The purpose of this review is to report an analysis of recent research related to measurement of attitudes in reading. It is based on a search of the literature published primarily since the late 1960's. The emphasis is on research related to attitudinal measuring tools rather than the results of research utilizing such tools. The review is organized into four sections. First, a number of studies reporting development of reading attitude measures are reviewed. Second, general criteria recommended for consideration in attitude scale development are presented. Third, some general comments about measuring reading attitude are made. Finally, a special reference section presents additional sources related to measurement and the affective domain, attitudinal measurement in education, and attitudinal measurement in reading. Reading attitude measures are noted as representing three categories: self report instruments, direct observation, and projective techniques. Criteria for reading attitude test development are item development, response set contamination, and statistical procedures. It is noted that adequate conceptualization and definition of reading attitude constructs is still a major problem in research in the field.
A Review of Research and Bibliography on Published Attitudinal Instruments in Reading

Preconvention Institute: The Affective Domain: Alternative Approaches to the Measurement and Improvement of Attitude Towards Reading

21st Annual Convention of the International Reading Association
Annaheim, California - May 10-11, 1976

"A plethora of variables such as attitudes, values, interests, motivation, anxiety, appreciation, adjustment and other personality characteristics are generally subsumed under the rubric of affective-behavior. The term non-cognitive is also used to distinguish these from the task-oriented cognitive variables such as aptitude and achievement."

"If you can't adequately define or measure the concept, it must belong in the affective domain."

"Affective objectives emphasize a feeling tone, an emotion, or a degree of acceptance or rejection. Affective objectives vary from simple attention to selected phenomena to complex but internally consistent qualities of character and conscience."

"We regret the lack of research dealing with clear-cut affective objectives of the school."

* Appreciation is expressed to Sharon Jeroski and Donna Dartnell, M.A. students in secondary reading - U.B.C., who helped locate information.
"A satisfactory theory which would lead to an optimal categorization system for the affective domain has yet to be developed."

"If desirable affective goals are to be realized as a result of the educational process, relevant formal learning situations have to be developed and the effect of such learning experiences will have to be systematically appraised."

Introduction

The purpose of this review is to report an analysis of recent research related to measurement of attitudes in reading. It is based on a reasonably thorough, although by no means complete, search of the literature published primarily since the late 1960's. The emphasis is on research related to attitudinal measuring tools rather than the results of research utilizing such tools.

The review is organized into four sections. First, a number of studies reporting development of reading attitude measures are reviewed. Second, general criteria recommended for consideration in attitude scale development are presented. Third, the paper concludes with some general comments about measuring reading attitude. Finally, a special reference section appears which presents additional sources related to measurement and the affective domain, attitudinal measurement in education, and attitudinal measurement in reading.

Ebbesen, 1970), (Jahoda & Warren, 1972). Standard bibliographic sources were used in locating the research on reading attitudes, attitudes in education and the affective domain. The ERIC computer data base was searched through the facilities of the Information/Knowledge Research Centre, Faculty of Education, University of British Columbia.

**Attitude Measures in Reading**

"Attitude is one of the most used yet ubiquitous terms in social science."

"This widespread usage has detracted from the operational clarity of attitude and rendered it a pot pourri term with no generally accepted definition. Unfortunately, for the conceptual state of the field, it seems that this is precisely why it has proved so attractive, since each individual has been allowed to tailor the term to suit his purposes. One of the advantages that attitude has is that it can be applied at many different levels of analysis."

"The development of instruments to assess attitude has been one of the most problematic areas in psychometrics. The main problem has been the accurate definition of the word attitude and the inability to isolate attitudes as discrete behavioral attributes."

"In attitude measurement the crucial question must always be whether a given technique constitutes an adequate operational definition of the underlying theoretical assumptions. What's accepted as a basis for inference inevitably turns on what is meant by attitude."

Numerous technical definitions of attitude have been put forth. Campbell suggested in 1963 that psychologists use 80 concepts that share the operational definition of the concept of attitude (as cited by Triandis (1971) p.4). According to him all the concepts deal with phenomena that are acquired and that modify later responses of the organism.

Perhaps the most influential definition has been that of Allport, presented in 1935, and emphasized in two succeeding editions of *The Handbook of Social Psychology* (1968).
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.

Features of the above can be found in most definitions of attitude. Shaw & Wright (1967) synthesized earlier definitions and suggest attitudes are relatively enduring systems of covert, implicit affective and evaluative reactions which are based on and reflect learned evaluative concepts or beliefs about characteristics of social objects or classes of objects.

Although it's not necessary to coin an omnibus definition of attitude to engage in attitudinal research in reading, the above definitions provide some building blocks to guide research in conceptualizing, operationalizing and quantifying attitudinal concepts relative to focal objects in reading. Attitudes about reading exist within the individual and cannot be seen or observed in direct fashion. However, the presence of attitudes toward reading can be inferred from various behavioral samples. Attitudes held by a person toward reading will tend to cause that person to notice things and do things selectively. "We use our attitudes as anchors by which we contrast extremely discrepant attitudes and assimilate similar attitudes." Attitude is a response thus permitting a person's attitude to be inferred and elicited by providing appropriate verbal and nonverbal stimuli. School and school related activities (reading) are appropriate areas for attitudinal assessment because they are socially salient in the life of every student. An attitude has intensity and is

---

1 McGuire (1969, p.149) develops a useful conclusion regarding attitudinal definition. "This extended discussion of how attitude is best defined from a provocative heuristic point of view must not obscure the point that, in a given experiment on attitude, the term can very readily be given an operational definition in terms of observable and scoreable responses. Typically, the person's attitude regarding an object is operationally defined as the response by which He indicates where he assigns the object along a dimension of variability."
a personal evaluation containing either positive, neutral or negative valence.

Finally, Good's (1973) definition can serve as a general referent for conceptualizing attitudes in education.

The predisposition or tendency to react specifically towards an object, situation, or value; usually accompanied by feelings and emotions; attitudes cannot be directly observed but must be inferred from overt behavior, both verbal and non-verbal.

Attitude assessment falls primarily within the area of measurement of affect. Affect has been measured by instruments ranging from personality tests and statements of values to galvanic skin response. Because of its abstract role as an intervening variable, and because of the covert, inferential nature of the quantification of attitudinal concepts, research allowing for a broad array of instrumentation options would seem most fruitful. Different types of behavioral specimens, used in conjunction, will allow for better inferences about attitudes. Indeed, the literature on attitudinal research, and educational research in general, consistently reiterates the desirability of research built on the principle of convergent validity using multiple measurement to provide better inference. The surplus meaning involved in the definition of attitude also provides a favorable valence toward multiple operationism.

The following section organizes some of the more accessible research on the development of attitudinal measures in reading. The small number of studies dealing with instrument development is indicative of the low activity to date in this area.

Although we have shelves full of writing proclaiming the desirability of developing positive attitudes towards books and reading and, to somewhat lesser extent, materials to use in attitude change in reading, with some exceptions few adequately standardized tools for measuring change in attitude have been developed.
Research reporting development of reading attitude measures can be organized under three headings: Self report instruments, direct observation and projective techniques.

**Self-reports as behavioral specimens**

In self report research behavioral measures of attitude are collected by the researcher confronting the respondent directly and eliciting responses about the focal object being studied (usually a stated proposition of some sort). The group administered inventory instrument, checked by the respondent, is perhaps the most widely used approach. It is also possible to utilize individually administered researcher checked (standard interviews), and individually administered respondent checked (mailed questionnaire) instruments.

The San Diego County Inventory of Reading Attitudes (1961) presents respondents with 25 questions about reading which are answered by a yes-no forced choice technique. The 25 items were selected from an original pool of 114 questions as being those which most effectively selected high and low attitude students in grades 1-6. No information is given on the source of the original item pool, the item discrimination techniques used, or the original subject sample. For norming purposes, the attitude score is the raw score converted to a stanine score. The theoretical range of scores is 0-25 raw scores (1-9 stanines). Higher scores indicate more positive attitudes with an average score being 16-18 raw score points (stanine 5). The norm group consisted of 757 elementary subjects from four schools in San Diego county. An odd/even split half reliability of .79 and a Spearman-Brown Prophecy coefficient of .89 were obtained. Significant differences were also obtained on scale scores of high and low attitude readers selected by teachers thus establishing validity by confirmation of independent teacher judgment. This instrument has received wide use -- perhaps because it has been one of the few
Kennedy and Halinski (1975) report development of a 70-item Reading Attitude Inventory based on a four point Likert (1970) technique of scale construction. The original behavioral specimens were collected from students in several secondary schools asked to respond to a number of open-ended statements. Student response terminology was used in developing a preliminary 90 item Likert inventory which was then administered to 500 secondary students. Factor analysis and item test correlations were used to pare the inventory to 70 items (data not reported). No information on scoring is provided, but it is assumed that the higher the score, the more positive the attitude. The revised instrument was administered to 977 secondary students with some students identifying themselves. All students provided standard demographic data such as sex, grade in English academic track, etc. In each signed section, teachers identified 3 students with the most negative and positive attitudes. ANOVA was used to determine internal consistency reliabilities with subgroups coefficients in excess of .90 and total group reliability reaching .95. An odd/even split half reliability of .93 was also reported. Significant differences were obtained on scale scores of positive and negative readers as selected by teachers. Females scored significantly higher than males, A Students significantly higher than B students, and B students significantly higher than lower grades. Similar trends, although not as strong, were evident in the scoring of tracked students. Signed and unsigned students did not differ significantly and the results of a 14 item "lie scale" provides additional evidence which suggests that generation of a "socially acceptable" response set was minimized. See also: (Fiddler, 1974).

Shirley (1969) queried 420 sophomore, junior and senior students in two high schools on introspective and retrospective changes on their concepts, attitudes and behavior experienced as a result of reading. Responses were
classified into four areas: influence on concept, attitude, behavior and combinatory. Self involvement was also classified. Data were gathered through a written questionnaire with the classification process validated through analysis of selected protocols by judges. Some validity data for the technique were obtained through supplementation of the self reports by teacher interviews and individual case studies. Positive influences correlated with amount of material read and other relationships were in the predicted direction providing modest validation for the critical incident approach.

Three attitude inventories, based on the Thurstone technique of scale construction (1970 a, b), were developed by Johnson and Jacobson (1968) to assess attitudes toward literal interpretation of Underdog, Anthropomorphic, and Culturally Alien thematic archetypes in reading selections. A questionnaire was used with 2,000 intermediate grade subjects to collect original source opinions. Seven judges sorted 600 opinions reflecting attitude toward each of the three themes into three groups of seven categories each spaced along a continuum of favorableness producing three attitude inventories (inventories not included -- no scoring data provided). Validity is implied, as is common with Thurstone scales, through judging procedures in item selection and scale construction and item content. Kuder-Richardson reliabilities of .82, .81, and .85, respectively, were reported for the three scales.

A modified semantic differential (Snider & Osgood, 1968) technique, couched in a semi-projective self-report approach, was utilized by Greenberg et al (1965) to investigate attitude of deprived Negro grade four subjects toward 13 concepts thought to be important for school learning (Best Friend, Myself, Smart Child, Mother, Father, Teacher, School, T.V., Reading, Homework, Playing, Arithmetic, Dumb Child). Eight adjective pairs were employed -- 2 potency and 6 evaluative -- to rate the 13 concepts. The 6 evaluative pairs
were selected from those that past research has indicated leaned heavily on the evaluative factor (attitude). The instrument was modified to a 3 point-scale to accommodate young subjects. Vocabulary was also adjusted. Pretesting with individuals was conducted but no data were reported. Five fourth grade (115) classrooms were used. Each subject had an evaluative and potency score for each of the 13 concepts. A positive attitude was scored plus, a negative attitude minus, and neutral 0. The possible range of evaluative scores was +1 to -6 (potency were multiplied by 3 to equate with evaluative scores). More positive scores indicated more positive attitudes. No norms were provided. No reliability or validity data were reported. However, the reliability of semantic differential studies has been established in general and ample evidence is available to support the notion that the evaluative factor is a valid indicator of attitudes and can be used as a generalized attitude measurement (Osgood, Suci, & Tannenbaum, 1970) (Heise, 1970). See also: (Kemper, 1970).

Gurney (1966) described development of a local Reading Attitude Survey, involving forced choice between 39 pairs of stimulus propositions consisting of activities frequently engaged in by children and reading statements. The pairings were tested with a pilot group of sixth graders half of whom were adjudged by their teachers to like reading and half to dislike reading. Some validity evidence can be deduced from the fact that differences were noted between the groups - those judged by teachers to like reading chose reading propositions more frequently.

Preliminary data on the norming of a Reading Attitude Inventory, consisting of self response forced choice between 36 pairs of propositions contrasting interest in reading with other interests, were reported by Heimberger (1970). Items are presented orally, organized into four categories: Recreational Reading, Work-Type Reading, Learning to Read, Social Values --
with subjects responding in writing. No data are given on source of original items or item discrimination techniques used in developing the inventory. Based on another sample, reliability is reported as being .74 but no sample description or method of computation are provided. 1969 norm group data, based on 1,093 second, third and fourth graders in 4 Western Pennsylvania counties, are presented in percentile equivalents. No scoring instructions are provided, it is assumed the range of scores is, theoretically, 0-36 -- the higher the score the more positive the attitude.

Dulin and Chester (1975) report an exploratory study which utilized paired comparisons of contradictory propositions about reading to determine if reading attitude operates on judgmental processes similar to those undergirding racial and social attitudes. Thirty paired pro and con arguments were developed coupled with an 11 point scale for each item ranging from -5 to 0 to +5. Five adjectival statements for each scale ranged from "very ineffective" to "highly effective". All negative statements were scored 1 to 11 in terms of negative to positive judgment, and all positive items were scored in reverse order. A high score was hypothesized to be related to positive attitude toward reading. No information was given on the construction of the 30 propositions or the item discrimination techniques used in developing the scale. Subjects consisted of 130 eleventh grade students from a Northern Wisconsin Secondary School. Respondents also checked three self-rating categories -- liking for reading, amount read, and the degree to which they valued or respected reading -- on a five point scale. Teacher ratings, on a similar scale, were also obtained. No reliability data were reported for the 30 propositions. Plausibility judgment scores did correlate significantly (.20 & .17) with one Self-Rating scale (liking for reading) and one Teacher Judgment scale (liking for reading) providing some validity evidence. Also, mean plausibility scores and respondent self-ratings on liking for reading did proceed in the
Bulleh (1970) (1972) explored factors involved in motivating children to read and reported development of the Bullen Reading Attitude Measure for use in grades 1-6. The paired comparison inventory consists of a primary pictorial form, with 6 subtests measured through 15 pairs of propositions, and an intermediate form using one or two word alternatives for 8 subtests measured through 24 propositional pairs. The attitude score is the summation of the number of times reading is preferred to other activities. The pairs contrast reading with equally desirable alternatives across subtests involving: home, school, desire to visit the library, desire to receive books at home, desire to purchase books as presents. The norming study used 291 subjects from 12 classes in grades 1-6 from Fall River, Massachusetts Schools. Test-retest reliability coefficients (3 week interval) are reported by grade and range from .46 to .82 for primary and .75 to .82 for intermediate. Spearman Brown split half reliabilities ranged from .67 to .82 and .86 to .90. A respondent self report of reading activity (25 question interview), a parental report of reading activities (8 item questionnaire), and comparison with teacher selection of 5 positive and 5 negative attitude students provide some validity evidence. On the respondent self report, all variables thought to be related to reading attitude correlated significantly at one or more grade levels. Correspondence between respondent answers and self reported reading activities, and parental reports and respondent answers, provide validity evidence. Teacher selection and attitude score was in the expected direction for younger students, but not for upper grades.

The scales consist of 15 Likert type items in five content areas -- English, mathematics, reading, social studies, and science developed for middle school through senior high school subjects. The preliminary item pool of several hundred statements was based on a search of the related literature and statements from teachers and students which they suggested as discriminating positive and negative attitudes towards school subjects. Content validity of each statement was assessed and items retained which possessed face validity and measured a broad range of content in each subject area. Scores for the 15 item reading scale are summed, totalling between 15 and 75, with high scores representing positive attitude. The *Manual for Administration and Interpretation* provides extensive data on scale construction, scoring and interpretation. Reliability is substantial and impressive data for content, factorial, convergent and divergent validity are also presented. Estes and Johnstone (1974) also present validity data using the multitrait-multimethod approach (Campbell & Fiske, 1959). The Estes Attitude Scales are well developed, within the state-of-the-art, and contain many interesting and useful features.

**Direct observation**

Relatively few studies utilizing direct observations as behavioral specimens in attitude research have been reported. Contrived and natural settings are utilized in some instances but major problems in reliability and validity exist in obtaining large numbers of observations for measuring purposes.

Healy (1965) assessed the relationship between initial reading experiences and changes in reading attitudes with two classes of fifth grade subjects. Attitude change was assessed by a time-sampling technique; recording reading behavior, competency in finding information, use of free time, and a questionnaire. An initial teacher observational judgment was made as to whether subjects liked reading, disliked reading or were neutral to it. The same judgment was made
at the end of the year. Number of books read was also used as an attitude indicator. Attitude changes affected the amount read and level of achievement and persisted into junior high school.

Rowell (1972) developed a Likert type 5 point observation checklist, A Scale of Reading Attitude Based on Behavior, for use with elementary subjects. The scale is scored 1 to 5 for each of 16 items and summed for a total attitude score. Adjectival labels range from "always occurs" to "never occurs." Original items were developed to reflect how children feel toward reading including reading for pleasure, reading in content areas, and reading in class. No data are presented on the original item pool or the item discrimination techniques used. Reliability and validity were determined with student teachers and supervisors who had contact with subjects over a prolonged period. Student teachers completed the scale on each child and supervisors on a randomly selected one-third of the class. Correlations (40 subjects) of .95, .91, .76, and .89, with an average coefficient of .88, provide satisfactory reliability. As a validity check, student teachers and teachers rated attitude to identify positive and negative attitude students. Correlations between ratings and reading attitude scores were .80, .84, .52, and .63 with an average of .70. Individual item scrutiny for 40 subjects also indicated 79 to 100 percent agreement between observers.

Projective techniques

In projective techniques, "the investigator interprets the response in terms of dimensions and categories different from those held in mind by the respondent while answering." The measure can be structured or unstructured but usually involves no criterion of success. Campbell et al (1964) provides a well known review of a large number of projective studies and Kidder and
Campbell (1970) evaluate structured indirect measures. Considerable controversy attends use and interpretation of projective methods. The reading research utilizing this approach in measuring attitudes usually involves considerable modification of the more familiar (TAT, Rorschach) projective techniques used in personality assessment.

Boning and Boning (1957) describe a list of 42 incomplete sentence stems, the Incomplete Sentence Projective Test, useful in eliciting attitudes toward reading. Supposedly, teacher observations confirmed consistency between written responses and actions.

The work of Lipsky (1971), using pictorial stimuli to uncover covert attitudes or feelings about reading, is an interesting approach. He utilized the School Apperception Method (SAM), developed by Solomon, Klein & Starr (1966), as a model for 9 ambiguously drawn pictures depicting the influence of four factors on fifth grade boys' attitudes toward reading: the home, the school, peers, and cultural values. Subjects made up a story, telling how the story began and ended, and what the people in the story were thinking and feeling. 194 statements were extracted from 10 high achieving and 10 low achieving readers. Five judges agreed on the covert positive and negative reading content on 165 out of 194 attitude statements. A second testing four weeks later yielded a reliability coefficient of .89. Item discrimination procedures reduced the final scale to 109 items. A second sample of 118 subjects, high and low achieving readers, were exposed to the scale with the SAM discriminating significantly between the two groups. The original interview technique offers some validity support as well.

Askov (1969) developed the Primary Pupil Reading Attitude Inventory to measure attitudes toward recreational reading. 20 second and third grade subjects were interviewed individually to determine preferred recreational
activities. Nine of the most frequent were depicted by an artist and three reading pictures added. Scales were developed for both sexes. Pairing resulted in 27 picture propositions allowing subjects to choose between reading and some other activity. Inclusion of 13 distractors resulted in a 40-item scale. Three second and third grade classrooms (73 subjects) comprised the sample. Test-retest reliability after one week produced a coefficient of .90. Some validity accrued due to the original interview technique. Using a different sample of 94 students, teachers selected 5 with high and low attitude toward reading. Comparisons revealed significant differences on inventory scores between the two groups. Very low correlations with achievement data were also obtained indicating that the instrument was tapping areas other than academic.

Lowery & Grafft (1968) developed a variation of the Projective Test of Attitudes (Lowery, 1966) to measure the relationship between supplemental use of paperback books and change in attitude toward reading. The PTOA interweaves three projective techniques. The Word Association Test, used to bypass defense of respondents, presents "book" and "word" randomly among a listing of neutral words. The subject is instructed to respond as rapidly as possible with the first three words that come to mind. The Thematic Apperception Test utilizes drawings of a child sitting at a desk with a book to elicit reading attitudes. The Sentence Completion Test asks respondents to finish incomplete sentence stems loaded to elicit negative and positive responses about reading. Scoring is by positive, negative or neutral response. A descriptive analysis was made with independent judges placing responses in five descriptive categories. Interrater reliability among judges was not estimated. Reliability, measured by consistency of response from pre to post test, was not calculated although Lowery (1966) reports good interrater
agreement and high reliability for a similar technique in another study. A change of 20 percent or more between pre and post tests was established as the criterion for attitude change. Students showed significant positive change in attitude as a result of exposure to paperbacks. The PTOA was judged to be a sensitive approach in detecting reading attitude change.

In two studies not explicitly concerned with attitude change in reading, Zimmerman and Allebrand (1965) estimated attitude toward academic achievement of good and poor readers using the TAT while Hake (1967) used a Reading Apperception Test (ten ambiguous reading oriented pictures) to estimate covert motivation of good and poor readers.

General Criteria for Attitude Scale Development

In reviewing reported reading attitude research one does not have to delve deeply to find inadequacies in scale construction and validation (although authors are not overly prone to point out limitations in their research!) The reader often has to deduce for himself, usually from a scanty report, that problems with restricted samples, response set bias, item analysis, and reliability and validity mar what could otherwise have been a useful piece of research. Difficulties lie both with the conducting and reporting of research. The reader (and the researcher to some extent) has to keep in mind certain generally accepted minimal criteria which serve as a bare-bones "psychometric report card" in reading and assessing research. The criteria presented here are aimed at the more general user rather than the laboratory researcher. The authors referenced in the first section of this review can be turned to for in-depth treatment of the topics presented. The criteria are divided into three sections following the rough chronology that takes place in scale construction. The purpose of the exercise in attitude measurement is to derive an adequate (reliable and valid) estimate of attitude without altering or destroying the
sought after attitude or creating another one in the process. This presents no small measuring problem and the criteria presented here are designed to further this end.

**Item development**

Once the image of the concept to be measured is formed (attitude definition) its relevant dimensions can be specified and decisions made as to the indicators which best represent the concept. Three factors are most important in item development:

1. Proper sampling of items to insure that the universe of content is well represented.
2. Carefully worded items couched in language that is easily recognized and understood by the respondent.
3. Pre-testing using appropriate item analytic procedures to eliminate or revise unsatisfactory items.

**Response set contamination**

Response set refers to the tendency on the part of subjects to contaminate responses by responding to attitude statements for reasons other than the content of the items. To minimize response set:

1. Make the scale interesting and pleasant.
2. Control acquiescence by occasionally switching response alternatives.
3. Control selection of socially desirable responses through forced choice answers; by pretesting items for socially desirable loadings; through use of a parallel social desirability scale.
4. Control faking of responses through use of "lie scales" cross checks in item wording, and "disguised" techniques in item presentation.

**Statistical procedures**

Several basic statistical requirements should be met in constructing scales. Few scales rate optimally on all criteria and poor performance on any one does not necessarily invalidate a scale. Reasonable performance should be expected on most of the criteria.
The sample should be representative of the group(s) studied.

Basic normative information should be reported.

Reliability coefficients should be included.

Data on homogeneity of test items should be provided including inter item, item test, and total test characteristics.

Predictive, content and construct validity should be established and relevant information reported.

Some Concluding Comments

A number of general comments can be offered on the basis of this review.

1. Information on the development and use of reading attitude instruments is difficult to locate and obtain being widely scattered in journal articles, books, technical reports, papers read at meetings, dissertations and, undoubtedly, in undisturbed piles of reports in the offices of researchers. Most published reports exhibit shortcomings in the amount and type of information they present.

2. The volume of information available on the assessment of reading attitudes is scanty in light of the considerable importance given to the relationship between positive attitude and reading achievement in the literature.

3. Quality instrument development and standardization, with some few exceptions, is rare. Scale construction methods can range from rigorous attempts to develop scales with well known characteristics to approaches that are highly subjective with few discernable criteria. Reading attitude scale construction often violates the basic rationale underlying the method chosen. Items appearing in many scales are often at variance with the scale construction technique used. Few adequate pilot studies on scale construction are reported.

4. Researchers tend to develop measures de novo when good existing techniques could be adapted, or to accept commonly used scales uncritically thus perpetuating the Matthew effect. Too little information is available on previous work and too few reviews which attempt to synthesize and cumulate research in the field are available.

5. Collection of behavioral specimens through self-report paper and pencil techniques dominates the methodology of attitude study in reading as in other areas. However, some attention has been given to other indirect and projective approaches. Little in the way of a conceptual framework holds research efforts together in anything
but an accidental pattern. The research typology represented in
the multiple indicator approach (argued by Summers (1970), Cook
and Sellitiz (1964) and many others) has yet to have any discernable
effect on attitude research in reading. Similar conclusions can
also be reached with respect to the notions of indirect, disguised
and unobtrusive measures put forth by Campbell (1970), Webb et al
(1966), and other writers. A recent paper by Greene and Zirkel
(1976) reviews literature on instrumentation of attitude measurement
in elementary reading. The authors opt for a multi-measure approach
emphasizing complementary, not duplicative, sources of data.
Observer report instruments, book counts, choice among self-report
approaches, and pictorial projective and activity preference
techniques are suggested as possible fruitful elements in a multi-
method approach to instrumentation.

When scale reliabilities are reported, test/retest, split-half and
alternate forms approaches are used in that order. Few studies
report more than one reliability check. Split-half reliabilities
are often reported as measures of temporal consistency. Item
discriminatory techniques utilizing intercorrelation tables and
factor and cluster analysis techniques are infrequently reported.

Establishing validity using predictive, content, concurrent and
construct indices is a persistent problem in reading attitude
research. In many instances, the content of items is the only
discernable quasi-validity information. Known groups techniques
figure prominently in validity estimates. Predictive validity
data seldom appear perhaps reflecting the state-of-the-art in
relating reading attitude to other psychological constructs.
With few exceptions, firmly established construct validity is
unique. One notable good exemplar is the recently published
instrument by Estes et al (1975) which is, technically and con-
ceptually, the best developed reading attitude scale to date.
Reliability and internal consistency data are strong. Evidence
touching on several kinds of validity is offered and convergent
validity results, using the Campbell and Fiske (1967) multitrait-
multimethod approach, are reported.

Adequate conceptualization and definition of reading attitude
constructs is still a major problem in research in the field.
More application of attitude theory from social psychology could
provide useful guidelines. Defining and measuring psychological
constructs is not always an easy matter. Measurement too often
predominates at the expense of conceptual frameworks. All the
statistics in the world are no substitute for good, hardnosed
preliminary thinking about what it is that is to be measured and
how it can best be operationalized and quantified.
More attention is being given to development of affective domain objectives in education. Measurement and evaluation specialists increasingly emphasize the "differential outcome hypothesis" in structuring and evaluating educational objectives schema (Glaser, 1972), (Ebel, 1973), (Tuckman, 1974 & 1975). Categories of outcome often include attitudes and values as affective objectives. However, measurement of such variables is still admittedly primitive. In their analysis of the teaching of affective responses, Khan and Weiss (1973) conclude that research on affective behavior is still in its infancy. A beginning has been made, but instrumentation and quantification procedures are proving to be more complex than they have been in the study of cognitive variables.

Attitude research will add an important dimension to the study of affective functioning. Attitudes constitute a significant source of behavioral variance and could serve to integrate and explain a wide range of behavior. In fifty years attitude research in social psychology has moved from a focus on measurement problems to research on the dynamics of attitude change, structure of attitude systems, functioning of attitudes within the total personality, and the mechanisms undergirding attitude change. A similar pattern will evolve in education and the study of reading attitudes, in particular, could play a significant role in such research. In their study of attitude toward school subjects, Estes et al (1975) point out:

The present authors believe that the composite evidence (correlations of attitude scale scores with self ratings, peer nominations, teacher ranking, course grades, achievement and extracurricular) .... provides a sound case for the convergent validity of the five Estes Attitude Scales. The data also suggest that reading attitude is the most centrally located variable in the nomothetic network of validating criteria. In fact, with the exception of self ratings of mathematics and peer nominations of science attitudes, reading attitude is found to be significantly correlated with all criterion measures. (p.19)

This offers rather interesting statistical validation for the centrality of
reading and reading processes in education.

In spite of what may appear to be glowing opportunities to break new frontiers in affective and attitude research as it relates to reading, a final word of hard realism should be rendered. Using Kuhn's classification of scientific behavior (The Structure of Scientific Revolution, 1962) one could place educational research somewhere between natural history (organized experience) and normal science (testing theories and applications). Expectations of early conceptual and measurement eurekas in attitudinal research would indeed be premature. It will continue to be "Little Science, soft science" in this area for some time to come.
REFERENCES


Ebel, R.L. The future of the measurement of abilities II. Educational Researcher, 1973, 2, 5-12.


San Diego County, California Department of Education. An Inventory of Reading Attitude. San Diego, California: Superintendent of Schools, Department of Education, 1961. ERIC ED 028 893.


Tuckman, B.W. The differentiated outcome hypothesis or when will we stop using conventional achievement as the only criterion for evaluation of conventional instruction? Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, April, 1974.

Tuckman, Bruce W. The differentiated outcome hypothesis. *JEDR Quarterly*, 1975, 9, 8-11.


ADDITIONAL SOURCES

The following listing includes additional sources of interest in three areas: (I) Measurement and the Affective Domain, (II) Attitudinal Measurement in Education, and III Attitudinal Measurement in Reading.

I Measurement and the Affective Domain

General Discussion


Inventories and Bibliographies of Affective Measures


Boric, Gary D. Handbook of Evaluation Instruments for Preparing Educational Personnel for the Handicapped. ERIC ED 088 908.


II Attitudinal Measurement in Education

Attitude Toward School


Finch, John Fall, 1974 South Carolina Student Survey and Attitude Inventory, May 1975. ERIC ED 111 876.


Lasseigne, Mary W. Assessing attitudes of junior high school students. Contemporary Education, 1974, 45, 139-141.


Attitude Measurement and Instrument Development


Harvill, Leo M. Evaluation of Several Methods for Measuring Young Children's Educational Attitudes. Vermillion, South Dakota, University of South Dakota, 1971. ERIC ED 056 059.


May, Frank B. An improved taxonomical instrument for attitude measurement. College Student Services, 1969, 2, 33-38.


Attitude Measurement in Subject Areas and in Relation to Curriculum Innovations


Estes, Thomas H. Measurement of students' attitudes toward what they study. In Harold H. Herber, & Peter L. Sanders (Eds.), Research in Reading in the Content Areas. Syracuse, New York: Reading and Language Arts Center, 1969.


### III Attitudinal Measurement in Reading


