teacher training for cultural awareness are urgently needed to answer questions such as: Do teachers who have been trained in cultural awareness demonstrate that they are better equipped to teach in multi-ethnic settings? Do the minority students under their tutelage have higher performance levels and/or continued academic interest? Do cultural awareness programs which involve the teachers in all aspects of the training, monitoring, and evaluation prove more successful than programs which do not involve the teachers in the planning? Are different types of culture awareness education needed for teachers from schools with significant numbers of culturally different students as opposed to schools with few such

students?

There is a need for a definitive classification of teaching behavior associated with cultural awareness. At present, research is limited to the use of general behavioral rating scales of affective teaching. It would greatly enhance the development of more affective teacher preparation in the area of minority education and/or multicultural education if there was a more precise understanding of what behavior characterizes teacher cultural awareness.

There is a need for a comparative study of the long range value of college teacher education programs which attempt to prepare new teachers for assignments in multi-cultural settings versus inservice training programs and workshops which attempt to re-educate teachers for such settings.

BIBLIOGRAPHY

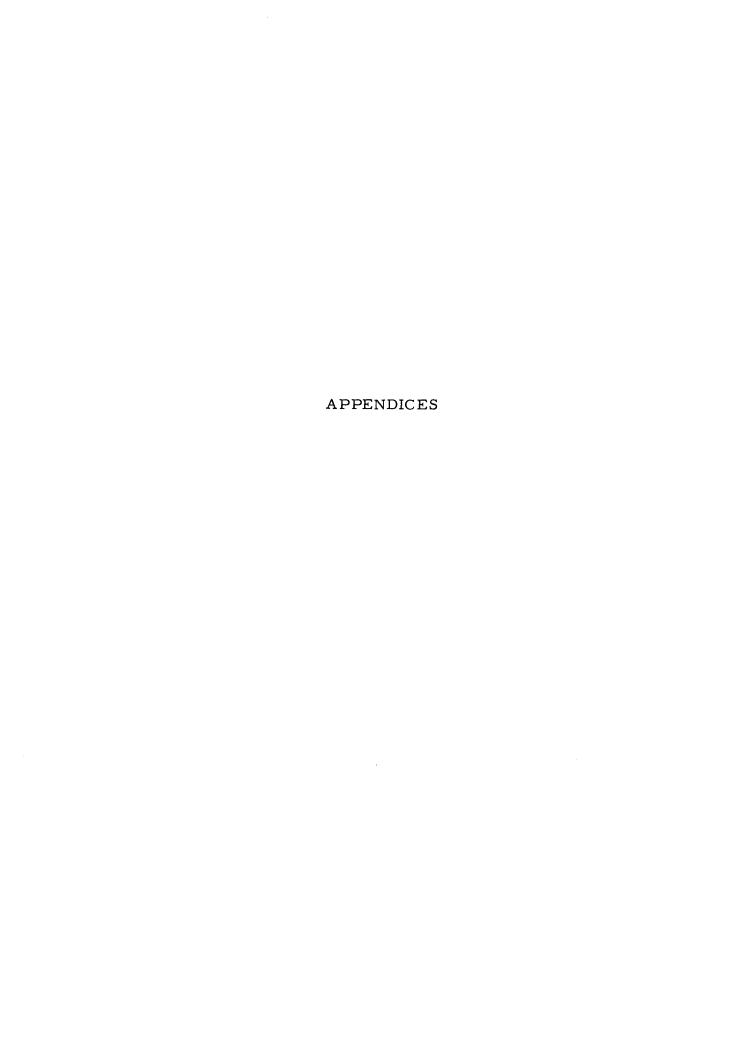
- Allport, Gordon W., Pattern and Growth in Personality, Holt, Rinehart and Winston, New York, 1961.
- Anastasi, Anne, 1961. Psychological testing. New York, The MacMillan Company.
- Banks, James A., Multicultural Education: In Search of Definitions and Goals. November 1974 (ERIC Ed 100 792)
- Baty, Roger M., Reeducating Teachers for Cultural Awareness:
 Preparation for Educating Mexican-American Children in
 Northern California, New York, Praeger Publishers, 1972.
- Beals, Alan. Culture in Process, Holt, Rinehart and Winston. N. Y., 1967.
- Brembeck, Cole S., and Hill, Walker H. Cultural Challenges to Education, D. C. Heath and Co., Lexington, Massachusetts, 1973.
- Brislin, Richard W., et al., Cross-Cultural Research Methods, John Wiley & Sons, N.Y., 1973.
- Bruning, James, and Kintz, B. L., Computational Handbook of Statistics, Scott, Foresman and Co., Glenview, Ill., 1968.
- Carter, Thomas P., 1969, Preparing teachers for Mexican American children. ERIC Clearinghouse on Teacher Education, Washington, D.C.
- Chilcott, John H., Greenberg, Norman C., Wilson, Herbert B., Readings in the Socio-Cultural Foundations of Education, Belmont, California, Wadsworth Publishing Company, 1969.
- Cole, M., Gay, J., Glick, J. A., Sharp, D. W., The Cultural Context of Learning and Thinking: An Exploration in Experimental Anthropology, New York, Basic Books, Inc., Publishers, 1971.
- Cole, Robert, Children of Crisis, New York, Dell Publishing Co., Inc., 1964.

- Cormack, Charles, an unpublished monograph on Cultural Awareness, 1975.
- Devore, Paul, "Variable Affecting Change in Inservice Teacher Education." Final Report, West Virginia University, Morgantown, W. V., September 1971, (ERIC Ed 070 764)
- Edwards, Allen L., 1957. Techniques of attitude scale construction. New York, Appleton-Century-Crofts.
- Eells, K, Davis, A, Havighurst, R. J., Herrick, Tyler, R., Intelligence and Cultural Differences, Chicago, University of Chicago Press, 1951.
- Garrison, Jesse H., and Kersh, Bert Y., The Comparison of Two Forms of a Classroom Simulation Test Designed to Enhance Future Teachers' Self-Definition and Teaching Style: Final Report, U. S. Department of Health, Education and Welfare, Office of Education, Project No. 9-I-033, October 1969.
- Guilford, J. P., Fundamental Statistics in Psychology and Education, McGraw-Hill Book Company, New York, 1956.
- Guilford, J. P., Psychometric Methods, McGraw-Hill Book Company, New York, 1954.
- Heath, G. Louis, Red, Brown, and Black Demands for Better Education, Westminster Press, Philadelphia, 1972.
- Henry, Jules, "A Cross-Cultural Outline of Education," Current Anthropology, I., 4 (July, 1960), pp. 267-305.
- Hudson, Herman, "The Black Studies Program: Strategy and Structure," The Journal of Negro Education, Vol. 41, Fall 1972, pp. 294-298.
- Ianni, F. A. J., and Storey, Edward, Cultural Relevance and Educational Issues: Readings in Anthropology and Education, Boston, Little, Brown, and Company, 1973.
- Jaramillo, Mari-Luci, 1973. "Today's schools are more culturally aware." Educational Leadership, Vol. 30, No. 7, p. 628-631.
- Johnson, H. S., Hernandez-M., W. J., Educating the Mexican-American, Valley Forge, Judson Press, 1970.

- Kerlinger, Fred N., 1973. Foundations of Behavioral Research. New York, Holt, Rinehart and Winston, Inc.
- Kroeber, A. L. and Kluckholn, Clyde, Culture: A critical review of concepts and definitions, Random House, New York, 1952.
- Landes, Ruth, Culture in American Education, New York, John Wiley & Sons, Inc., 1965.
- Likert, R., Roslow, S., and Murphy, G., "A Simple and Reliable Method of Scoring the Thurstone Attitude Scales," Journal of Social Psychology, Vol. 5, 1934, pp. 228-238.
- McCloskey, Elinor F., Urban Disadvantaged Pupils: A Synthesis of 99 Research Reports, Northwest Regional Educational Laboratory, Portland, 1967.
- Mehrens, William A., and Lehmann, Irvin J., Measurement and Evaluation in Education and Psychology, Holt, Rinehart and Winston, New York, 1973.
- Oppenheim, A. N., 1966. Questionnaire design and attitude measurement. Basic Books.
- Oregon Rules for Certification of Teachers, Specialists, and Administrators, 1975 Teacher Standards and Practices Commission, State of Oregon, Salem, January 1, 1975.
- Parsons, Talcott, The Social System, Tairstock Publications, London, 1952.
- Poppleton, P. K., and Pilkington, G. W. A comparison of four methods of scoring an attitude scale in relation to its reliability and validity. British Journal of Social and Clinical Psychology, 1964, Vol. 3, 36-39.
- Reece, Carolyn, 1974. Black Self-concept. Children Today, Mar-Apr 1974.
- Riessman, Frank, The Culturally Deprived Child, New York, Harper & Brothers, 1962.
- Riessman, Frank, The Culturally Deprived Child: A New View, School Life, April 1963, p. 7.

- Rokeach, Milton, Beliefs, Attitudes and Values, Jossey-Bass Inc., San Francisco, 1968.
- Rosenberg, Milton, et al., Attitude Organization and Change. Yale University Press, New Haven, 1960.
- Ryans, David, Characteristics of Teachers, American Council on Education, Washington, D. C., 1960.
- Selinger, Alphonse D., The American Indian High School Dropout: The Magnitude of the Problem, Portland, Northwest Regional Education Laboratory, 1968.
- Sherif, Carolyn and Sherif, Muzafer, 1967. Attitude, ego-involvement, and change. New York, John Wiley and Sons.
- Sidowski, Joseph B., edited by, Experimental Methods and Instrumentation in Psychology, McGraw Hill Book Co., New York, 1966.
- Skeel, Dorothy J., "Determining the Compatability of Student Teachers for Culturally Deprived School by Means of a Cultural Attitude Inventory," Unpublished doctoral dissertation, Pennsylvania State University, 1966.
- Spindler, George D., Education and Culture: Anthropological Approaches, New York, Holt, Rinehart and Winston, 1963.
- Spindler, George D., 1974. Education and Cultural process. New York, Holt, Rinehart, and Winston.
- Stenhouse, Lawrence, Culture and Education, Thomas Nelson and Sons, London, 1967.
- Stent, Madelon D. et al., Cultural Pluralism in Education: A Mandate for Change, Appleton-Century-Crofts, New York, 1973.
- Stone, James C., Teachers for the Disadvantaged, Jossey-Bass, Inc., San Francisco, 1969.
- Summers, Gene F., Attitude Measurement, Rand McNally, Chicago, 1971.
- Swick, Kevin J. and Lamb, Morris L. (ERIC Ed 073 025) Development of Positive Racial Attitudes, Knowledge and Activities in

- Pre-Service Social Studies Teachers, Southern Illinois Univ., January 1972.
- Swick, Kevin J. and Lamb, Morris L. (ERIC Ed 070 023) A
 Conceptually Based Social Studies Model for the Training of
 Teachers: Emphasis on Improved Racial Attitudes, Knowledge,
 and Activities, Nov. 1972.
- Thorndike, Robert L. and Hogen, Elizabeth, Measurement and Evaluation in Psychology and Education, John Wiley and Sons, New York, 1969.
- Thurstone, L. L. Attitudes can be measured. American Journal of Sociology. 1928, Vol. 33, p. 529-554.
- Thurstone, L. L. and Chave, E. J. The Measurement of Attitude. Chicago: University of Chicago Press, 1929.
- Tittle, C. R. and Hill, R. J., Attitude Measurement and Prediction of Behavior: An evaluation of conditions and measurement techniques. Sociometry, 1967, Vol. 30, 179-213.
- Triandis, Harry C., 1971. Attitude and attitude change. New York, John Wiley and Sons.
- Tylor, E. B., Primitive Culture, John Murray, London, 1871.
- Von Maltitz, Frances Willard, Living and Learning in Two Languages: Bilingual-Bicultural Education in the United States, McGraw-Hill Book Company, New York, 1975.
- Winer, B. J., Statistical Principles in Experimental Design, McGraw-Hill, New York, 1962.
- Wischler, Irving R. and R. E. Bernberg, 1950. Indirect methods of attitude measurement. International Journal of Opinion and Attitude Research. Vol. 4, p. 209-228.
- Zirkel, Perry Alan and Greene, John F., Cultural Attitude Technical Report, April 1974 (ERIC Ed 102 196)



APPENDIX I

Cultural Awareness Questionnaire Fall 1972 Crawford

1.	Minorities get more breaks than the average	Strong ly agree	Agree	Uncertain	Disagree	Strongly disagre e
1.	person.					
2.	All students are treated alike within a classroom.			0		
3,	Parents of low income families are not motivating their children toward education.	0	0		0	_
4.	All students need to learn about and appreciate others who come from other cultural backgrounds.	0	0	0	_	
5.	Competition in school is a valuable tool in the learning process.	Ö	0	0	0	0
6.	Minority students need to be taught the values and attitudes which characterize American society.			 7 -		· , .
7.	A high incidence of theft is prevelant among minority and lower class students.	0	0		0	_
8.	Friendship at school between racially mixed students should be encouraged.	_	0	0	0	0
9.	People of minority background are needed as teachers in our public schools.	0	0	0	0	
10.	Children from low income homes need special consideration regarding school lunches.	0	0	0	0	0
11.	Schools encourage cultural diversity.					
12.	In school we minimize the differences of students from minority cultural backgrounds.	0	0	_	_	_
13.	Schools need to give special attention to	П		П		П

14.	It takes special training for a white middle class person to understand and teach children from poor and minority homes.				_	0
15.	Teachers are usually better equipped to under- stand the needs of their students than are the students' parents.			_		
16.	Minority children have happy dispositions.					
17.	Students in a racially mixed school have a real advantage in their education.					
18.	The minorities are undermining our government and schools.					
19.	Lower class students lack the drive to achieve that is found among middle class students.					
20.	The subordinant position within our society occupied by minorities and the poor accounts for their lack of self-assurance and self-worth.	0				
21.	Treating all students alike regardless of race, creed or class is the goal of the public schools.		_			
22.	A teacher should not be hired who is not willing to work under a person of another racial back- ground.				0	
23.	The term culture refers to all behavior and life style of a people.					
24.	People may be considered culturally distinctive if they eat different types of food, dress differently than others, and have different beliefs.	0			0	
25.	The schools should attempt to involve the parents of minorities and the poor by letting them help decide what should be taught in the curcirulum.	0	0			_
26.	Minority groups should have representation on the school board in proportion to their numbers in a community.	0				
27.	Teachers need to receive special training in how to understand and teach students from different ethnic, racial and class backgrounds.	0				

28.	Administrators of schools need special training in how to work with minority and low income people.					
29.	Vocational training is probably the best program for students from minority backgrounds.					
30,	Minorities don't even know their own language.					
31.	Cultural values shape a student's personality.					
32.	Minority youths have personality problems.					
33.	A child's school behavior reflects his culture.					
34.	A child's behavior tells us about his feelings.					
3 5.	Minority youth behave differently.					
36.	The socioeconomic background affects the ability of a child to learn.					
37.	A child's capabilities are linked to his socio- economic level.			0		
38.	Minority and lower class kids have bad grammar.					
39.	A minority child's silence shows his resistence.					
40.	Silence is a method of communicating a child's feelings.					
41.	Bad behavior is characteristic of lower class kids.					
42.	Parents of minority kids care little about their school behavior.					
43.	Poor behavior is characteristic of minority kids.					
45.	Language skills reflect the socioeconomic level of a child's family.					
46.	A child's language skills reflect the level of communication at home.					
47.	Students from minority homes use a great deal of nonverbal communication.					
48.	Middle class students have a facility for oral communication.			0		
49.	Middle class students behave well in school.	П	П	П	П	П

50.	Well ormed personality's are characteristic of					
	middle class children.					
51.	Middle class kids have good self concepts.					
52.	The self concept of minority children is related to acceptance at school.					
53.	The self concept of a child is related to the families socio-economic level.	<u> </u>		0		
54.	Minority children tend to have a poor self concept,					
55.	Teachers need to take a special interest in the minority child.	0				
56.	A teacher needs to give special attention to minority children.	0	-	0	0	
57.	A teacher needs to give special help to minority children.		_	0	₂ 🗖	
58.	It is wrong for a teacher to give special help to a minority child.		0			
59.	It is wrong for a teacher to give special attention to minority children.	0				
60.	Minority children need special attention.					
61.	Minority children have a good self concept.					
62.	Minority children communicate well in their own language.		0	0		_
63.	Minority children are alert.					
64.	Minority children are competitive.					
65.	Minority kids are troublemakers.		0			
66.	Teachers are generally good models for kids from minority homes.		0	0		Ö
67.	The school needs children from every ethnic group.		0		0	0
68.	Schools reflect the values of the majority.					
69	Middle class teachers teach middle class values.	п	П	п	п	п

70.	Middle class values are for everyone.			
71.	White middle class values fail to meet the needs of minorities and the poor.			
72.	Teachers should teach their students about minority values.			
73.	Every student needs to know the school believes his culture is valuable.			
74.	Teachers are among the worst bigots.			
75.	Middle class white cultural values are presented as the norm in the schools.			
76.	Students should respect the teacher.			
77.	Minority students are shy.			
78.	Minority students are aggressive.			
79.	Minority students are beligerent.			
80.	Minority students are nonparticipants in school.			
81.	Minority students expect favors.			
82.	Minority children are seldom good students.			
83.	Minority children are frequently superior students.			
84.	Language skills are poorly developed among minority students.			
85.	Minority students are good at sports.			
86.	Minority students take a lot of the teacher's time.			
87.	Minority students receive little attention from their teachers.			
88.	Minority children should be discouraged from expressing themselves at school in their cultural dialect.	0	0	
89.	Teachers are to be examples of society's values.			
90.	Teachers must set a good moral standard.			

91.	Teachers should not force their values on their students.					
92.	Teacher education fails to prepare teachers to teach non-middle class children.	0				
93.	Teacher education provides the teacher with adequate cross cultural awareness of minorities.					
94.	Cross cultural awareness is an area that needs emphasis in teacher education programs.					
95.	Cross cultural awareness is lacking in the public schools.					
96,	Valuable cultural alternatives are available when minorities exist within a community.					
97.	Extra classroom time is needed for presenting minority cultures.					
98.	We must seek to develop a common culture for all Americans.		0			
99.	All Americans share a common culture.					
100.	The children of the poor resist being sent to school.					
101.	Children from poor and minority homes show appreciation when treated fairly.					
102.	Middle class children don't appreciate what they have.					
103.	Teachers favor middle class children.					
104.	Schools cater to the demands of middle class parents.					
105.	Schools cater to the demands of minority parents.		0			
106.	Minority parents care about the schools influence upon their children.	_				
107.	Minority parents take an interest in their children's performance at school.					
108.	Teachers who are of a minority background them-	п	п	п	п	п

109.	Teachers of minority backgrounds themselves are over sensitive to the needs of minority children.			
110.	Teachers of minority background themselves have great cultural awareness.			
111.	Schools of predominantly lower class and minority children suffer from lack of parental support.			
112.	Minority kids are highly aware of cultural differences within the community.			
113.	Discrimination against minorities is common on the part of teachers.			
114.	Discrimination in the school takes a subtle form.			
115.	Discrimination is part of human nature.			
116.	Minorities discriminate against the majority.			
117.	Discrimination is a sickness within our society.			
118	Minorities would prefer to have segregated schools.			
119.	Minorities want to control the schools.			
120.	Minorities would like to control the schools in their neighborhoods.			
121.	Minorities should be in control of the schools in their neighborhoods.		O	
122.	Cultural pluralism is a desirable goal for our society.			
123.	Cultural pluralism is a means of accepting various cultural groups within our society.			
124.	Cultural pluralism won't work in our society.			
1 25.	Minorities do have inferior IQ'a on an average.			
126.	Valuable results are obtained when IQ tests are used cross culturally.			
127.	Invalid results are to be expected when using IQ tests cross culturally.	0		

128.	Cross cultural comparisons on the basis of IQ testing is hazardous.				
129.	Socioeconomic level is an important variable when testing in the public schools.		.		
130.	Middle class children possess high IQ's.				
131.	America is a cultural melting pot.				
132.	In America the melting pot concept is a myth for people of color.		0	0	
133.	Minorities are discriminated against because they are inferior.			_	
134.	Minorities occupy the lowest socioeconomic status in America.				
135.	Minority children are unconfortable in a pre- dominantly white middle class school. Whites are uncomfortable around minorities.	0	_		
136.	Food and dress may distinguish one cultural group from another in America.				
137.	There are no significant cultural differences between the majority of Americans and the racial minorities.		_		_
138.	Skin color may be used to distinguish between the races.	_			
139.	A race is distinguish by its skin color.				
140.	Racial classifications are inadequate means of distinguishing between people.				
141.	Grouping children according to race helps the school teachers and administrators.				
142.	Children should not be grouped in the school according to race.				_
143.	Children prefer to associate with others of their own race.				
144.	Assimilation of cultural minorities is a desirable goal for the public schools.				

145.	Assimilation is a racist approach to handle cultural minorities within our society.			
146.	Minority groups should be assimilated by the cultural majority in our society.			
147.	School children must be taught basic rules of behavior.			
148.	A public school shouldn't try to meet the cultural needs of each of its students.			
149.	Schools should attempt to meet the needs of each student.			
150.	Lower class students have realistic life goal expectations.			
151.	Lower class students are objective in their estimation of present and future opportunities.			
152.	Lower class students have unrealistic life goals expectations.			
153.	Integration is a desirable goal for minority students within the public schools.			
154.	Present integration in the schools is based on acceptance of middle class behavior and values.			
155.	Inferior intelligence is expected among minority children because of smaller average brain size.			
156.	A low nutritional level of the lower class and minority children contributes to poor performance in school.			
157.	No significant difference in brain size is found among the races of the world.			
158.	Oral communication is difficult for minority children.			
159.	Minority and lower class children have effective oral communication skills.			
160.	IQ tests are valid measures for cross cultural comparisons			0

APPENDIX II

Confidential Questionnaire Related to Teaching (check and fill in all items)

Teacher's Aide □ Teacher □ Schoo	l Administrator 🗆	Male □ Female □	
Married □ Single □ Age: 20-25 □	26-30 □ 31-40 □	□ 41-50 □ 51-65 □	
Ethnic Background: White Indian] Mexican-Ameri	can 🗆 Black 🗆 Other 🗆	
The grade you teach now	Number of years i	n district	-
Teaching Experience: K-3 □ 4-6	□ 7-9□ 10-12	☐ Total years of teach	ning experience
Have you taught in a classroom with 5 American, Black, etc.) Yes □ no		y different students (Indi	ian, Mexican-
How many years experience of teaching Where have you had this experience	•		_ _
	City	State	
-	City	State	_
Where did you grow up?City		State	
Was it: rural □ small town □	urban 🗆		
Academic Training	where		
No. of years of college	where		
Bachelors degree earned	where		<u> </u>
No. of years of graduate training	where		
Masters degree earned	where		
No. of years beyond masters	_		

INSTRUCTIONS FOR QUESTIONNAIRE

Please fill in an answer for each question.

Write additional comments on the back of the Questionnaire if you like.

Avoid, when possible, answering in the "uncertain" column.

Form A

Code 1	No. :					
		Strongly agree	Agree	Uncertain	Disagree	Strongly
1.	Minority children have a poor self concept. (The way they feel about themselves.)			0		
2.	Minority groups should be assimilated by the cultural majority in our society.			0		
3.	We must seek to develop a common culture for all Americans.	0		0		
4.	Language skills are not well developed among minority students.					
5.	Minority parents take little interest in their children's performance at school.			0		
6.	Teachers need to give special attention to minority children.					
7.	Discrimination against minorities is common on the part of teachers.					
8.	Assimilation is a racist approach to handle cultural minorities within our society.					
9.	English language tests discriminate against children of minority backgrounds.					
10.	It is wrong for a teacher to give special attention to minority children.					
11.	Students in a racially mixed school are short changed in their education.					

12.	The minorities are undermining our government and schools.	0				
13.	Bilingual, bicultural education should be a part of every public school.			<u> </u>		
14.	Cross cultural awareness is lacking in the public schools.	0	0			
15.	Teachers are not able to understand the needs of students as well as their parents.	0				
16.	Teacher education fails to prepare teachers to teach non-middle class children.	0		0		
17.	Minority students are good at sports.					
18.	Minority students take a lot of the teacher's time.	0			–	
19.	Minority groups should have representation on the school board in proportion to their numbers in the district.		0	<u> </u>	0	0
20.	Treating all students alike regardless of race, creed or class is the approach of the public schools.	0	0		0	_
21.	Teachers teach their students about minority cultural values.					
22.	Assimilation of cultural minorities is not a desirable goal for the public schools.	0	0			_
23.	Language skill is a measure of students' intelligence.			<u> </u>	D	0
24.	Vocational training is probably the best progra for students from minority backgrounds.	.m.	0	0		
25.	Teachers do not favor middle class children.					
26.	A teacher should not be hired who is not willing to work under a person of another racial background.	0	0			0

27.	A high incidence of theft is prevalent			
	among minority students.			
28.	Minorities want to control the schools.			
29.	Middle class teachers teach middle class values.			
30.	There are no significant cultural differences between the majority of Americans and the racial minorities.			
31.	Oral communication is difficult for minority children.		, 	
32.	Teachers of minority backgrounds themselves are over sensitive to the needs of minority children.			
33.	Poor behavior is characteristic of minority kids.			
34.	The self concept of minority children is related to acceptance at school.			
35.	A public school shouldn't try to meet the cultural needs of each of its students.			
36.	Public schools cannot handle children lacking English language skills.			
37.	A minority child's silence is not indicative of resistance to the teacher.			
3 8.	Minority parents don't care about the school's influence upon their children.		_	
3 9.	Teachers do not force their values on their students.			
40.	Minority children communicate well in their own language.		_	
41.	Minorities would prefer to have segregated schools.			
42.	Minorities ought to be in control of the schools in their neighborhoods.			

43.	Minority children resist being sent to school.				
44.	Minority parents do not motivate their children toward education.				
45.	Minority children should be discouraged from expressing themselves at school in their cultural dialect.	0			
46.	Discrimination is part of human nature.				
47.	The schools must attempt to involve the parents of minority children by letting them help decide what should be taught in the curriculum.				0
4 8.	Minority students are non-participants in school.		0		
49.	Present integration in the school is based on acceptance of middle class behavior and values.				
50.	Cultural pluralism is not a desirable goal for our society.				
51.	Parents of minority kids care little about their children's school behavior.				
52.	America is a cultural melting pot.				
53.	Minority children have effective oral communication skills.		0		
54.	Valuable cultural alternatives are available when minorities exist within a community.				
55.	Teachers need to take a special interest in the minority child.				
56.	Minorities would like to control the schools in their neighborhoods.				
57.	Schools ignore the differences of students from minority cultural backgrounds.				
58.	In America the melting pot concept is a myth for people of color.				П

APPENDIX III

Affective Rating Scale

USE OF EXPRESSION

dynamic voice qualities facial expression smiles, frowns bodily movements monotonous, dull voice passive, fixed facial expression, stiff, stilted posture

1 2 3 4 5 6 7 8 9 10

REACTIONS WITH PUPILS

responsive attentive
evidences interest and concern
elicits responses
accepts and recognizes students ideas

ignores, seems insensitive interest only in "lesson" gives questions/answers stylized/dogmatic

1 2 3 4 5 6 7 8 9 10

STYLE OF PRESENTATION

flexible and open exciting to students interactive

rigid, stiff boring, unreal to students follows preconceived plan

1 2 3 4 5 6 7 8 9 10

APPENDIX IV SUBJECT SCORES on the

DCA AND AFFECTIVE RATING SCALE (Raw Data)

Subjects	<u>Je</u>	Use of Expression			Reaction with Pupils				Style of Presentation				
	DCA scores	r #1	er #2	r #3		r #1	r #2	r #3	-	r #1	r #2	r #3	
Sub	DC _	rater	rater	rater	total	rater	rater	rater	total	rater	rater	rater	total
1	250	6	7	5	18	6	6	4	16	5	6	7	18
2	235	3	3	2	8	4	3	2	9	5	4	2	11
3	235	8	7	8	23	8	7	8	23	9	7	8	24
4	231	7	6	7	20	6	5	6	17	6	6	5	17
5	229	3	2	1	6	2	3	1	6	3	2	1	6
6	219	5	5	6	16	5	5	4	14	5	7	6	18
7	217	4	4	4	12	3	3	4	10	3	4	5	12
8	216	7	7	6	20	7	7	7	21	7	7	7	21
9	215	6	5	7	18	5	5	6	16	4	5	6	15
10	215	4	3	4	11	2	2	3	7	3	4	3	10
11	214	3	4	3	10	4	5	4	13	4	5	4	13
12	213	3	3	2	8	2	3	2	7	3	4	2	9
13	213	4	6	5	15	4	5	3	12	4	6	4	14
14	211	4	3	4	11	3	4	3	10	5	4	4	13
15	211	9	8	7	24	9	9	8	26	9	8	8	25
16	210	5	5	3	13	4	5	3	12	5	6	5	16
17	210	8	9	9	26	7	8	9	24	8	9	9	26
18	205	4	3	5	12	4	4	6	14	5	4	6	15
19	201	2	4	3	9	2	4	3	9	3	5	3	11
20	200	6	6	4	16	6	7	5	18	5	7	6	18
21	200	7	6	8	21	5	5	7	17	6	6	8	20
2 2	199	3	4	4	11	4	3	2	9	5	3	2	10
23	199	4	3	4	11	2	2	1	5	2	2	1	5
24	195	6	4	4	14	5	4	4	13	7	5	5	17
25	194	8	7	6	21	7	6	4	17	8	7	6	21
26	190	2	2	2	6	2	2	2	6	2	3	2	7
27	189	5	7	6	18	3	5	4	12	6	8	6	20
28	187	8	8	7	23	5	2	4	11	7	8	6	21
29	185	9	8	8	25	7	7	8	22	8	8	8	23
30	1 <i>7</i> 7	4	4	5	13	4	4	5	13	4	4	5	13

APPENDIX V

DCA SUB-CATEGORY SCORES
AND SPLIT HALF SCORES
(Raw Data)

Cts	Composite DCA scores		DCA Sub-Category Scores						
Teacher Subjects		3. A	<u>~</u>	~	s rela-	rela- to ies	elated orities rica	Split Half DC Test Scores	
		minority language usage	minority students	minority parents	teachers rela- tionship to minorities	schools rela- tionship to minorities	items related to minorities in America	Odd	Even
1	250	38	45	38	48	44	3 7	130	12
2	235	31	48	33	42	39	42	118	11
3	235	32	50	31	45	3 7	40	118	11
4	231	3 7	45	32	39	42	36	125	10
5	229	35	42	34	39	40	39	1 23	10
6	219	32	41	32	39	3 7	38	113	10
7	217	33	43	25	41	3 7	38	113	10
8	216	33	43	32	36	3 7	35	114	10
9	215	31	3 7	28	35	40	44	104	11
10	215	34	3 7	27	44	39	34	104	11
11	214	28	3 7	32	42	3 7	38	104	11
12	213	27	38	32	41	3 7	38	109	10
13	213	31	41	30	38	36	3 7	105	10
14	211	31	42	28	44	31	35	110	10
15	211	30	45	32	32	31	41	105	10
16	210	30	40	30	3 7	35	38	109	10
17	210	30	38	29	38	36	39	102	10
18	205	28	45	27	33	34	38	101	10
19	201	21	40	27	43	33	3 7	97	10
20	200	29	38	27	36	38	32	106	9
21	200	27	34	26	41	34	38	100	10
22	199	32	38	28	32	3 7	32	99	10
23	199	28	34	28	40	36	33	103	9
24	195	27	40	29	34	33	32	102	9
25	194	27	38	25	33	35	36	93	10
26	190	28	33	30	37	31	31	95	9
27	189	32	47	28	27	29	2 6	97	9
28	187	24	3 7	26	33	35	32	93	9
29	185	26	38	28	3 7	30	26	95	9
30	177	24	40	24	33	28	28	88	8

APPENDIX VI

Reliability

This appendix is devoted to a discussion of three common testing procedures for obtaining reliability estimates: test-retest, equivalent test, and split half tests.

Test-retest estimates of reliability are obtained by administering a test to a group of individuals and then readministering the same test at a later date to the same group. Any variation in scores from one testing to the next is regarded as error when no treatment or training has been given the group being tested (Mehrens, 1974).

The advantage of the test-retest method is that it gives an estimate of the stability of the instrument over an extended time period. The difficulty with this procedure is that testing itself may cause a reaction which contaminates the results. Another problem is that practice and memory may affect the test results the second time the instrument is given. When the interval between testings is short there is greater likelihood of this form of contamination occurring. Time and scheduling limitations may also preclude the use of this procedure (Anastasi, 1961; Mehrens, 1974; Thorndike, 1969).

Split half or subdivided tests are probably the most widely used procedures for estimating reliability. These involve a single test being divided into two presumably equivalent halves. One method is to examine the content and difficulty of each item and then make a

Another method is to simply put all the odd-numbered items in one half-test and all the even numbered items in the other. This procedure is followed only for scoring, not for administration of the test (Anastasi, 1961; Thorndike, 1969).

The correlation between the two scores derived from the split half procedure provide a measure of the accuracy of the test's measurement of the individual. It must be remembered that the correlations are between the two half length tests and may not be directly applicable to the full-length test (Thorndike, 1969).

The advantage of the split half method is that only one test needs to be administered. Contamination from practice and memory are controlled using this procedure. The disadvantage of the split half method is that the scores are obtained within a given moment of time and could change for an individual from day to day and yet not be reflected in this type of reliability coefficient (Mehrens, 1973; Thorndike, 1969).

Reliability measurement using the equivalent form procedure requires that two parallel forms of a test be available and it requires sufficient time for two testings in the same day. Thorndike (1969) indicates that practical consideration of convenience and expediency work against the equivalent form procedure.

APPENDIX VII

Validity

This appendix is devoted to a discussion of three common types of validity as classified by the American Psychological Association, the American Educational Research Association, and the National Council of Measurement Used in Education.

Content validity is concerned with the substance, matter or topics of a measuring instrument. Content validation is guided by the
question: Is the content of the instrument representative of the area
or universe of content of the subject being measured. Content validation is judgmental. Each item of a test is judged on the basis of
whether or not it is representative of the universe or subject area
(Kerlinger, 1973; Thorndike, 1969).

For an understanding of criterion-related validity it is helpful to establish the meaning of criterion. David Ryans (1960) provides the following definition:

A criterion is a standard description, or definition, which is accepted in understanding research and is used to provide a frame of reference for judging whether or not some phenomenon occurs (and often the degree to which it occurs). It is a base, often of a rather arbitrary nature and ultimately involving value judgments, against which comparisons may be made (p. 26).

In other words criterion-related validity involves comparing test scores with one or more outside criteria which is thought to measure the attribute being studied. An illustration of criterion related validity might be, how accurate are the scores of a scholastic aptitude test in predicting the success of students in college. Predictive validity can be estimated by determining the correlation between the test scores and grades of students in college (Kerlinger, 1973; Thorndike, 1969).

The problem with criterion-related validity is finding a suitable criterion measure. As stated in the definition by Ryans it is a matter of judgment whether a test is related to a particular criterion (Ryans, 1963; Kerlinger, 1973).

With construct validity the researcher is more interested in the property being measured than in the instrument itself. According to Kerlinger (1973):

The significant point about construct validity, that which sets it apart from other types of validity, is its preoccupation with theory, theoretical constructs, and scientific empirical inquiry involving the testing of hypothesized relations (p. 461-462).

Evidence of validity is based upon both rational and empirical data. Judgment and evidence are brought together in the process of validation (Ryans, 1960).

APPENDIX VIII

Rating Scales

This appendix is devoted to a discussion of three common types of rating scales: category rating scales, numerical rating scales, and graphic rating scales.

Category rating scales consist of several categories from which the observer selects the one that best represents the behavior being observed. Such categories may include "expression" with ratings for:

very expressive

expressive

not very expressive

no expression

"Numerical rating scales are perhaps the easiest to construct and use. They also yield numbers that can be directly used in statistical analysis," according to Kerlinger (1973: p. 547).

The graphic rating scale employs lines and bars with descriptive phrases for each item along the graph. They have the advantage of representing a continuum and are easy to use (Kerlinger, 1973; Guilford, 1954).

The weaknesses of rating scales needs mention. These fall into two categories, one extrinsic and the other intrinsic. Kerlinger (1973) suggests that because rating scales are seemingly easy to

construct and use they are often used indiscriminately forgetting their intrinsic defects.

"The intrinsic defect of rating scales are their proneness to constant or biased error" (Kerlinger, 1973: p. 548). Among the most constant rating errors, the "halo effect" is significant. This involves the tendency for an observer to rate in the direction of his own likes and dislikes.

Other important sources of constant error are the "error of severity" and the "error of leniency." Here the tendency of the observer is to rate all individuals too low or too high. A corresponding error is the "error of central tendency" which eliminates extremes by rating right down the middle (Kerlinger, 1973; Thorndike, 1969).

Finally, in discussing criterion measures Ryans (1960) observes:

Probably in most situations involving natural behavior, or field behavior, it is necessary to employ criterion measures which have technical shortcomings (p. 39).